BAŞKENT ÜNİVERSİTESİ

TEMEL İSTATİSTİK YÖNTEMLERİ

PROF. DR. SERPİL CULA PROF. DR. ZEHRA MULUK
BAŞKENT ÜNİVERSİTESİ TİCARİ BİLİMLER FAKÜLTESİ

III. Baskı

BAŞKENT ÜNİVERSİTESİ 2018

ISBN 978-605-68615-2-9

BAŞKENT ÜNİVERSİTESİ
Bağlıca Köyü
Eskişehir Yolu 20 km
06530 Ankara
Tel: 0 312 234 10 10

BİRİNCİ BASKININ ÖNSÖZÜ

İstatistik, doğru bilgi edinmek ve bu bilgiyi doğru sunmak için vazgeçilmez bir bilim dalıdır. Günümüzün rekabetçi ortamında verilerden yararlanıp, gerekli bilgiyi elde etmek ve bu bilgiler doğrultusunda politika üretmek, iş çevreleri için kaçınılmaz bir durumdur. Çağımızda eleman alımlarında bilgiyi doğru alıp yorumlayabilen kişiler giderek daha çok tercih edilmektedir. Yeterli istatistik bilgisinin alınması bu nedenle büyük önem kazanmaktadır. Nitekim üniversitelerimizin bir çok bölümünde istatistik zorunlu bir temel ders olarak okutulmaktadır. Bu kitabı hazırlamaktaki amacımız da özellikle Başkent Üniversitesi Ticari Bilimler Fakültesinin tüm bölümlerinde temel istatistik dersini almakta olan öğrencilere bu konuda beceri kazandırmak ve katkı sağlamaktır.

Kitap oniki bölümden oluşmaktadır. Her konuda çok çeşitli örneklerle, uygulamaya ağırlıklı olarak yer verilmiştir. Buna ek olarak, öğrencilerin edindikleri bilgileri pekiştirebilmeleri için çeşitli problemler bölüm sonlarında sunulmuştur.

Kitabın hazırlanmasında bilgi ve katkılarından yararlandığımız değerli bilim adamlarına teşekkür borçluyuz; özellikle ülkemizdeki ilk istatistik bölümünü kuran ve bizlere katkısı büyük olan Sayın hocamız Prof. Dr. Alaettin Kutsal'a başta olmak üzere, uzun yıllar birlikte çalıştığımız değerli hocamız ve arkadaşımız Sayın Doç. Dr. Tülay Saraçbaşı’na ve titiz inceleme sonucu yapmış olduğu önerilerle bu çalışmaya katkıda bulunan Sayın Prof. Dr. İsmail Erdem'e en içten teşekkürlerimizi sunarız.

Akademik yaşamımızın her aşamasında olduğu gibi bu çalışmamızda da desteklerini cömertçe veren ailelerimizin tüm bireylerine teşekkür borcumuz büyüktür.

Mayıs 2006

Prof. Dr. Zehra Muluk-Yrd. Doç. Dr. Serpil Cula

İKİNCİ BASKI İÇİN ÖNSÖZ

Birinci baskının önsözünde belirtildiği gibi bu kitabın amacı istatistik dersini alan öğrencilerin istatistiği daha kolay anlamalarına katkı sağlamaktır.

Birinci baskının kullanımında hedef kitlemizden gelen geri bildirim, konu sonu problemlerinin cevaplarının verilmesi şeklinde oldu. Bu istek doğrultusunda konu sonu problemlerinin bazıları için sonuçlar kitabın son bölümünde verildi. Birinci baskının kullanımı sırasında bazı problemlerin değiştirilmesi uygun bulundu. Bu değişiklikler ve konu içinde uygun görülen düzeltmeler bu baskıda yapıldı.

Bu baskı ile kitabımızın daha geliştiği ve yararlı olduğu görüşündeyiz. Saygılarımızla...

Ağustos 2010

Prof. Dr. Zehra Muluk-Yrd. Doç. Dr. Serpil Cula

İÇİNDEKİLER

Sayfa

BİRİNCİ BÖLÜM

İKİNCİ BÖLÜM

ÜÇÜNCÜ BÖLÜM

DÖRDÜNCÜ BÖLÜM

BEŞİNCİ BÖLÜM

ALTINCI BÖLÜM

YEDİNCİ BÖLÜM

SEKİZİNCİ BÖLÜM

DOKUZUNCU BÖLÜM

ONUNCU BÖLÜM

ONBİRİNCİ BÖLÜM

ONİKİNCİ BÖLÜM

1. BİRİNCİ BÖLÜM

1.1. GİRİŞ

1.2. İSTATİSTİĞİN TEMEL KAVRAMLARI

1.3. BAZI ÖRNEKLEME YÖNTEMLERİ

1.4. BİLGİ ELDE ETME YOLLARI

1.5. İSTATİSTİKSEL ÇIKARSAMA ARAÇLARI

PROBLEMLER

1.1. GİRİŞ

Bu bölümde istatistiğin tanımlanması, gereğinin kavranması, kullanılacak bazı terimlerin açıklamalarının yapılması amaçlanmıştır. Bölüm içinde aşağıda verilen sorulara cevaplar aranacaktır.

İstatistik nedir? Neden istatistiğe gerek duyarız? Kitle nedir? Denek (örneklem birimi) nedir? Değişken nedir? Veri nedir? Kaç türlü değişken vardır? Her zaman kitleye ulaşmak mümkün müdür? Örneklem nedir? Örneklem nasıl elde edilir? Örnekleme yöntemleri nelerdir? Örneklemden nasıl yararlanılır? Ölçme teknikleri nelerdir? Bilgi elde etme yolları nelerdir? Deneysel araştırma nedir? Saha araştırması nedir?

İstatistik nedir? Neden istatistiğe gerek duyarız?

Kuruluşlar, dünyada gittikçe artan rekabeti karşılamak için gereken bilgi, beceri ve gerekli yeteneklere sahip insanları bulmak için, üniversite ve yüksekokul mezunlarını istihdam etmektedir. Ancak bu mezunlarda aranan özellik yalnız yüksek eğitimli olmaları değil, iş dünyasının karar verme araçlarını anlama, uygulama ve yorumlama yeteneğine sahip olmalarıdır. İstatistik, kişiye yukarıda belirtilen süreç için gerekli bilgileri veren bir bilim dalıdır.

Bir çok kuruluşta karar verme aracı olarak geniş veri tabanları bulunmaktadır. Ancak karar vericiler bu verileri etkin olarak kullanma konusunda zorlanmaktadırlar. Uygulamalı istatistik, öğrencilere veri kümelerini kullanabilmek ve etkin bilgiye dönüştürebilmek için gereken araçları sunar. Bilimsel anlamda karar verme, bilgilere dayanmalıdır. Örneğin yeni bir ürününü piyasaya sürecek bir satıcı, satış yapacağı kitleyi tanımak ve satabileceği nitelikte bir ürün üretmek isteyecektir. Bunun için hedef kitleyi tanımlamalıdır. Tanımlanan kitleden bilgi toplanmalıdır. Bilgi tüm kişilerle tek tek görüşülerek mi toplanacaktır. Bu mümkün olmayabilir. O halde topluluğu temsil edebilecek bir alt grup alınabilir. Toplanan bilgiler nasıl ölçülmelidir ki satıcıyı doğru karar vermeye yöneltsin. Bu ve buna benzer soruların bilimsel olarak cevaplanabilmesi için istatistiğe gerek vardır. İstatistiğin yüksek öğretiminin hemen tüm dallarında gerekli bir ders olarak okutulmasının nedeni budur.

İstatistik sözcüğü, rakamla elde edilen bilgilerin belli kurallarla anlaşılır ve yorumlanabilir duruma getirilmesidir. Örneğin:

- Türkiye'de bir araştırma firmasının yaptığı araştırmaya göre Kanal X'in seyredilme oranı % 65'dir.

- Bir şirketin kalite sorumlusu müşteri memnuniyet oranının bir önceki yıla göre %1 gerilediğini belirledi.
- Bir turizm firması 2005'de düzenlediği turlardan 4 milyon dolardan fazla kazandı.

Yukarıda verilen örneklerin hepsi birer istatistiktir. Çoğu kişi bu tür istatistikleri ders almadan da bilir. Bu istatistikler belirli olayları tanımlayan rakamlardır. Her gün okuduğunuz gazetede borsadaki fiyatlar, suç oranları, hükümet organlarının bütçeleri hakkında tanımlayıcı istatistiksel bilgiler verilir. Böylesi tanımlayıcılara her yerde rastlayabiliriz. İstatistik bilimi karar vermeye yardımcı olacak çeşitli yöntemler içerir.

1.2. İSTATİSTİĞİN TEMEL KAVRAMLARI

İstatistik biliminin bazı temel kavramları vardır. Bu kavramlardan bazıları aşağıda açıklanacaktır.

Kitle

Araştırıcının ilgilendiği ortak özellikleri taşıyan birimlerin oluşturduğu topluluğun tümüne kitle (population) denir. Başkent Üniversitesinde okuyan öğrencilerin eğitim sorunlarını araştıracağımız bir çalışmada, Başkent Üniversitesinde araştırma döneminde kayıtlı olan tüm öğrencilerin oluşturduğu topluluk o çalışma için kitledir. İç Anadolu bölgesinin sosyal ve ekonomik yapısının incelendiği bir araştırmada, İç Anadolu da yaşayan tüm ailelerin oluşturduğu topluluk kitledir. Bir sigorta şirketinde kaza sigortası yaptırıp hasar tazminatı alan müşterilerle ilgili bir araştırmada sigortalı araçlardan kaza yapıp sigortadan masrafları alanların tümü; ihracat yapan bir firmanın belli bir yıl için ihracat kalemleri ile ilgili bir araştırmada bir yıllık tüm ihraç kalemleri; Ankara'da yaşayan ailelerde yapılacak bir gelir araştırmasında, Ankara'da yaşayan ailelerin tümü kitleye örnektir. Kitle sonlu ya da kavramsal olabilir. Sonlu kitlede sınırlar bellidir. Örneğin, ihraç edilmek için kontrol edilecek 100000 makarna kolisinde kitle sonludur. Ancak, makarna üreten bir fabrikada üretim hattından bilgi almak istenmiş ve zaman kısıtı koyulmamış ise kitle kavramsal olacaktır.

Parametre

Kitleye ait bilgilere parametre denir. Kitle büyüklüğü (N), kitlede incelenen değişkene ait ortalama (μ), varyans (σ^2) birer parametredir.

Örneklem

Kitleye ulaşmanın mümkün olmadığı durumlarda kitleden çekilmiş, kitleyi tanımlayabilecek, kitlenin küçültülmüş bir kopyası kullanılır ki buna **örneklem** (sample) denir. Örneklemden elde edilen bilgiler kitleyi tanımak için kullanılır. Örneklemden elde edilen bilgilere **istatistik** denir. Araştırmalarda örneklem ile çalışmak ekonomik ve daha kısa zamanda sonuç almamızı sağlar. Ancak istatistiklerin doğru ve güvenilir olması için kitleden örneklem seçim işleminin doğru yapılması gerekir. Doğru seçilmemiş bir örneklem hatalı bilgi verecektir. Örneklemin doğru seçilmesi için geliştirilmiş yöntemler vardır, bu yöntemlere **örnekleme yöntemleri** denir. Örnekleme yöntemlerinden bazılarına Bölüm 1.3'de kısaca değinilecektir.

Denek

Kitlede ya da örneklemde yer alan her bireye ya da birime denek (subject) denir. Örneğin Başkent Üniversitesinde öğrencilerle ilgili bir araştırmada her bir öğrenci bir denektir. İç Anadolu bölgesinde yaşayan ailelerde sağlık koşullarını konu alan bir araştırmada her bir aile bir denektir.

Değişken

Kitlede ya da örneklemde deneklerin nitelik ya da nicelik belirten özelliklerinin her biri değişkendir. Örneğin, bir şirketteki müşteri profilinin araştırıldığı bir çalışmada, çalışmaya alınan her müşteri bir denek, bu deneklerden alınan cinsiyet, yaş, aylık alışveriş tutarı, A ürününde tercih ettiği marka birer değişkendir. Değişkenler yapı bakımından nicel ve nitel olmak üzere ikiye ayrılır. Nicel değişken nicelik gösterir. Örneğin; yaş, bir cismin ağırlığı, bir şirketin kar oranları, belli bir ürünün ihracat miktarı, bir ailedeki kişi sayısı birer nicel değişkendir. Nitel değişken nitelik gösterir. Örneğin cinsiyet, milliyet, din, dil, ırk, bir ihracattaki ürün çeşitleri vb. değişkenler nitel değişkenlerdir.

Veri

İstatistiksel tekniklerin kullanılabilmesi için öncelikle ilgilenilen konuyu aydınlatacak değişken ya da değişkenlere ait bilgi toplanması gerekir. Birimlerden bilgi toplama işlemi çok önemli bir konudur. Araştırma sonunda kitleden ya da örneklemden elde edilen sayısal veya niteliksel bilgiler **veri** olarak tanımlanır. Veriler yapılarına göre sürekli ve kesikli olabilir.

Ölçme

Değişkenlerin aldığı değerleri belirlemek için ölçme yapılması gerekir. Nitekim, 19. yüzyıl bilim adamlarından Lord Kelvin, "Üzerinde konuştuğunuz şeyi ölçebilir ve sayısal olarak ifade edebilirseniz o şey hakkında bir şey biliyorsunuz demektir. Ama bunu yapamıyorsanız bilginiz hem yetersiz hem de istenilen nitelikte değildir. Bildiğinizi belki bilgi başlangıcı sayabilirsiniz. Fakat ne olursa olsun sizin düşünce düzeyinizde bilimsel aşamaya ulaştığınız pek söylenemez" diyerek ölçmenin önemini vurgulamıştır.

Ölçümler değişkenin özelliğine göre;

1- Sınıflanabilir (gruplanabilir, nominal)
2- Sıralanabilir (ordinal)
3- Aralıklı (interval)
4- Orantılı (ratio)

olmak üzere dört ana başlık altında toplanabilir.

Sınıflanabilir ölçeklerde ölçme, bireylerin belli bir özellik itibarıyla aynı olup olmamasına dayanır. Bir grup insanın; sigara içip içmediğine göre, cinsiyete göre, milliyete göre, sevdiği spor dallarına göre sınıflanması bu ölçme yöntemine örnek olarak verilebilir.

Sıralanabilir ölçüm türünde temel özellik sıralamadır. Burada ayrılığın yanı sıra daha üstün ya da daha düşük kavramları kullanılarak sıralama yapılır. Spor karşılaşmalarında uygulanan dereceler, ürünlerde iyiden kötüye ya da tersine yapılan sıralamalar, bir tütün eksperinin verdiği dereceler, idari personelin unvanlarına göre sıralanması vb. bu ölçme yöntemine ilişkin örneklerdir.

Aralıklı ölçümün temel özelliği bir başlangıç noktasının ve bir bitim noktasının olmasıdır. Sınavlarda kullanılan puan sistemi, üniversite giriş sınavlarında kullanılan puanlar örnek olarak verilebilir.

Orantılı ölçü türünün temel özelliği her zaman kabul edilen değişmez bir başlangıç noktasının bulunmasıdır. Bu ölçüm türünde birimler de standarttır. Uzunluk ölçüleri, ağırlık ölçüleri, hacım ölçüleri, etki süresi vb. örnek olarak verilebilir.

Nitel yapıdaki değişkenlere ait veriler sınıflanabilir ya da sıralanabilir ölçeklerle, nicel yapıdaki değişkenlere ait veriler ise aralıklı ya da orantılı ölçeklerle ölçülürler.

Tam Sayım

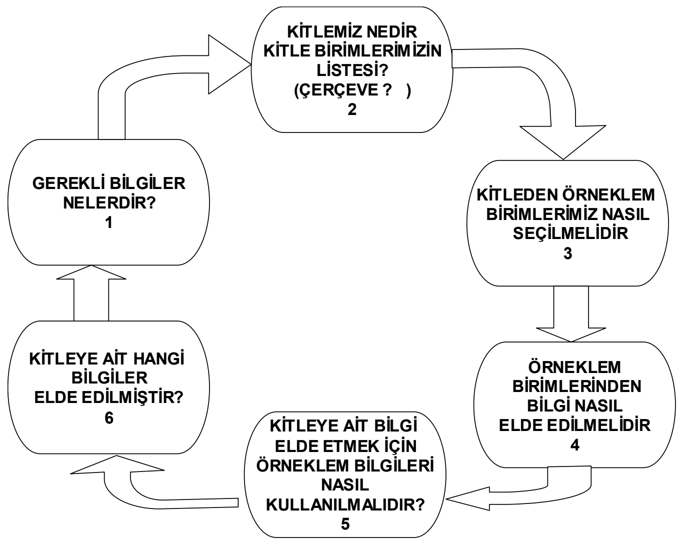
Kitlenin tüm deneklerinden bilgi toplanmış ise **tam sayım** yapılmıştır. Aşağıda verilen durumlar dışında tam sayım çok sık kullanılmaz.

a- Kitle çok küçük ve kitleye ait ayrıntılı bilgi elde etmek isteniyor ise kullanılır. Örneğin, bir şirket kendisi ile devamlı alışveriş yapan bayileri hakkında bilgi toplamak istiyorsa tam sayımı kullanması uygun olur.

b- Bazı durumlarda aynı kitleden değişik zamanlarda küçük örneklemler seçerek araştırma yapma gereksinimi duyulabilir. Bu durumda kitleye ait parametreleri belirlemek için, bir defaya mahsus olmak üzere tam sayım yapılabilir. Ülkemizde yapılan nüfus sayımı bir tam sayımdır.

Bilgi Elde Etme Süreci

Bölüm 1.1'de belirtildiği gibi istatistik çeşitli nedenlerle var olan bilgiyi elde etme ve anlamlandırma bilimidir. Bu sürece araştırma diyebiliriz. Araştırma yapma gerekçesi, çeşitli nedenlerle gerekli olan bilgi ihtiyacıdır. Bu bilgiyi elde etme süreci Şekil 1.1'de verilmiştir.



Şekil 1.1 Örnekleme Kullanılarak Bilgi Elde Etme Süreci

Şekil 1.1'de değinilen ilk adım bilgi elde etmek için gereksinim duymaktır. İkinci adım çerçeveyi oluşturma, üçüncü adım örneklemin seçimidir. Dördüncü kısımda kitleye ait bilgi elde etmek için örneklem bilgilerinin nasıl kullanılacağı, beşinci kısımda da bilgi elde etme yolları üzerinde durulacaktır. Bu adımlar aşağıda ele alınacaktır.

1.3. BAZI ÖRNEKLEME YÖNTEMLERİ

Örneklemenin doğru yapılabilmesi kitlenin doğru tanımlanabilmesine bağlıdır. Kitleyi tanımlayan bilgiye çerçeve denir. Örneğin; Bir okulda öğrenciler ile ilgili bir araştırma yapılacak ise tüm öğrenciler kitledir. Bu kitleden örneklem seçip o küçük topluluk ile çalışılacak ise bu küçük topluluk örneklemdir. Örneklemi doğru seçebilmek için kullanılan bilgi, öğrencilerin okul numaralarından oluşan bir veri tabanı olabilir ki bu bilgiye de çerçeve denir. Bir banka müşterileri ile ilgili bir araştırma yapılacak ise tüm banka hesap numaralarından oluşan veri tabanı çerçevedir. Her zaman çerçeveyi oluşturmak kolay olmaz. Bazı durumlarda çerçeve tanımlanamaz. Eğer bir şarap fabrikasına üzüm alımı yapılıyor ve kamyonlardan seçilen bir miktar örnek üzerinde kalite kontrolü uygulanacak ise tüm kamyondaki üzüm salkımlarının numaralanması mümkün olamaz. Bir marketler zinciri, müşterilerinin tümü üzerinde bir araştırma yapmayı planlamış ise burada da çerçeveyi oluşturmak mümkün değildir. Çerçevenin tanımlanabilmesi ya da çerçevenin oluşturulamamasına bağlı olarak örnekleme yöntemleri olasılığa bağlı olan ve olasılığa bağlı olmayan (keyfi) örnekleme olarak ikiye ayrılır. Olasılığa bağlı olmayan örnekleme araştırıcının sübjektif kararlarına bağlı olarak kitleden seçilen bir gruptur. Genelde tercih edilmez. Bu tür örneklemede parametreleri iyi tahmin etmek olanaksızdır. Ancak kitlenin tanımlanamadığı durumlarda, başka yol olmayacağı için kullanılan bir yöntemdir.

1.3.1. Olasılığa bağlı olmayan örnekleme yöntemleri

Keyfi örnekleme, dilim örneklemesi ve kota örneklemesi olasılığa bağlı olmayan örnekleme türlerindendir.

Keyfi örnekleme; N tane birim içinden, n tanesinin araştırıcı tarafından tamamen kişisel görüş ve tecrübelerine dayanarak seçilmesidir.

Dilim örneklemesi; Kitleyi meydana getiren birimler çok geniş bir alana yayılmış ise, örneği oluşturmak üzere tüm coğrafya alanlarından birim seçilmesi, fazla para ve zaman harcaması gerektirir. İmkanlar buna elverişli değilse, kitleyi tüm özellikleri ile temsil edebilen küçük bir parça örneği meydana getirebilir. Bir ülkeyi temsil etmek üzere bir şehrin seçilmesi dilim örneklemesi olarak düşünülebilir.

Kota örneklemesi; Kitle, ele alınan nitelik bakımından farklı kısımlardan oluşuyorsa, örneklemin kitleyi temsil gücünü artırmak için her kısımdan önemi ile orantılı olarak keyfi örnekleme ile birim seçilmek suretiyle yapılan örneklemedir.

1.3.2. Olasılığa bağlı örnekleme yöntemleri

Olasılığa bağlı örnekleme yöntemleri, kitledeki her birime belirli ve sıfırdan farklı bir olasılıkla örneğe girme şansı verir. Olasılığa bağlı örnekleme yöntemleri kitlenin yapısına göre değişir.

Basit rasgele örnekleme; Eğer kitle tekdüze ise (homojen) yani kitlede bulunan denekler birbirlerinden çok farklı bir yapı göstermiyorlarsa, olasılığa bağlı örnekleme yöntemlerinden olan basit rasgele örnekleme ya da sistematik örnekleme yöntemleri kullanılır. Basit rasgele örnekleme ile sistematik örnekleme yönteminin kullanılma yerleri farklılık gösterir.

Sistematik örnekleme; Bu örnekleme genelde kitlede seri olarak numaralama ya da adreslemenin olabildiği durumlar için uygundur. Örneğin; bir kalite mühendisi günlük üretimi 10000 olan bir ürünün günlük kontrolleri için 10 ürünü incelemek isterse, bu 10 ürünün seçimi için sistematik örneklemenin kullanılması uygun olur.

Tabakalı örnekleme; Kitle bütün olarak tekdüze yapıya sahip değil, ancak kendi içinde tekdüze olan alt parçalara ayrılabiliyorsa tabakalı (katmanlı) örnekleme kullanılır. Tabakalı örneklemede tabakaların doğru oluşturulması önemlidir. Her tabaka kendi içinde homojen olmalıdır. Seçilecek örneklemin büyüklüğü, tabakaların büyüklüğüne orantılı olarak dağıtılır. Her tabakada bulunan deneklerin o tabaka için örnekleme çıkma şansları eşit olmalıdır.

Olasılığa bağlı örnekleme istatistikte önemli bir yer tutar. Bu konuda daha ayrıntılı bilgi için kaynaklarda verilen örnekleme kitaplarına başvurulabilir.

1.4. BİLGİ ELDE ETME YOLLARI

Bilgi elde etme sürecinde dördüncü adım, karar vermek için gerekli olan bilgiyi elde etmek deneysel araştırmalar ve saha araştırmaları olmak üzere iki ana başlık altında toplanabilir. Deneysel araştırma ile bilgi elde etme olayında araştırıcı şartları oluşturur. Oluşturulan şartlar altında, incelenen değişkenin davranışı ile ilgili bilgi toplanır. Saha araştırmalarında ise durumu belirlemek için bilgi toplanır. Bu tür çalışmalarda araştırıcı olaya müdahale edemez, gerçekte var olan durumu öğrenmeye çalışır.

Deneysel araştırma ile bilgi elde etme

Araştırmalarda bağımlı ve bağımsız değişkenler vardır. Bağımsız değişkenlerin bağımlı değişkeni etkilediği düşünülür. Deneysel araştırmada araştırıcı deneyi yaparak bağımsız değişkenlerin çeşitli düzeylerini kullanarak en iyi sonucu elde etmeye çalışır. Bu nedenle araştırma araştırıcı tarafından kurgulanır. Deneysel çalışmalarda istatistiğin kullanılması 1920 yılında bugün kullandığımız anlamda R.A. Fisher ile başlamıştır. Günümüzde endüstride üretilen ürünün kalitesini yükseltmek, tarımda en iyi ürünü elde edebilmek, tıpta tedavi yöntemlerini karşılaştırmak için deneysel araştırmalar kullanılmaktadır.

Saha araştırmaları ile bilgi elde etme

Durum tespiti yapmak, geçmiş bilgilerden yararlanarak gelecek için kestirim yaparak karar vermek, saha araştırmaları ile gerçekleştirilir. Bu tür araştırmalarda eski kayıtlar kullanılarak, gözlem yolu ile, anket çalışması uygulanarak deneklerden bilgi toplanır. Anket çalışmaları da telefon anketi, posta yoluyla gönderilen anket ve araştırıcının kişi ile yüz yüze yaptığı anketler olarak ayrılmaktadır.

a- Telefon anketleri

Diğer anket türlerine göre daha ucuz ve etkili bir veri toplama aracıdır. Seçim yıllarında, siyasal gruplar seçim kampanyalarına ilişkin konularda ve adaylarının ne durumda olduğunu anlamak için anketler düzenlerler. Bu anketlerin kısa bir süre içinde yapılmaları gerektiği için, genellikle telefonla arama yöntemi kullanılır. Telefon görüşmelerinde hatta fazla kalınamayacağı için görüşme genellikle 1 ile 3 dakika arasında bitecek şekilde hazırlanmalıdır. Şüphesiz, çeşitli nedenlerle bilgi alamadığımız kişiler olacaktır.

Telefon anketi için uygulanan basamaklar aşağıda verilmiştir:

1- Konu tanımlanmalıdır: Anketin ne için yapıldığı açıkça belirtilmeli, amaçlar tanımlanmalıdır. Araştırma için belirli bir bütçe ayrılmalıdır.

2- Kitle tanımlanmalıdır: Kitle o konuda bilgi sahibi olan kişilerin tümüdür. Anket yalnızca bu grup arasında yapılmalıdır.

3- Anket soruları saptanmalıdır: Anketi kısa tutmak için soru sayısı sınırlandırılmalı, önemli sorular önce sorulmalıdır. Mümkünse, belirli yanıt seçenekleri verilmelidir.

4- Uygun bir anket hazırlanmalıdır: Anketin sonunda demografik sorular sorulmalıdır. Demografik sorular yanıtlayanın karakteristiklerine, öğrenimine ve öz niteliklerine ilişkin sorulardır (yaş, gelir, vs.). Anketin amacı giriş bölümünde açık bir biçimde ifade edilmelidir. Anketi cevaplayan kişilerin kimliklerinin gizli kalacağı vurgulanmalıdır.

5- Anket bir ön denemeye tabi tutulmalıdır: Anket kitle içinden çok küçük bir grup üzerinde denenmelidir. Anketin uzunluğu, soruların açıklığı, kolaylığı kontrol edilmelidir. Ankette eksiklik olup olmadığı incelenmeli, gerekiyorsa değişiklik yapılmalıdır.

6- Örneklem büyüklüğü ve örnekleme yöntemi saptanmalıdır: Örneklem büyüklüğü sonuçlarımıza ne kadar güvenmek istediğimize, sonuçların ne derece doğru olmasını istediğimize, maliyete ve kitlenin üyeleri arasındaki görüş farklılıklarının ne kadar olduğuna bağlıdır. İstatistiğin örnekleme alanında örneklem büyüklüğünün nasıl hesaplanacağı anlatılmaktadır. Bu konu hakkında bilgi Bölüm 7.4'de verilecektir.

7- Örneklem seçilmeli, telefon numaraları elde edilmelidir: Telefon numaraları bilgisayarda oluşturulmuş ya da geçerli bir listeden alınmış olabilir. Yanıt vermeyen telefonlar için bir geri arama kuralı geliştirilmelidir. Telefon eden kişiler açık ve dürüst olacak şekilde eğitilmeli, anketi yanıtlayan kişiyi yönlendirmemelidir. Alınan cevaplar veri sayfasına kaydedilmelidir.

b- Yazılı Anketler

İnsanların görüşünü öğrenmek ve gerçek verileri toplamak için en sık kullanılan yöntem yazılı anket yapmaktır. Yazılı anket bazen potansiyel yanıtlayıcıya doğrudan verilir, bu tür çalışmalara yüz yüze görüşme ile yapılan anket çalışmaları denir. Bazen de uygulama posta ile yapılır. Anket postaya veriliyorsa, bazı masraflar vardır. Bunlar yanıtlayana gönderilen ve yanıtlayanın gönderdiği posta masrafları, anket geliştirme ve baskı masrafları ile veri analizleri için gerekli masraflardır.

Yazılı anketler daha ayrıntılı olabilir ve bu nedenle de tamamlanmaları telefon anketlerinden daha uzun sürebilir. Burada, çok zaman istemeyen ve kolayca tamamlanabilecek bir anket düzenlemeye dikkat edilmelidir. Yazılı bir ankette hem açık hem de kapalı uçlu sorulara yer verilebilir. Yanıtlayanın, tanımlanmış seçeneklerden oluşan kısa bir listeden seçme yapmasını gerektiren sorular kapalı uçlu sorulardır. Örneğin, "Hangi siyasal partiyi tutuyorsunuz? Cumhuriyet Halk Partisi mi? Doğru Yol Partisi mi? DSP mi? AKP mi? Yoksa başka bir parti mi?" Bu tür sorular kapalı uçlu sorulardır. Yanıtlayanların kendi seçtikleri diğer sözcük ya da beyanlar ile yanıtladıkları sorular açık uçlu sorulardır. Açık uçlu sorular, yanıtlayana daha fazla esneklik sağlar; fakat yanıtların analizi daha zor olabilir.

Yazılı anketlerde yapılması gerekenler telefon anketlerinde anlatılanlar gibidir. Yazılı anket sonunda demografik sorular sorulmalıdır. Girişte anketin amacı açıklanmalıdır. Kişilerin kimliklerinin gizli kalacağını vurgulanmalıdır. Anketin düzeni açık ve çekici olmalıdır. Yanıtlar için yeterli yer bırakılmalıdır. Anketin amacını açıklayan bir kapak sayfası eklenmelidir. Anket posta yolu ile uygulanacak ise anketin geri gönderilmesi için bir zarf eklemekte yarar vardır. Posta ile gönderilen anketlerde, düşük bir yanıtlama oranı beklenebilir; çoğunlukla %5 ile %20 arasında. Dolayısıyla 200 yanıt isteniyorsa, 1000 ile 4000 arasında anket postalamak gerekmektedir.

Yanıtlayanın doğru ve güvenilir bilgiler vermesini sağlamak için, yazılı anketlerin biçimlendirilmesine dikkat edilmelidir. Yazılı yanıtlar için yeterli boşluğun bırakılması gerekir. Anketin nasıl doldurulacağına ilişkin talimatlar açık olmalıdır. Yazılı bir anketin göze hoş görünmesi gerekir. Anketin görünüşü yanıtlama hızını etkileyeceği için uzman kişilerce hazırlanmalıdır.

Yazılı anketlere verilmiş yanıtlar elle mi girilecek yoksa optik okuyucu ile mi taranacağına karar verilmelidir. Uygulanan yaklaşım anketin tasarımını etkileyecektir. Çok sayıda anket dağıtılmış ise tarama tercih edilmelidir. Bu yöntem ile veri bilgisayar ortamına girilirken daha az hata yapılacak ve veri toplama süreci hızlanacaktır. Ancak tarama yöntemi seçilirse, olası yanıt biçimleri sınırlı olacaktır. Anket, yanıtlayıcılara doğrudan veriliyorsa, yüksek bir yanıtlama oranı beklenebilir. Örneğin, büyük olasılıkla üniversitede, dönem sonunda bir ders değerlendirme formu doldurulması istenebilir. Konuya ilişkin bir anket formu aşağıda örnek olarak verilmiştir.

ÖĞRETİM SORUMLUSU VE DERS DEĞERLENDİRME FORMU

Açıklama: Bu anketin amacı, öğrencilerimizin öğretim sorumluları ve dersler hakkındaki görüşlerinin alınmasına yöneliktir. Sonuçların yararlı olması için objektif biçimde yanıtlamanız önem taşımaktadır. Anketi isim belirtmeyiniz. Değerlendirmenizi optik form üzerinden yapınız (En düşük not=1, En yüksek not=5).

1. Öğretim sorumlusu derse ait beklenti ve hedeflerini açıkça ifade etti.
2. Açık ve anlatılır bir anlatım tarzına sahipti.
3. Karmaşık kavramları çok iyi açıkladı.
4. Derse zamanında girip çıktı.
5. Ders konularına hakimdi.
6. Derslere hazırlıklı geldi.
7. Ders araç gereçlerini etkin kullandı.
8. Derste sınıf içi tartışmalara ağırlık verdi.
9. Sınıf/laboratuar ortamına hakimdi.
10. Öğrencilerle etkileşimi olumluydu.
11. Ders kitabı çok uygun olarak seçilmişti.
12. Sınav soruları derste işlenen konularla ilgiliydi.
13. Sınav süreleri yeterliydi.
14. Sınav ve projeleri adil bir şekilde değerlendirdi.
15. Sınav sorularını görme ve soru sorma olanağı tanıdı.
16. Bu ders diğer derslere oranla çok zordu. (Çok kolaydı=1,...,Çok zordu=5)

Yukarıda verilen değerlendirme formu dönem içerisinde ve dönem sonunda öğrencilerin öğretim üyesini değerlendirmek için kullandıkları bir örnek formdur. Ancak bu form tamamen bir örnek olup geliştirilmesi bazı soruların eklenip, bazı soruların çıkartılması kullananların deneyimlerine göre gerçekleştirilebilir.

Günümüzde internet yaygın olarak kullanılmakta ve yazılı anketler internet aracılığıyla deneklere ulaştırılmaktadır. İnternet aracılığıyla hazırlanacak olan anketlerin elektronik ortamın özelliklerine uygun olarak hazırlanması gerekir.

c- Doğrudan gözlem ile veri toplamak

Çok sık kullanılan bir başka bilgi elde etme yoludur. Bu yöntem, veri toplama sürecinde deneklerin fiziksel olarak gözlenmesini ve verilerin kaydedilmesini gerektirir. Gözlem yaparken karşımıza çıkacak ana kısıtlamalar gözlemi yapmak için gereken süre ve paradır. Gözlemlerin etkili olabilmesi için, eğitimli gözleyiciler kullanılmalıdır ki bu da masrafı artırır. Kişisel gözlem zaman alır. Kişisel algılamalar özneldir. Farklı gözleyicilerin, bir durumu aynı şekilde göreceğinin ve aynı şekilde bildireceğinin hiçbir garantisi yoktur. Kişisel görüşmeler amaçlara bağlı olarak yapılandırılmış (soruların yazılı olduğu görüşmeler) ya da yapılandırılmamış (genel birkaç soruyla başlayıp, verilen yanıtlara göre başka soruların sorulduğu görüşmeler) olabilir. Bu yöntemde açık ya da kapalı uçlu sorular kullanılabilir.

Veri Toplamaya İlişkin Sorunlar

Verilerin toplanmasında dikkat edilmesi gereken bazı noktalar vardır. Karar vermeye yardımcı olmak üzere veriye ihtiyaç duyuluyorsa, öncelikle dolaylı yoldan elde edilen verilerin uygun olup olmadığı araştırılmalıdır. Bunun nedeni, dolaylı yoldan elde edilen verilerin toplanmasının doğrudan elde edilen verilere kıyasla daha hızlı ve ucuz olmasıdır. Bununla birlikte, dolaylı yoldan elde edilen verilere güvenilmeden önce, bu verilerin doğru toplanıp kaydedildiğinden emin olabilmek için kaynaklar kontrol edilmelidir. Bilinmeyen kaynaklardan gelen bilgilerden kaygı duyulabilir. Bu, web üzerinden edinilen bilgiler için de geçerlidir. Veri toplamayla ilgili başka genel sorunlar da vardır. Bunlardan biri de, veri toplamanın yanlı olabilmesidir. Örneğin, kişisel görüşmede, görüşmeyi yapan kişi soru sorma şekliyle, ses tonuyla ya da üzerinde konuşulan konuya bakışıyla önyargısını isteyerek veya istemeyerek hissettirebilir. Bir anketle veri toplama sürecine müdahale edebilecek bir başka yanlılık kaynağı da yanıtsız yanlılıktır. Daha önce belirtildiği gibi, postayla gönderilen anketlerin sorunu büyük bir yüzdenin geri dönmemesidir. Telefonla yapılan aramalarda da bazen kişiler bulunamaz ya da yanıtlamayı reddedebilir. Kişisel görüşmelerin denekleri de görüşme yapmayı reddedebilir. Yanıt alamamayla ilgili bir sorun vardır. Yanıt verenlerin sağladıkları veriler, veri toplama için denek seçiminde de yanlılık olarak kendini gösterebilir. Buna seçim yanlılığı denir. Üniversitedeki atletizm ücretlerinin artırılmasının yararları konusunda bir çalışma yapılıyorsa, verileri futbol maçı seyreden öğrencilerden toplamak amaca hizmet etmeyecektir. Akşam saatlerinde bir telefon anketi yapılıyorsa, gece çalışan insanlar ankete dahil edilmiyor demektir. Gece çalışanlarla gündüz çalışanların görüşleri, gelir ve öğrenim düzeyleri, vs aynı değilse, veriler yanlı demektir. Ayrıca, görüşülen kişiler yalan söylüyorsa, telefon anketleri ve kişisel görüşmelerle elde edilen veriler hatalı olacaktır.

Veri toplamada karşılaşılan bir başka sorun da ölçme hatalarıdır. İyi eğitilmemiş iki teknisyenin vereceği ölçümler aynı nesne için birbirinden çok farklı olabilir.

Kişisel gözlem yoluyla veri toplamak da problemler yaratabilir. Kişiler aynı olayı ya da nesneyi farklı görme eğilimindedirler. Buna gözlemci yanlılığı denir.

1.5. İSTATİKSEL ÇIKARSAMA ARAÇLARI

Bilgi elde etme sürecinde beşinci adım istatistiksel çıkarsamalardır. Karar vericinin kitlenin bir alt kümesini temel alarak, kitle hakkında bir yargıya varmasını sağlayan araçlardır. İstatistiksel çıkarsama araçlarında iki ana konu vardır. Tahmin ve hipotez testi. Tahmin ve hipotez testi için kullanılan yöntemler birbiriyle yakın ilişkili fakat farklı amaçlara hizmet etmektedirler.

Tahmin

Tahmin (kestirim), kitle hakkında bilgi elde etmek istendiğinde kitleye ulaşmanın zor ya da olanaksız olduğu durumlarda örneklemden yararlanarak bilgi elde etme yoludur. Bu durumda örneklemin doğru seçilmiş ve uygun bir tahmin yönteminin kabul edilmiş olması gerekir. Aşağıda verilen örnekler konuyu daha iyi açıklayacaktır.

Örnek 1.1: Kanal tercihi örneği

Televizyon kanalları bir önceki sene hangi magazin programını kaç kişinin izlediğini kesin olarak bilemezler. Tüm izleyicilere sormak da kolay değildir. Bunun yerine, bir araştırma şirketinin verdiği programların seyredilme sıklığı bilgilerine güvenmeleri gerekir. Bu bilgiler, araştırmanın yapıldığı dönem boyunca programları açan izleyicilerin sayısına ilişkin tahminlerdir. Bu çalışma tüm ülkeden seçilen az sayıda hane temel alınarak hesaplanır. Hesaplanan sayı tüm kitle için bir tahmindir. Kitle için TV programlarının seyredilme sıklığı bilgisi olarak kullanırlar. Bu bilgiler özellikle reklamcılar tarafından TV kanalları ile yapılan sözleşmelerde yararlı olmaktadır. Ancak bu bilginin doğru olabilmesi için seçilmiş olan hanelerin tüm kitleyi tanımlaması şarttır. Eğer seçilen haneler belli bir gelir grubuna ait ise bu küçük grubun tüm ülkeyi tanımladığı söylenemez. Böyle bir örneklemden yapılan tahminler de hatalı olur.

Örnek 1.2: Sanayi sayımı örneği

Ülke düzeyinde uzun vadeli ekonomik kararlar verilirken sanayi sektöründe olan değişimler, istihdam, yatırım vb. bilgilerin bilinmesi gerekir. Ancak tüm iş yerlerini kapsayan bir araştırma yapmak da mümkün değildir. Bu bilgiler belli aralıklarla yapılan örnekleme çalışmaları ile elde edilir. Örneğin ülkemizde veri toplama görevini yasal olarak üstlenen T.C. Başbakanlık Türkiye İstatistik Kurumu (TÜİK) sanayi sayımı ve ülke çapında daha başka sayımları, örneklem yöntemini kullanarak yapmaktadır.

Hipotez Testi

Televizyon reklamlarında çeşitli ürünler için bir çok iddialar vardır. Örneğin, "Goodyear lastikleri en azından 60000 mil dayanır" ya da "Doktorlar Bayer Aspirini diğer markalardan daha çok tercih ediyor". "General Elektrik ampulleri diğer markalardan daha uzun süre dayanır" ya da "Müşteriler ABC deterjanını diğer deterjanlara tercih ediyorlar". Bu iddiaların boş sözler mi, yoksa gerçek verilere mi dayandığı araştırılabilir. Tüketici raporu gibi, tüketici araştırma kuruluşları bu türden iddiaları düzenli olarak kontrol etmektedirler. Örneğin deterjan örneğinde, tüketici raporu, iki marka arasında müşteri tercihleri bakımından fark olmadığı hipoteziyle yola çıkarak, bir grup müşteriyi seçip ABC deterjanını ve diğer deterjanları marka bilmeden kullanmalarını isteyebilir. Örnek veriler sonucunda, tercihler arasında önemli bir farklılık görülüyorsa, müşteri tercihleri arasında fark olmadığı hipotezi reddedilir. Yalnızca küçük bir farklılık görülürse, tüketici raporu yazarları hipotezi reddedemezler. Bu konu Bölüm 8'de verilecek olan hipotez testleri bölümde ayrıntılı olarak ele alınacaktır.

PROBLEMLER

1- Aşağıdaki sorulardan doğru olanları (D) ile, yanlış olanları (Y) ile işaretleyiniz.

a- Kitle özelliklerinin rakamla belirtilen değerine parametre denir.

b- Bir değişkenin değeri denekten deneğe değişmez.

c- Homojen birimlerden oluşan kitlede tabakalı örnekleme kullanılır.

d- Bir grup çocuğun boy uzunluğuna ait veriler, sürekli nicel veri özelliği gösterir.

e- Ortalaması 195 olan kitleden, 58 denekli örneklem çekilmiştir. Örneklemin varyansı 21, ortalaması 198'dir. Verilen bilgilerden 21 istatistiktir.

f- Bir araştırmada ailenin eğitim durumuna göre sınıflama yapılmaktadır. Buradaki ölçüm türü nominaldir.

2- 2000 sayfalık 3 ciltlik bir ansiklopedide yanlış yazılan sözcüklerin incelenmesi için tüm 2000 sayfa içerisinden 100 sayfa rasgele seçilmiştir. Her sayfadaki yanlış sözcük sayıları sayılarak sonuçlar listelenmiştir.

a- Kitle nedir?

b- Örneklem nedir?

c- Hangi tür örnekle yöntemi kullanılmıştır?

d- Buradaki değişken nedir? Hangi tür değişkendir?

3- Yaş ortalaması 14 olan hükümlü çocukların yeniden toplumlaşma davranışlarını incelemek için bir araştırma yapılmıştır. Bunun için Ankara Kalaba ve İzmir Şirinyer ıslah evleri çalışma alanı olarak seçilmiştir. Ankara Kalaba'da 1025, İzmir Şirinyer'de 980 çocuk barınmaktadır. Bu çocuklar içinden işledikleri suç çeşidine göre 5 grup oluşturulmuştur. Suç çeşidindeki çocukların sayısı ve farklı ıslah evleri dikkate alınarak toplam 216 çocuk seçilip anket uygulanmıştır. Anketteki sorulardan bazıları:

Yaşı, ıslah evine gelmeden önce nerede yaşadığı (köy, kaza, kent içi, gecekondu), evden kaçıp kaçmadığı, kaç kez kaçtığı, kardeş sayısı, vb. Bu örneklemde ortalama kardeş sayısı 6±1,3 olarak bulunmuştur.

a- Bu çalışmada kitle nedir?

b- Yukarıda verilen açıklamada parametre değeri verilmiş midir? Verilmiş ise kaçtır?

c- Yukarıda verilen açıklamada değişkenleri, bu değişkenlerin türlerini yazınız.

d- Çalışmada kullanılan örneklem türü nedir?

e- Yaş için kullanılan ölçüm türü nedir?

f- Yukarıda verilen soruda örneklem değeri verilmiş midir? Örneklem değeri nedir?

4- Çevre kirliliğinin yoğun olduğu bir bölgede toplam 20000 kişi yaşamaktadır. Bölge kent içi ve gecekondu olmak üzere ikiye ayrılmıştır. Her iki bölgeden, bölge nüfuslarına orantılı olarak 1000 kişi seçilmiştir. Seçilen kişilere anket uygulanmıştır. Anketteki bazı sorular; eğitiminiz, evinizde TV var mı? Çevrenizi kirli buluyor musunuz? Aylık geliriniz nedir? vb. Bilgilere göre aşağıdaki soruları cevaplayınız.

a- Bu araştırmanın kitlesi nedir?

b- Bu araştırmada örnekleme yöntemi nedir?

c- Bu araştırmada denekler nedir?

d- Yukarıdaki soruda kaç değişken verilmiştir?

e- Araştırmada veriler hangi yönteme göre toplanmıştır?

f- Eğitimin ölçüm türü nedir?

5- Şeker hastalığının nedenleri üzerine yapılan araştırmada cinsiyet, yaş, ağırlık ve daha önceki kuşaklarda hastalığın gözlenip gözlenmediği sorulmuştur. Bu araştırmadaki değişkenleri ve değişkenlerin özelliklerini belirtiniz.

6- Nüfusu 4 milyon olan bir yerleşim bölgesine yeni bir radyo istasyonu kurulacaktır. Bu bölgede yaşayan 12 yaş ve üstündeki kişilerin radyo yayınlarından beklentilerini belirlemek amacıyla bir araştırma düzenlenmiştir. Araştırma kapsamına giren kişilerden 4000'i rasgele seçilmiştir. Araştırma sorularından bazıları aşağıdadır:

a- Yaşınız:

b- Cinsiyetiniz:

c- Radyoyu hangi sıklıkla dinlersiniz:

(Her gün; Ara sıra; Hafta sonu; Hiç)

d- Evinizdeki radyo sayısı:

(0 1 2 3)

Aşağıdaki soruları yukarıdaki araştırmaya göre cevaplayınız.

a- Bu araştırmanın kitlesi nedir?

b- Bu araştırmada örnekleme yöntemi nedir?

c- Yaşın ölçüm türü nedir?

d- Cinsiyetin veri özelliği nedir?

e- c şıkkında sorulan sorunun veri özelliği nedir?

f- d şıkkında sorulan sorunun veri özelliği nedir?

7- 2005-2006 Güz döneminde B.Ü. Erkek öğrenci yurtlarına başvuran öğrenciler bilgi formu doldurmuştur. Doldurulan bu formlar fakültelere göre düzenlenmiştir. Her fakülteden 25'er öğrenci formu rasgele seçilerek toplam 150 öğrenci formu incelenmiştir. Formlarda: Ailenin aylık geliri; Kardeş sayısı; Doğum yeri; Diploma derecesi vb. sorular sorulmuştur.

a- Bu araştırmanın kitlesi nedir?

b- Bu araştırmada örnekleme yöntemi nedir?

c- Bu araştırmada değişken sayısı nedir?

d- Ailenin aylık gelirinin veri özelliği ve ölçüm türü nedir?

e- Kardeş sayısının veri özelliği nedir?

f- Doğum yerinin veri özelliği ve ölçüm türü nedir?

8- Bir araştırmada seçim yanlılığı nedir?

9- Kota örneklemesini hangi durumlarda kullanırız?

10- Dondurma üreten bir şirket, müşterilerinin çocuk sayılarını belirlemek için bir anket hazırlıyor. Hazırladığı sorulardan birinin cevap bölümünde çocuk sayısı için

1.....2....3....4 ve daha çok

olmak üzere hazırlanmıştır.

a- Bu soru için ölçüm türü nedir?

b- Bu bilgileri kullanarak ankete katılanların ortalama çocuk sayısını hesaplayabilir miyiz?

11- Bir araba üretim firması garanti sebebi ile firmanın karşılaştığı problemlerin çözümünde şikayet çeşitlerinin, bunların oranlarının, maliyetin, harcanan zamanın önemli olduğunu belirlemiştir. Bu bilgileri alabilmek için izlenecek yolu, bilgi almak için hazırlanacak formu düşününüz ve bir taslak form hazırlayınız.

12- Bir süpermarketler zinciri yöneticisi olarak, müşterinin markette yapılan değişikliklerden hoşlanıp hoşlanmadıklarını öğrenmek için, bir çalışma yapmak istiyorsunuz. Müşterilerin market içinde malların düzenlenişine, erişim kolaylığına, kasaların yerine ilişkin yaklaşımları incelenecektir.

a- Hangi örnekleme yöntemini kullanırsınız?

b- Bu bilgileri hangi bilgi elde etme yolunu kullanarak bulursunuz?

13- 12. problemde bir telefon anketi seçtiğinizi varsayınız.

a- Soracağınız soruların listesini hazırlayınız.

b- Verileri bilgisayarın okuyabileceği bir biçime nasıl dönüştüreceğiniz konusunda bir plan hazırlayınız.

c- Yanıt alamama durumunda izleyeceğiniz yöntem nedir?

14- Kırsal ve kentsel alanda yaşayan 0-5 yaş arası çocuğu olan kadınların bebek beslenmesi ile ilgili davranışlarını araştırmak için anket düzenlenmiştir. Anketteki sorulardan bazıları:

a- Evlenme yaşınız nedir?

b- Doğumdan ne kadar süre sonra anne sütü verdiniz? (hemen, 1 saat sonra, 1 gün sonra)
c- Anne sütünden önce başka besin verdiniz mi?
d- Bebeğiniz ek besinlerle birlikte anne sütünü toplam kaç ay aldı?

Araştırmadaki soruların veri ve ölçüm türlerini yazınız.

15- Bir özel televizyon kanalı TV müşterilerinin yeni bir spor programı isteyip istemedikleri konusunda bir anket yapmak için telefon ile anket yapma tekniğini uygulayacaktır. Böyle bir araştırma için tüm evreleri açık olarak belirterek planlamayı yapınız.

16- Bir beyaz eşya firmasının üretim müdürü olarak, firmanın karşı karşıya olduğu garanti iade problemlerin çözümünde kullanılmak üzere bir veri toplama süreci kurmak istiyorsunuz. Bir veri onay formu kullanmanız önerildi. Uygun bir veri onay formu geliştirin ve bu formun nasıl kullanılacağını açıklayıcı örnekler vererek gösteriniz.

17- 11. soru için aşağıdaki soruları cevaplayınız.

a- Toplayacağınız verilerin değişken türü nedir?
b- Veri toplama yönteminden hangilerini kullanırdınız?

18- Bir üniversitenin gıda hizmetleri yönetici yardımcısı olduğunuzu düşününüz. Üniversite kafeteryasında yemek kuyruğunda çeşitli müşteri davranışlarına ilişkin veri toplamanız istendi. Özellikle de öğle yemeklerinde seçimler ve istekler ile ilgilenmeniz istensin. Kafeterya her gün öğle yemeğinde 8 çeşit sunuyor. Müşterilerin kuyrukta yaptıkları yemek seçimlerine ilişkin veri toplamak için bir veri onay formu hazırlayınız.

19- Sipariş ile börek baklava imalatı yapan bir şirket, müşteri memnuniyetini belirlemek istiyor. Şirketin dört çeşit börek, beş çeşit baklava yaptığını düşünerek bir araştırma planlayınız. Bu araştırma için kitlenizi, çerçevenizi, örnekleme yöntemini, örneklem büyüklüğünü, kullanacağınız bilgi toplama formunu, bilgi toplama yöntemini belirterek araştırmanızı planlayınız.

20- Bir işyeri sınavına başvuran 300 kişiden rasgele 10 kişi seçilmiştir. Seçilen 10 kişi ile ilgili bazı değerler aşağıda verilmiştir.

Denekler	Aldığı puan	Cinsiyeti
1	38	E
2	48	E
3	56	K
4	58	K
5	62	K
6	65	E
7	68	E
8	68	E
9	71	K
10	79	K

a- Bu problemde kitleyi yazınız.

b- Hangi örneklem yöntemi kullanılmıştır?

c- Değişkenler nelerdir? Değişken türlerini ve veri yapısını yazınız.

d- Puanlar hangi ölçek türüne girer?

2. İKİNCİ BÖLÜM

2.1. GİRİŞ

2.2. SIKLIK DAĞILIMLARI

2.3. NİCEL DEĞİŞKENLERDEN ELDE EDİLEN SÜREKLİ VERİLERDE SIKLIK DAĞILIMLARI

2.4. KESİKLİ NİCEL VERİLERDE SIKLIK DAĞILIMLARI

2.5. NİTEL DEĞİŞKENLERDEN ELDE EDİLEN VERİLERDE SIKLIK DAĞILIMLARI

2.6. GRAFİKLER

2.7. FARKLI SINIFLANDIRMALAR

PROBLEMLER

2.1. GİRİŞ

İstatistik bilimi, karar verme sürecinde toplanan verilerden anlam çıkarmada kullanılan bir araçtır. Bu tanımın ilk kullanımı, toplanan bilgilerin özetlenmesi şeklinde olacaktır. Sonucun görsel olarak sergilenmesi işlemi de grafikler aracılığı ile yapılır. Bu bölümde sıklık dağılımları ve grafik çizim yöntemleri verilecektir.

2.2. SIKLIK DAĞILIMLARI

Sıklık dağılımları verilerin özetlenmesi işlemidir. Bu dağılımlarda, bir yığın halinde bulunan yüzlerce veri, uygun çoklukta sınıflara yerleştirilir. Bu sınıfların sıklıkları, yüzdelikleri verilerek yığın halinde bir anlamı olmayan verilerin anlaşılır olması sağlanır. Bu işlemin doğru yapılabilmesi için geliştirilmiş kurallar bu bölümde anlatılacaktır.

2.3. NİCEL DEĞİŞKENLERDEN ELDE EDİLEN SÜREKLİ VERİLERDE SIKLIK DAĞILIMLARI

Nicel değişkenlerden elde edilen veriler sürekli ve kesikli olarak ikiye ayrılmıştır. Öncelikle sürekli veriler üzerinde durulacaktır. Burada görülen temel kurallar, kesikli veriler için de kullanılacaktır. Ancak kesikli verilerde aşağıda tanımı verilecek olan sınıf değerlerinin tam sayı olmasına dikkat edilmesi gerekmektedir.

Sıklık dağılımını oluşturabilmek için aşağıda verilen bazı tanımların bilinmesi gerekmektedir.

Dağılım Sınırları: Bir veri topluluğunda (dağılımda) yer alan en küçük ve en büyük değerlerdir.

Dağılım Genişliği: Bir veri topluluğunda yer alan en küçük ve en büyük değerler arasındaki farktır.

Sınıf: Eşit ya da birbirine yakın değerli verilerin oluşturduğu gruptur. Bu gruplar birbirinden ayrık kümeler oluşturur.

Sınıf Sayısı: Eşit ya da birbirine yakın değerli verilerin oluşturduğu ayrık kümelerin sayısıdır. Sınıf sayısı “k” ile gösterilir. Bu sayının 7 ile 20 arasında olması tercih edilir. Dağılımdaki veri sayısı çok olduğunda, sınıf sayısı aşağıda verilen Sturges’in formülü kullanılarak bulunabilir:

$$k=1+3,3 \log n \tag{2.1}$$

Burada n, eldeki veri kümesindeki gözlem sayısını gösterir. k en yakın tam sayı olarak alınır.

Alt Sınır: Bir sınıfta yer alan en küçük değerdir.

Üst Sınır: Bir sınıfta yer alan en büyük değerdir.

Sınıf Aralığı: Ard arda gelen iki sınıfın alt sınırları ve üst sınırları aralarındaki fark olup, w ya da c ile gösterilir. Sınıf aralıkları eşit ya da farklı olabilir, eşit olduğunda aşağıda verildiği biçimde elde edilir:

$$w=c=(\text{en büyük değer}-\text{en küçük değer})/k \tag{2.2}$$

w veya c hesaplama sonucunda bulunan sayı veya daha büyük bir sayı olarak alınır.

Sınıf Değeri: Bir sınıfın alt ve üst sınırlarının ortalamasıdır ve S_i ile gösterilir.

Sıklık: Bir sınıfta yer alan gözlem sayısıdır, f_i ile gösterilir. Her bir sınıfa düşen sıklıkların toplamı, toplam gözlem sayısına eşittir.

$$\sum_{i=1}^{k} f_i = n \tag{2.3}$$

Burada f_i; i. sınıfın sıklığıdır.

Göreli Sıklık: Her bir sınıfa düşen gözlem sayısının toplam gözlem sayısına oranıdır. i. sınıfın göreli sıklığı (p_i) aşağıdaki eşitlik ile verilmiştir:

$$p_i = \frac{f_i}{n} \tag{2.4}$$

Sınıf göreli sıklıkların toplamı 1'e eşittir ($\sum_{i=1}^{k} p_i = 1$).

Birikimli Sıklık: Sınıf sıklıklarının üst üste eklenmesi ile bulunan sıklıklardır.

Göreli Birikimli Sıklık: Göreli birikimli sıklıklar, sınıf birikimli sıklıklarının toplam denek sayısına bölünmesiyle elde edilir.

2.3.1. Sürekli Nicel Verilerde I. Tür Sıklık Dağılımları

Örnek 2.1: Bir çiftlikteki 200 inekten 45 inek basit rasgele örnekleme ile seçilmiş ve bir günlük süt üretimleri kilogram olarak alınmış ve Tablo 2.1'de verilmiştir. Bu verilere ilişkin değerlerin sıklık dağılımını oluşturalım.

Tablo 2.1 İneklerin Kilogram Olarak Günlük Süt Üretim Değerleri

10,37	22,50	14,59	11,95	23,81
18,25	8,89	11,95	16,44	22,50
26,40	15,95	10,85	9,36	20,84
24,95	23,11	12,72	11,50	36,02
14,67	13,52	17,84	17,35	27,85
24,76	18,12	18,00	30,23	28,20
17,45	18,49	21,50	17,89	27,85
5,45	23,35	16,30	15,57	22,50
17,49	17,83	14,59	10,00	25,50

Sıklık dağılımını oluşturmadan önce, veriler küçükten büyüğe sıralanarak işlemlerde kolaylık sağlanabilir.

Tablo 2.2 İneklerin Kilogram Olarak Günlük Süt Üretimlerinin Sıranmış Değerleri

5,45	12,72	17,35	18,49	24,76
8,89	13,52	17,45	20,84	24,95
9,36	14,59	17,49	21,50	25,50
10,00	14,59	17,83	22,50	26,40
10,37	14,67	17,84	22,50	27,85
10,85	15,57	17,89	22,50	27,85
11,50	15,95	18,00	23,11	28,20
11,95	16,30	18,12	23,35	30,23
11,95	16,44	18,25	23,81	36,02

1. ADIM: Özetlemenin kaç sınıfta yapılacağı belirlenmelidir. Sınıf sayısı,

$$k=1+3,3 \log 45=6,45$$

olarak elde edilir. Sınıf sayısının tam sayı olması gerekir. Bu örnek problemde sınıf sayısı 6 veya 7 olarak alınabilir.

2. ADIM: Sınıf aralığı saptanır.

$$w = c = \frac{36,02 - 5,45}{7} = \frac{30,57}{7} = 4,36$$

Sınıflar oluşturulurken işlemlerde kolaylık sağlaması için, bu yukarı yuvarlanır. Örneğin aralık değeri 5 alınabilir.

3. ADIM: İlk sınıfın alt sınırı belirlenir. İlk sınıfın alt sınırı eldeki gözlemlerin en küçük değerine eşit veya daha küçük bir değer olmalıdır. Burada dikkat edilecek nokta, sınıfların ayrık kümeler oluşturması ve her gözlemin bir sınıfta yer almasıdır. Burada birinci sınıfın alt sınırı olarak 5,00 alınmıştır. İkinci sınıfın alt sınırı 5+w=5+5=10 olmalıdır. Bu şekilde 7 sınıf olacak biçimde alt sınırlar oluşturulur.

4. ADIM: Sürekli nicel verilerde, I. tür sıklık dağılımının birinci sınıfının üst sınırını oluşturmak için, ikinci sınıfın alt sınırı alınır ve birinci sınıfın üst sınırı olarak belirlenir. Böylece örneğe ilişkin birinci sınıfın üst sınırı 10 olacaktır, ancak sıklıklar belirlenirken 10, birinci sınıfa değil, ikinci sınıfa dahil edilir. Ardışık sınıfların üst-alt sınırları aynı olur. Ancak, sınıf aralıkları kapalı-açık veya açık-kapalı aralıklar şeklinde belirlenerek, sınıfların ayrık olması sağlanmalıdır.

5. ADIM: Her bir sınıfa düşen sıklık sayıları belirlenir ve sıklık kolonu oluşturulur.

Tablo 2.2'den yararlanarak birinci sınıfın sıklığı, 5,00-10,00 arasında bulunan 5,45; 8,89; 9,36 olan 3 değerdir. İkinci sınıfın sıklığı 10,00-15,00 arasında bulunan değerlerin sayısıdır ve 11'e eşittir. Tüm sınıflara düşen sıklıklar bu şekilde bulunur.

6. ADIM: p_i göreli sıklıklar hesaplanır.

Örneğin birinci sınıf için göreli sıklık,

$$p_1 = \frac{f_1}{n} = \frac{3}{45} = 0{,}07$$

ikinci sınıf için,

$$p_2 = \frac{f_2}{n} = \frac{11}{45} = 0{,}24$$

olacaktır.

7. ADIM: Birikimli ve göreli birikimli sıklıklar belirlenir.

Birikimli sıklıklar ya belli bir değerden az olan verilerin toplanması ya da belli bir değerden büyük olan verilerin toplanması ile oluşturulur. I. tür sıklık dağılımında, sınıfların üst sınırı o sınıfa dahil olmadığı için birikimli sıklıklar, sınıfların üst sınırlarından daha az olan değerler olarak yorumlanır. Eğer birikimli sıklık kolonu belli değerlerden daha büyük olanları verecek ise alt sınır ve daha fazla olan değerler toplanarak yorumlanır.

Birinci sınıfın sıklık değeri, aynı zamanda birinci sınıfın ...den az birikimli sıklık değeridir. Tablo 2.3'ün birinci sınıfının sıklığı 3 olduğundan birinci sınıfın ...den az birikimli sıklık değeri de 3 olarak yazılır. İkinci sınıfın ...den az birikimli sıklık değeri, birinci sınıfın ...den az birikimli sıklığı ile ikinci sınıfın sıklık değerlerinin toplamı olup

3+11=14 olacaktır. Tüm sınıflar için aynı işlemler uygulanır, ...den az birikimli sıklık kolonu oluşturulur. ... ve çok birikimli sıklığı da aynı işlem son sınıftan başlanarak, ilk sınıfa kadar uygulanarak elde edilir. Göreli birikimli sıklıklar da, her bir sınıfa düşen birikimli sıklıkların, toplam gözlem sayısına bölünmesiyle elde edilir. Örneğin birinci sınıfın ...den az göreli birikimli sıklık değeri; 3/45=0,07'dir.

Aşağıda verilen tablo yukarıdaki adımlardan sonra oluşturulan tablodur.

Tablo 2.3 İneklerin Kilogram Olarak Günlük Süt Üretim Değerlerine İlişkin Sıklık Dağılımı

	Sıklık	Göreli Sıklık	Birikimli Sıklıklar		Göreli Birikimli Sıklıklar	
Sınıflar	f_i	p_i	... Den Az	...Ve Çok	...Den Az	... Ve Çok
5,00≤..<10,00	3	3/45=0,07	3	45	0,07	1,00
10,00≤..<15,00	11	11/45=0,24	14	42	0,31	0,93
15,00≤..<20,00	14	14/45=0,31	28	31	0,62	0,68
20,00≤..<25,00	10	10/45=0,23	38	17	0,84	0,37
25,00≤..<30,00	5	5/45=0,11	43	7	0,96	0,15
30,00≤..<35,00	1	1/45=0,02	44	2	0,98	0,04
35,00≤..<40,00	1	1/45=0,02	45	1	1,00	0,02
	Σf_i=45	Σp_i=1,00				

Tablo 2.3'den yararlanarak yapılabilecek yorumlar: Sıklık dağılımından, karar vermeye yönelik yorumlar yapılabilir. 10 kg'dan az süt veren kaç inek vardır sorusu, birinci sınıfın den az birikimli sıklık değeri ile cevaplandırılır. Beşinci sınıfın den az birikimli sıklığından çiftlikteki hayvanlardan 43 tanesi 30 kg'dan az süt üretir sonucu söylenebilir. Dördüncü sınıfın den az göreli sıklık sütunundan, çiftlikteki ineklerin %84'ü, 25 kg'dan az süt üretmektedir sonucu çıkarılmaktadır. Süt verimi 5 kg ve fazla ineklerin sayısı 45, 10 kg ve fazla süt verimi olan ineklerin sayısı 42'dir.

Soru: Bu çiftlikteki hayvanların kaçı, 17,5 kg'dan az süt vermektedir?

17,5 kg doğrudan tabloda bulunmaz ama doğrusal öteleme (interpolasyon) uygulanarak bu değere karşılık gelen sayı aşağıdaki biçimde hesaplanabilir:

15.............. 14
17,5.............x
20.............. 28

$$\frac{17{,}5-15}{20-15} = \frac{x-14}{28-14} \Rightarrow x=21$$

21 inek yaklaşık olarak 17,5 kg'dan az süt vermektedir denilebilir.

2.3.2. Sürekli Verilerde II. Tür Sıklık Dağılımları

Bu tür sıklık dağılımı, TÜİK ve diğer kaynaklarda sınıflandırma yapılan verilerde görülmektedir. Daha çok ondalıklı verilmemiş veriler için uygun bir yöntemdir. I. tür sıklık dağılımından farklılığı, alt ve üst sınırların belirlenmesindedir. Bir sonraki sınıfın alt sınırı o sınıfın üst sınırındaki değerin son hanesinden bir fazla yani son basamağa 1 eklenmiş halidir.

Örnek 2.2: Bir iş yerinde çalışanların çocuklarına ait bir araştırmada sorulardan ikisi cinsiyet ve yaş ile ilgili olup, alınan cevaplar aşağıda verilmiştir:

Yaş	7	7	7	7	8	...	25	25	26
Cinsiyet	Erkek	Erkek	Kız	Erkek	Kız	...	Erkek	Kız	Kız

Verilere ilişkin sıklık dağılımı Tablo 2.4'de verilmiştir. Bu tablonun önceki sınıflamadan farkı, bir önceki sınıfın üst sınırı bir sonraki sınıfın alt sınırının son hanesinden bir eksik değere sahip olmasıdır. Örneğin Tablo 2.4'ün birinci sınıfının üst sınırı 9 ile bitmiş, ikinci sınıfın alt sınırı 10 ile başlamıştır. Diğer sınıflar da benzer şekilde oluşturulmuştur. Bu tür sınıflamada sıklıklar oluşturulurken, elimizde bulunan verilerin sınırları da dahil olmak üzere, sınırların aralarında değer alan veriler sayılır. Örneğin 10 ile 12 dahil olmak üzere bu değerler arasında değer alan veriler sayılmış ve bu yaş grubuna dahil 30 çocuk olduğu saptanmıştır.

Tablo 2.4 Örnek 2.2 için Yaşlara göre Sıklık Dağılımı

Sınıflar Alt Sınır-Üst Sınır	Sıklık	Sınıf Değeri (S_i)
7-9	20	(7+9)/2=8
10-12	30	(10+12)/2=11
13-15	150	14
16-18	60	17
19-21	30	20
22-24	5	23
25-27	5	26

2.3.3. İki Değişkeni Birlikte İçeren Birleşik Sıklık Dağılımları

Örnek 2.2'deki veriler için Tablo 2.5'de verildiği gibi bir sütun da cinsiyet için açılmıştır. Eğer kızlar için ayrı, erkekler için ayrı bir sıklık dağılımı istenirse ve cinsiyete göre kararlar alınmak istenirse, Tablo 2.5'de verilen sıklık dağılımı elde edilir. Örneğin 10 ile 12 dahil olmak üzere bu yaş grubunda olan erkek çocuk sayısı 18, kız çocuk sayısı 12 olarak bulunmuştur.

Tablo 2.5 Örnek 2.2 İçin Cinsiyet ve Yaşa göre Sıklık Dağılımı

Sınıflar Alt Sınır-Üst Sınır	Sıklık Erkek	Sıklık Kız
7-9	8	12
10-12	18	12
13-15	90	60
16-18	40	20
19-21	20	10
22-24	4	1
25-27	1	4
	Σf_i=181	Σf_i=119

Tablo 2.5'den yararlanarak göreli sıklıklar, birikimli sıklıklar vb. bulunabilir. Birinci sınıfta bulunan 8 erkek çocuk için göreli sıklık değeri, 8/181=0,044'dür. Bunun anlamı, erkeklerde 7-9 yaş grubuna ait çocuklar, toplam erkek çocuk sayısının %4,4'ünü oluşturmaktadır. Bayanlar için de ayrı bir sıklık sütunu olduğu için aynı tablodan kız öğrenciler için de bazı bilgiler elde edilebilir.

İki değişken üzerinden yapılan sınıflamalar yorum kolaylığı ve bir tablo ile daha fazla bilgi elde etme olanağı sağlar.

2.4. KESİKLİ NİCEL VERİLERDE SIKLIK DAĞILIMLARI

Kesikli nicel verilerde sıklık dağılımını oluşturma, Tablo 2.4'de verilen sıklık dağılımında ki gibidir. Burada önemli olan sınıf değerinin tam sayı olmasına dikkat edilmesidir.

Örnek 2.3: Bir yayınevinin günlük satış miktarları 1 yıl için incelenecektir. Bu yıl içinde seçilen 22 gün için kitap satış sayıları aşağıda verilmiştir:

19	39	58	75	135	195	196	200	235	254	255
286	312	314	356	370	371	373	373	430	433	490

Sınıf sayısı 8 olacak biçimde verilere ilişkin sıklık dağılımını oluşturunuz.

Örnek 2.1 için verilen adımlar bu örnek için de yapılabilir. Sınıf aralığı,

w=(490-19)/8=58,87~59

olmak üzere, elde edilen sıklık dağılımı Tablo 2.6'da verilmiştir.

Tablo 2.6 22 Gün İçin Kitap Satış Sayılarına İlişkin Sıklık Dağılımı

				Birikimli Sıklıklar		**Göreli Birikimli Sıklıklar**	
Alt Sınır-Üst Sınır	S_i	f_i	p_i	**...Ve Az**	**...Ve Çok**	**...Ve Az**	**...Ve Çok**
19-77	48	4	0,182	4	22	18	100
78-136	107	1	0,045	5	18	23	82
137-195	166	1	0,045	6	17	27	77
196-254	225	4	0,182	10	16	45	73
255-313	284	3	0,136	13	12	59	54
314-372	343	4	0,182	17	9	77	41
373-431	402	3	0,183	20	5	91	23
432-490	461	2	0,092	22	2	100	9

Tablo 2.6'dan görüldüğü gibi tabloda birikimli sıklıklar, ...ve az , ...ve çok kolonları olarak verilmiştir. ...ve az kolonunun birinci değeri, birinci sınıfa düşen sıklık değeridir. Sonraki satırlar, bir sonraki sıklıklar ile sırasıyla toplanarak elde edilir. ...ve çok kolonu ise, aynı işlem son sınıftan başlanarak yukarı doğru toplanarak elde edilir.

...ve az sıklıklarının yorumları için üst sınır, ...ve çok sıklıklarının yorumları için alt sınır kullanılır. Örneğin, 77 ve az 4 kitap satışı, 136 ve az 5 kitap satışı, 196 ve çok 16 kitap satışı görülmektedir.

2.5. NİTEL DEĞİŞKENLERDEN ELDE EDİLEN VERİLERDE SIKLIK DAĞILIMLARI

2.5.1. Sınıflanabilen Verilerde Sıklık Dağılımları

Sınıflanabilen verilerde sınıflar birbirlerinden bağımsız olarak elde edileceği için, sıklık dağılımında sadece sınıf, sıklık ve göreli sıklık kolonları bulunur.

Örnek 2.4: Bir turizm sezonunda ülkemize gelen 1000 turistin 300'ü Antalya, 280'i Bodrum ve geriye kalanı diğer tatil bölgelerini tercih etmiştir. Seçilen tatil bölgelerine ilişkin sıklık dağılımını oluşturalım.

Burada üç sınıf vardır ve bu sınıflar arasında geçiş yoktur. Sıklıklar ne kadar çoğaltılırsa da daha başka bir sınıf oluşturmak mümkün değildir. Aşağıda tercih edilen tatil bölgelerine ilişkin sıklık dağılımı verilmiştir.

Tablo 2.7 Tercih Edilen Tatil Bölgelerine İlişkin Sıklık Dağılımı

Tatil Bölgeleri	f_i	p_i
Antalya	300	300/1000=0,30
Bodrum	280	280/1000=0,28
Diğer	420	420/1000=0,42
	Σf_i=1000	

2.5.2. Sıralanabilir Verilerde Sıklık Dağılımları

Veriler belli bir sıralama ölçütüne göre sıralanabilen sınıflara ayrılır ve her sınıfa düşen gözlemlerin sayısı saptanarak sıklık dağılımı elde edilir. Sıklık dağılımında sınıf, sıklık, göreli sıklık, ve az ile ve çok birikimli sıklıklar ve bunlara ilişkin birikimli göreli sıklık kolonları yer alır.

Örnek 2.5: Bir iş yerinde çalışanların memnuniyetleri araştırılmak istenmiştir. Bunun için 77 kişi rasgele seçilmiş ve elde edilen sonuçlar aşağıda verilmiştir.

Tablo 2.8 Çalışanların Memnuniyetlerine İlişkin Sıklık Dağılımı

			Birikimli Sıklıklar		**Göreli Birikimli Sıklıklar**	
Memnuniyet Durumu	f_i	p_i	**...Ve Az**	**...Ve Çok**	**...Ve Az**	**...Ve Çok**
İyi Değil	29	0,377	29	77	0,376	1,000
Kısmen İyi	31	0,402	60	48	0,779	0,623
İyi	7	0,091	67	17	0,870	0,220
Çok İyi	6	0,078	73	10	0,948	0,129
Mükemmel	4	0,039	77	4	1,000	0,051

Çok iyi ve az memnuniyeti olan 73 kişi vardır. Çok iyi ve fazla memnun olan çalışanların örneklemdeki oranı % 12,9'dur.

Örnek 2.6: Üniversitemizde ders notları F, D-, D, D+, C-, C, C+, B-, B, B+, A-, A olarak sıralanmaktadır. 2004-2005 ders yılında Matematiksel İstatistik I dersini alan 61 öğrencinin harf notlarının sıklık dağılımı aşağıda verilmiştir.

Tablo 2.9 Notlara İlişkin Sıklık Dağılımı

			Birikimli Sıklıklar		Göreli Sıklıklar	
Notlar	**f_i**	**p_i**	**...Ve Az**	**...Ve Çok**	**...Ve Az**	**...Ve Çok**
A	8	0,13	8	61	0,13	1,00
A-	7	0,11	15	53	0,25	0,87
B+	1	0,02	16	46	0,26	0,75
B	1	0,02	17	45	0,28	0,74
B-	13	0,21	30	44	0,49	0,72
C+	4	0,07	34	31	0,56	0,51
C	7	0,11	41	27	0,67	0,44
C-	3	0,05	44	20	0,72	0,33
D+	6	0,10	50	17	0,82	0,28
D	3	0,05	53	11	0,87	0,18
D-	1	0,02	54	8	0,88	0,13
F	7	0,11	61	7	1,00	0,11

Yukarıda verilen tablodan yararlanarak aşağıdaki soruları cevaplayınız.

1. Matematiksel İstatistik I dersini alan öğrencilerin % kaçı C+ notunu almıştır?
2. Matematiksel İstatistik I dersini alan öğrencilerin % kaçı A notunu almıştır?
3. Matematiksel İstatistik I dersini alan öğrencilerin % kaçı C- ve az not almıştır?
4. Matematiksel İstatistik I dersini alan öğrencilerin % kaçı B ve çok not almıştır? Başarı sınırı C notu olduğuna göre bu sınıfın başarı oranı ne kadardır?

2.6. GRAFİKLER

Verilerin çizgisel gösterimlerine grafik denir. Grafikte, tabloda yer almayan hiçbir yeni bilgi söz konusu değildir. Bir koordinat düzleminde x ekseni sürekli ya da nicel değişkenlerde sınıf değerlerine; nitel değişkenlerde sınıflara ve y ekseni de her iki değişken için sıklığa ayrılır. Sıklıklar çok büyük değerler ise y ekseni için göreli sıklıklar da kullanılabilir.

Bir grafikte aşağıdaki bilgiler bulunmalıdır.

a) Grafiğin kısa ve öz bir adı olmalı ve grafiğin üst bölümüne yazılmalıdır.

b) Verilerin kaynağı sol alt bölüme yazılmalıdır.

c) Her iki eksende kullanılan ölçü birimleri belirtilmelidir.

d) Dipnot varsa, sağ alt kısma yazılmalıdır.

Grafikler oluşturulurken x ve y ekseninde kullanılacak olan ölçek birimine göre farklı çizimler elde edilebilir. Bu da kişiye yanıltıcı bilgiler verebilir. Bu konuda ayrıntılı bilgi Kohler, 1985'den elde edilebilir.

2.6.1. Nitel Değişkenlerden Elde Edilen Verilerde Grafikler

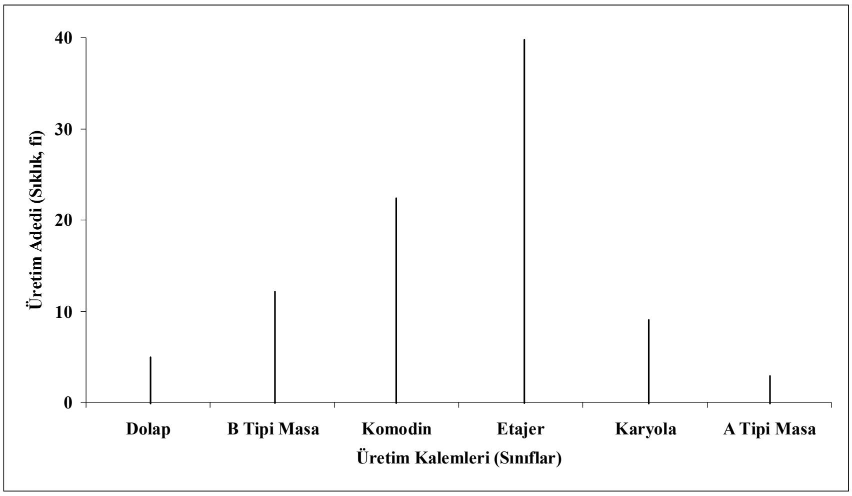
Çubuk Grafiği

Düzlem üzerindeki (x,y) (x=Sınıfı temsil eden nokta; y=Sınıf sıklığı) noktaları, xy koordinat sisteminde işaretlenir ve bu noktalardan x eksenine kadar dikmeler indirilirse, çubuk grafiği elde edilir. Bu grafik kesikli verilerde ve nitel verilerde kullanılan grafik türüdür.

Örnek 2.7: Bir mobilya fabrikası 1 haftada 5 dolap, 12 B tipi masa, 22 komodin, 39 etajer, 9 karyola ve 3 A tipi masa üretmektedir. Sıklık dağılımını oluşturunuz ve çubuk grafiğini çiziniz.

Tablo 2.10 Üretim Kalemlerine İlişkin Sıklık Dağılımı

Üretim Kalemleri	f_i	p_i
Dolap	5	0,06
B Tipi Masa	12	0,13
Komodin	22	0,24
Etajer	39	0,44
Karyola	9	0,10
A Tipi Masa	3	0,03

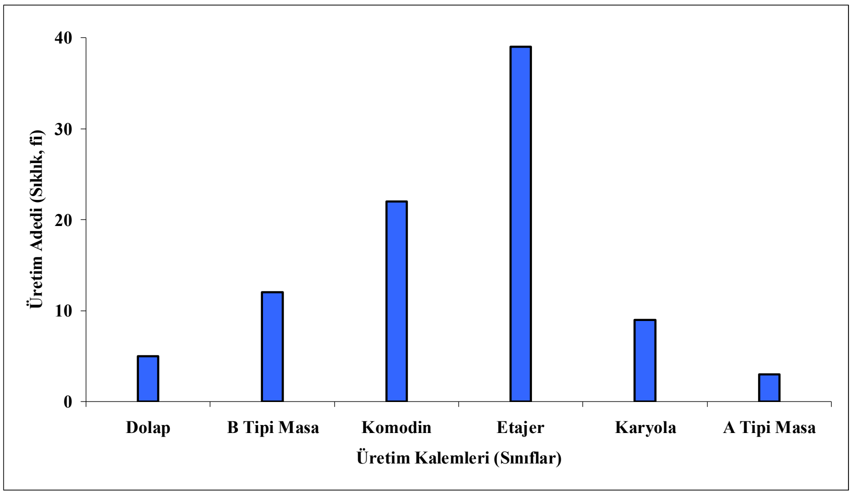


Şekil 2.1 Örnek 2.7'de Verilen Değerlere İlişkin Çubuk Grafiği

Ayrık Dikdörtgenler

Sınıfı temsil eden aralık üzerine, sınıf sıklığını yükseklik kabul eden ayrık dikdörtgenler çizilir.

Örnek 2.8: Örnek 2.7'deki verilere ilişkin ayrık dikdörtgenler grafiğini oluşturunuz.



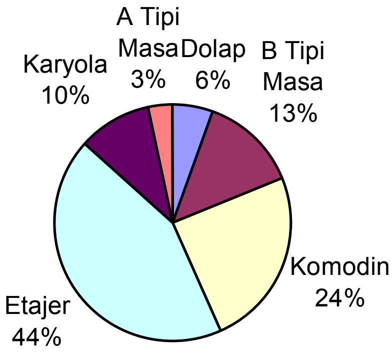
Şekil 2.2 Örnek 2.7'de Verilen Değerlere İlişkin Ayrık Dikdörtgenler Grafiği

Daire Dilimleri Grafiği

Bu grafikte her bir sınıfa düşen sıklık bir dairenin parçası ile gösterilir. Bu grafiği çizebilmek için öncelikle göreli sıklıklar hesaplanır, sonra da her bir göreli sıklık değeri 360^0 ile çarpılarak, o sınıfa düşen daire dilimlerini belirleyen merkez açılar bulunur. Tüm sınıflar için bu uygulama yapıldığında daire dilimlenmiş olur.

Örnek 2.9: Örnek 2.7'de verilen verilere ilişkin daire dilimleri grafiğini çiziniz.

Üretim Kalemleri	f_i	p_i	Merkez açı ($p_i \times 360^0$)
Dolap	5	0,06	21,6
B Tipi Masa	12	0,13	46,8
Komodin	22	0,24	86,4
Etajer	39	0,44	158,4
Karyola	9	0,10	36
A Tipi Masa	3	0,03	10,8



Şekil 2.3 Örnek 2.7'de Verilen Değerlere İlişkin Daire Dilimleri Grafiği

2.6.2. Nicel Verilerde Grafikler

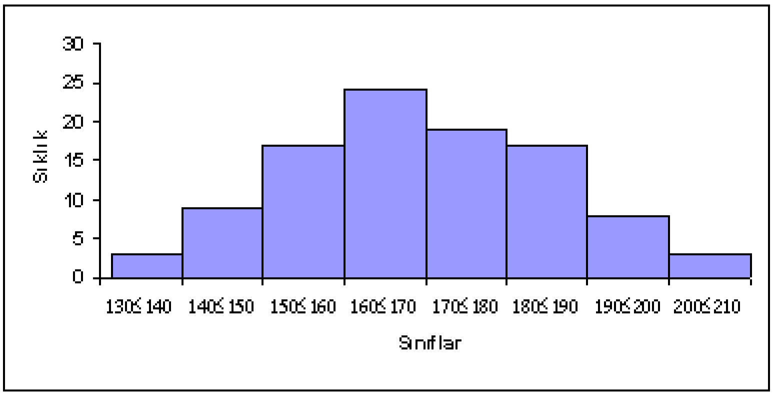
Histogram (Dağılım Dikdörtgenleri)

Yüksekliği sıklıklarıyla orantılı, taban uzunlukları x ekseni üzerinde ve genellikle sınıf aralığına eşit olarak yan yana yapışık olarak oluşturulan dikdörtgenler histogram grafiğini meydana getirirler. Farklı sınıf aralığına sahip dağılımlarda histogram yerine başka bir grafik türü kullanmak gerekebilir.

Örnek 2.10: Rasgele seçilen 100 ineğin ağırlıklarına ilişkin (kg olarak) verilerin sıklık dağılımı Tablo 2.11'de verilmiştir. Tabloya ilişkin histogram grafiğini oluşturunuz.

Tablo 2.11: İneklerin Ağılıklarına İlişkin Sıklık Dağılımı

Sınıflar	f_i
130≤..<140	3
140≤..<150	9
150≤..<160	17
160≤..<170	24
170≤..<180	19
180≤..<190	17
190≤..<200	8
200≤..<210	3



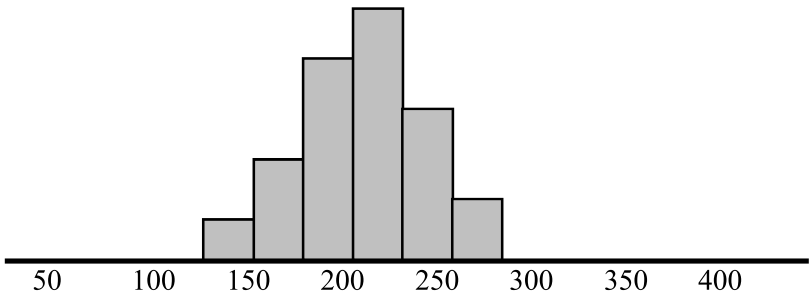
Şekil 2.4 Tablo 2.11'de Verilen Tabloya İlişkin Dağılım Dikdörtgenleri Grafiği

Dağılım dikdörtgenleri grafiklerinden yararlanarak aşağıda belirtilen 3 farklı bilgi sağlanabilir.

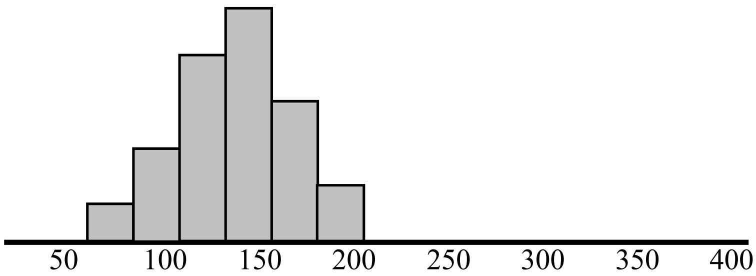
1. Verinin merkezinin yaklaşık olarak nerede olduğunu gösterir.
2. Verideki yayılımı gösterir.
3. Dağılımın şeklini gösterir (Basık, sağa-sola çarpık, merkezde toplanmış, çan şeklinde, vb.).

Şekil 2.5'de yayılımı aynı, konumları farklı aynı değişken için çeşitli çizimler verilmiştir.

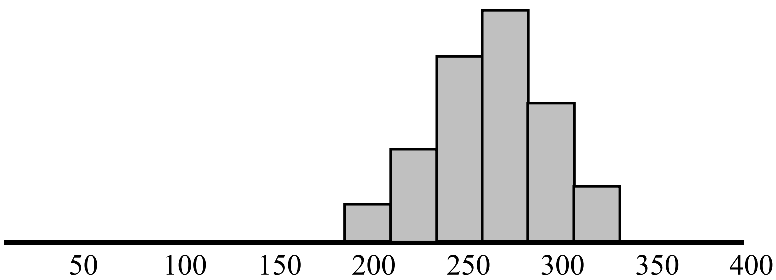
a)



b)



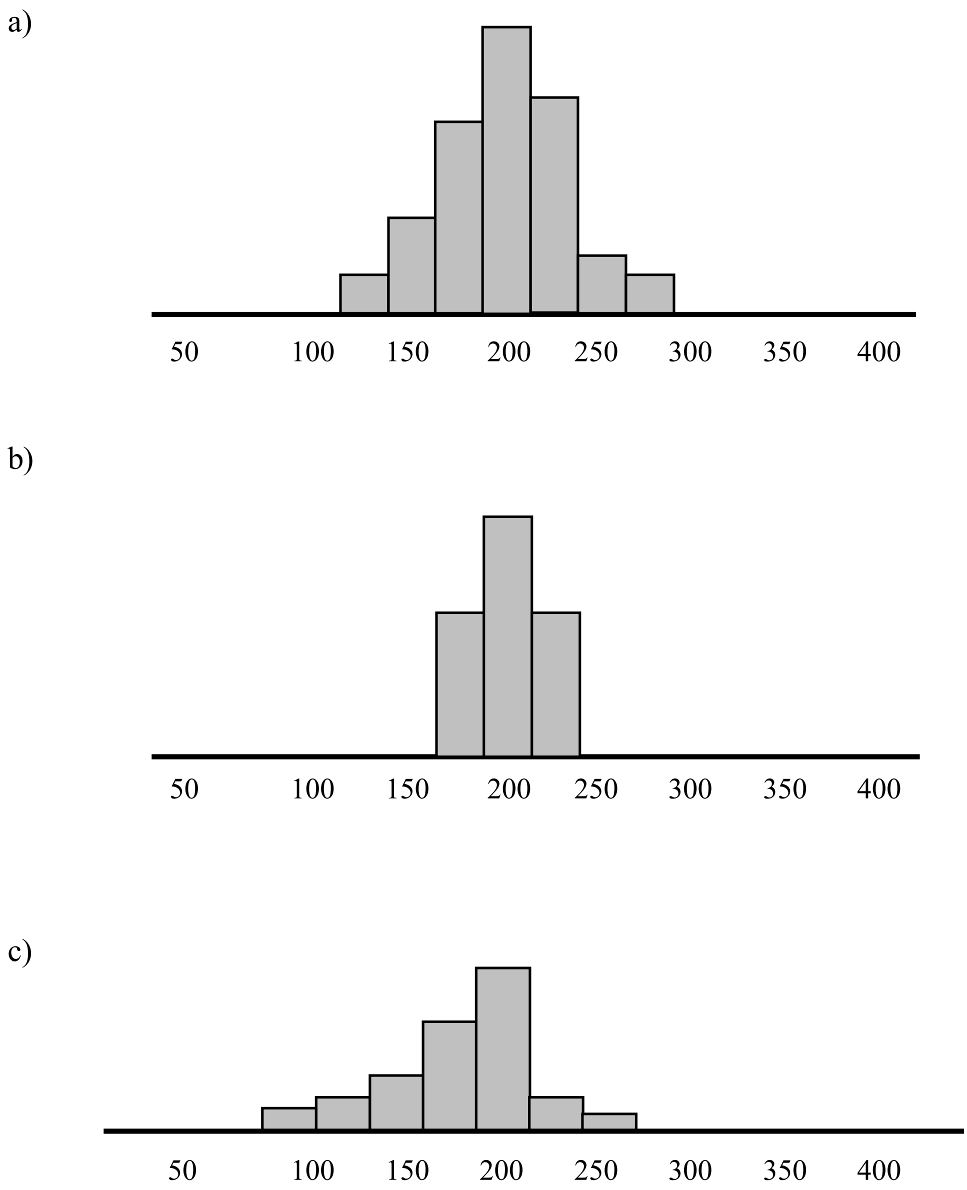
c)



Şekil 2.5 Aynı Yayılımlı, Farklı Konumlardaki Örneklem Dağılımları

Dağılım dikdörtgenleri grafikleri a, b ve c aynı değişken için üç farklı veri grubundan oluşturulmuştur. İlk bakışta verilerin çoğunluğunun nerede olduğu görülmektedir. Eğer bu veriler bir tatil köyünde kurulan pazarda A, B ve C mağazalarının 1 ay boyunca günlük satış getirileri ise, C mağazasının en başarılı olduğu söylenebilir.

Dağılım dikdörtgenleri grafikleri verilerin yayılımı ile ilgili de görsel bilgi verir. Aynı değişkene ait veriler için karşılaştırma olanağı sağlar. Aşağıda konuya ilişkin grafik verilmiştir (Şekil 2.6).



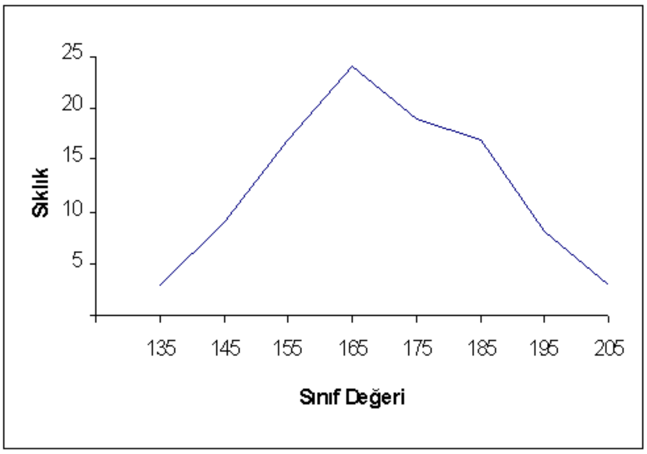
Şekil 2.6 Aynı Konumlu, Farklı Yayılımlı Örneklem Dağılımları

Şekil 2.6, bir tatil köyünde bulunan A, B ve C mağazalarının günlük satış getirilerini gösterdiğinde hangi mağaza daha istikrarlıdır? Her üç mağazanın da ortalama günlük satışları 200 dolar civarında iken, üçüncü mağazanın (c) bir istikrar sağlayamadığı görülmektedir. A mağazasının satış getirileri ise B mağazasına göre daha fazla yayılım göstermektedir.

Dağılış Poligonu

Sınıf değerlerinin sıklıklarla birleştirilen noktalarının kesik çizgilerle birleştirilmesi ile elde edilen grafik türüdür. Grafiği çizebilmek için öncelikle sınıf değerlerinin bulunması gerekir.

Örnek 2.11: Tablo 2.11'de verilen tabloya ilişkin dağılış poligonu grafiğini oluşturunuz.

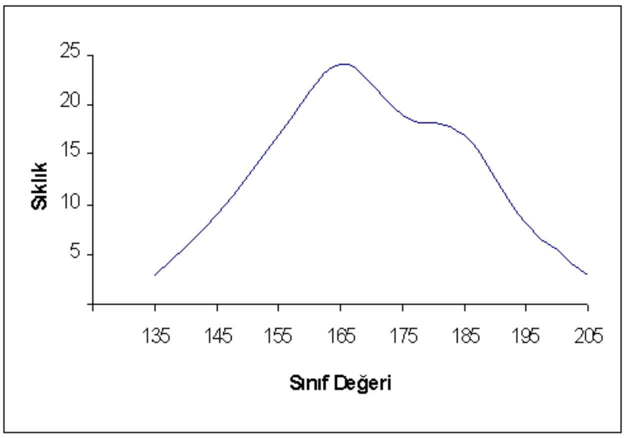


Şekil 2.7 Tablo 2.11.'de Verilen Tabloya İlişkin Dağılış Poligonu Grafiği

Dağılış Eğrileri

Poligonlar için belirlenen noktaların kesik çizgiler yerine eğrilerle birleştirilmesidir. Sınıf aralıkları küçüldüğünde çizilebilir. Ya da kuramsal dağılımlar için kullanılır.

Örnek 2.12: Tablo 2.11'de verilen tabloya ilişkin dağılış eğrileri grafiğini oluşturunuz.



Şekil 2.8 Tablo 2.11'de Verilen Tabloya İlişkin Dağılış Eğrisi Grafiği

2.7. FARKLI SINIFLANDIRMALAR

Farklı Aralıklı Sıklık Dağılımı

Verilerin birimleri arasında çok fark varsa, bazı sınıfların sıklıkları çok büyük bazı sınıfların sıklığı çok az hatta sıfır ise, tek bir sınıf aralığı ile sıklık dağılımının oluşturulması anlamlı olmayabilir. Bu durumda farklı sınıflar için farklı sınıf aralıkları kullanılabilir. Aşağıda farklı sınıflandırmaya ilişkin örnek verilmiştir.

Örnek 2.13: Bir sigorta şirketi müşterilerinin 2005 yılının ocak ayına kadar devamlı kaç yıldır sigorta yaptırdıklarının dökümünü özet olarak aşağıdaki biçimde belirlemiştir.

Devamlı sigorta yaptıran yıl	Müşteri sayısı
5	6
6	8
7	-
8	1
9	-
10	9
11	-
12	3
13	-
14	2
15	1

Verilere ilişkin sıklık dağılımını oluşturunuz.

Tablo 2.12 Örnek 2.13'deki Verilerin Sıklık Dağılımı

	Sınıflanmış Veriler	
	Yıllar	**Müşteri Sayısı**
w=1	5	6
	6	8
	7-9	1
w=3	10-12	12
	13-15	3

Örnek 2.14: Küçük ticari işyerlerinin gayri safi satış tutarı (G.S.S.T.) (TL)'ye göre farklı aralıklı sınıflamaya sahip sıklık dağılımı aşağıda verilmiştir (Yıl 1998).

G.S.S.T (YTL):	İŞYERİ SAYISI:
0..............................10.000	----
10.001..............................50.000	308
50.001..............................100.000	434
100.001............................250.000	4397
250.001............................500.000	8393

Açık Uçlu Sıklık Dağılımları

Bazen kitlede bulunacak en büyük değer ile en küçük değer için herhangi bir bilgimiz olmayabilir. Bu durumda ilk sınıfın alt sınırı ve son sınıfın üst sınırı belirsizdir. Bu tür sıklık dağılımına açık uçlu sıklık dağılımı denir. Aşağıda konuya ilişkin bir uygulama verilmiştir.

Örnek 2.15: 1985 sayım sonuçlarına göre şehirlerin nüfus sayılarına göre dağılımı:

Nüfus	Şehir Sayısı
-750	0
751-1000	1
1001-2000	12
2001-3000	53
3001-4000	40
4001-5000	56
5001-6000	44
6001-7000	40
7001-8000	31
8001-9000	28
9001-10000	22
10001-15000	93
15001-2000	48
20001-25000	30
25001-50000	62
50001-100000	37
100001-	41

PROBLEMLER

1- Aşağıda verilen ham verilerin 6 sınıflı sıklık dağılımını oluşturunuz.

13	13	14	17	18	18	19	21	21	22
23	23	24	25	25	26	26	27	27	28
29	30	32	32	35	36				

2- 30 öğrencisi olan bir sınıfın istatistik dersinden 30 üzerinden aldıkları notlar aşağıda verilmiştir:

2	2	3	4	6	7	7	8	8	9
10	10	10	10	11	11	11	12	12	12
13	13	14	16	18	20	25	25	28	28

a- 7 sınıflı sıklık dağılımını oluşturunuz.

b- Öğrencilerin % kaçı 15 puan ve daha az puan almıştır?

3- Bir psikoloji laboratuarında 20 deney hayvanının öğrenme derecelerine göre sıklık dağılımı düzenlenmiştir. Verilere ilişkin göreli sıklık ve birikimli sıklık kolonlarını oluşturunuz.

Öğrenme Derecesi	**Sıklık**
Çok Zayıf	2
Zayıf	3
Orta	6
İyi	4
Çok İyi	5

4- Bir firma piyasaya süreceği gıda maddesi için bir anket düzenlemiştir. Ankette bulunan bilgilerden biri de paketleme ile ilgilidir. Ankete katılan 1000 kişiden 30'u mamulün açık satılmasını, 400'ü karton kutuda, 350'si naylon torbada, geri kalanı kapalı kağıt torbada satışa sunulmasını uygun bulmuştur. Bu sonuçları gösterebilecek bir sıklık dağılımı oluşturunuz ve uygun bir grafik çiziniz.

5-

Hane halkı sayısı:	1	2	3	4	5	6
Sıklık:	20	24	17	15	9	5

Yukarıdaki bilgiye göre veri türü nedir? Verilere ilişkin uygun grafik türlerinden birinin çizimini yapınız.

6- Bir kliniğe yüksek tansiyon şikayeti ile başvuran 200 kişiden, 100'ü hiç sigara içmediğini, 60'ı az sigara içtiğini, 40'ı çok içtiğini belirtmiştir. Sıklık dağılımını oluşturunuz.

7- Bir kümeste elde edilen yumurtalardan rasgele seçilen 30 yumurtanın ağırlıkları gram olarak aşağıdaki şekilde bulunmuştur.

50,0	52,5	55,4	56,0	57,3	58,4	50,2	52,6	55,5	56,2
57,4	58,8	50,6	53,5	55,6	56,8	57,8	59,0	51,3	54,0
44,7	56,9	58,0	59,6	52,2	55,4	55,8	57,0	58,0	60,0

a- Verilerin 10 sınıflı sıklık dağılımını oluşturunuz.

b- Yumurtaların % 30'u kaç gramdan daha ağırdır?

c- Yumurtaların yüzde kaçı 57 gr'dan daha ağırdır?

8- Sınıf mevcudu 100 olan ekonomi dersinin genel sınav notlarına ilişkin sıklık dağılımı aşağıda verilmiştir. Eksik yerleri tamamlayınız.

Notlar	Sıklık	Göreli sıklık	...Ve az	...Ve çok
F	40			
D		0,20		
C			72	
B				
A				4

9- Aşağıdaki sıklık dağılımındaki

a- Boş yerleri doldurunuz.

b- Deneklerin kaç tanesi 2,2 ve çoktur?

Alt Sınır- Üst Sınır	S_i	f_i	...Ve az	...Ve çok
-1,6			3	
-1,8			5	
-2,0			7	
-2,2			18	
-2,4			27	
-2,6			35	
-2,8			45	

10- Deney hayvanı yetiştiren bir kuruluşta A türünden 30, B türünden 35, C türünden 70, D türünden 65 hayvan bulunmaktadır.

a- Değişken türünü belirtiniz.

b- Verilerin sıklık dağılımını düzenleyiniz.

c- Uygun bulduğunuz bir grafik çiziniz.

11- Prematüre bebeklerden 10 tanesinde ölüm ve canlı durumu ile cinsiyete göre özellikleri aşağıda verilmiştir:

Ö: Ölü C:Canlı E:Erkek K:Kız

Doğum sonucu	Ö	Ö	C	C	C	C	Ö	C	Ö	Ö
Cinsiyet	K	K	E	E	E	K	K	K	E	K

a- Verilerin sıklık dağılımını oluşturunuz.

b- Ölü doğumlarda kız bebeklerin ve canlı doğumlarda erkek bebeklerin yüzdesi nedir?

12- 33 kişinin bir günde içtikleri sigara miktarları aşağıda verilmiştir. Sıklık dağılımını 7 sınıf için oluşturunuz.

5	15	20	8	20	12	10	10	12	15	15
18	20	3	2	3	3	3	6	10	20	15
20	15	15	20	21	20	3	6	5	17	21

13- Televizyon üreten bir firma 30 kişiye bir hafta içinde kaç saat televizyon izlediğini sormuş ve aşağıdaki bilgileri elde etmiştir: Verilere ilişkin sıklık dağılımını 8 sınıf için oluşturunuz.

7	6	3	9	5	10	2	14	5	4
1	7	4	6	3	8	5	6	8	7
3	12	3	6	5	8	11	6	2	5

14- Ankara'daki 300 pastanenin günlük ortalama satışlarına ilişkin sıklık dağılımı oluşturulmuş ve aşağıda verilmiştir. Verilere ilişkin uygun bir grafik çiziniz.

Satış (kg)	f_i
50≤..<100	20
100≤..<150	30
150≤..<160	45
160≤..<170	70
170≤..<180	75
180≤..<190	20
190≤..<200	20
200≤..<210	20

15- 20 kişilik bir sınıfta öğrencilerin boy uzunlukları aşağıda verilmiştir Verilere ilişkin sıklık dağılımını oluşturunuz

1,65	1,68	1,68	1,69	1,69	1,69	1,70	1,75	1,76	1,76
1,76	1,76	1,78	1,79	1,79	1,79	1,79	1,80	1,81	1,86

16- 2004 yılı verilerine göre, Türkiye'de bulunan yabancıların 114315'i ikamet amacıyla, 22107'si öğrenim görmek amacıyla ve 21140'ı çalışmak amacıyla Türkiye'dedir (Kaynak: Emniyet Genel Müdürlüğü). Sıklık dağılımını oluşturunuz ve uygun bulduğunuz bir grafiği çiziniz.

17- 2003 yılında, yurt dışında lisans eğitimi yapan 11318, yüksek lisans yapan 3096 ve doktora yapan 1809 öğrenci vardır (Kaynak: Milli Eğitim Bakanlığı, Yükseköğretim

Genel Müdürlüğü). Verilere ilişkin sıklık dağılımını oluşturunuz ve uygun bir grafik çiziniz.

18- Soru 4'de yapılan ankette sorulardan biri anketi cevaplayanların yaşları ile ilgilidir. Rasgele seçilen 26 kişiden alınan cevaplar aşağıda verilmiştir:

18	19	20	21	22	23	24	30	35	38	40
41	43	44	45	46	47	50	51	52	54	58
60	61	65	70							

Verilere ilişkin sıklık dağılımını oluşturunuz.

19- Soru 18'de oluşturulan sıklık dağılımı için birikimli sıklıkları ve göreli sıklıkları oluşturunuz. Bu tablolardan nasıl yararlanabilirsiniz?

20- Soru 19'da oluşturulan sıklık dağılımını kullanarak firmanın piyasaya süreceği gıda maddesinin hangi yaşlara cevap vermesi gerektiği konusunda politikası için bir yorum yapınız.

3. ÜÇÜNCÜ BÖLÜM

3.1. GİRİŞ

3.2. NİCEL VERİLERDE KONUM ÖLÇÜLERİ

3.3. NİTEL VERİLERDE KONUM ÖLÇÜLERİ

3.4. NİCEL VERİLERDE DEĞİŞİM ÖLÇÜLERİ

3.5. NİTEL VERİLERDE DEĞİŞİM ÖLÇÜLERİ

PROBLEMLER

3.1. GİRİŞ

İstatistiksel yöntemlerle bir dağılımı betimlemek için çeşitli ölçümler bulunmuştur. En çok kullanılan ölçüm, bir dizi verinin oluşturduğu dağılımın konum ölçüsüdür. Konum ölçüleri, verilerin dağılımlarının birbirine göre yerlerini, uzaklıklarını kısacası, konumlarını belirtirler. Dağılım hakkında bilgiler elde etmemizi sağlarlar. Bu bilgilerden yararlanarak sıklık dağılımını tanımak ve karşılaştırmak olanağı vardır. Ancak çok yönlü karşılaştırmalar yapabilmek için konum ölçüleri yeterli değildir. Veriler arasındaki değişimden kaynaklanan farklılıkların ölçüsü olan değişim ölçülerinin de bilinmesi gerekir. Aşağıda sırasıyla konum ve değişim ölçüleri açıklanacaktır.

3.2. NİCEL VERİLERDE KONUM ÖLÇÜLERİ

Nicel verilerde en sık kullanılan konum ölçüleri aritmetik ortalama, tepe değeri, ortanca'dır. Ağırlıklı ortalama, geometrik ortalama, harmonik ortalama, yüzdelikler, çeyrek değerler ve ondalıklar da diğer konum ölçüleridir.

Aritmetik Ortalama

Kitle büyüklüğü N ve X_j, j. gözlem değeri olmak üzere kitle için aritmetik ortalama (μ),

$$\mu = \frac{\sum_{j=1}^{N} X_j}{N} \qquad (3.1)$$

eşitliği ile elde edilir. Aritmetik ortalamanın özellikleri aşağıda verilmiştir:

1) Dağılımda yer alan tüm değerler dikkate alınarak hesaplanır.
2) Dağılımdaki verilerin aritmetik ortalamadan ayrılışları toplamı sıfırdır.
3) Dağılım sınırları (uç değerler) aritmetik ortalama üzerinde etkilidir. Dağılım sınırları aşırı değerler ise, aritmetik ortalama değeri dağılım hakkında tutarlı bilgi vermez. Bu durumda dağılım sınırlarından etkilenmeyen ortanca, çeyrek değerler, tepe değeri gibi diğer konum ölçülerinden yararlanılır.

Örneklemden elde edilen ortalama $\overline{X}$ ile gösterilecektir. Örneklem aritmetik ortalaması sınıflandırılmış ve sınıflandırılmamış veriler için farklı hesaplanır.

a) Sınıflandırılmamış (Sıklık Dağılımı Düzenlenmemiş) Verilerde Aritmetik Ortalama

$X_1, X_2, ..., X_n$ örneklem verilerine ilişkin veriler ise aritmetik ortalama;

$$\overline{X} = \frac{\sum_{j=1}^{n} X_j}{n} \tag{3.2}$$

biçiminde bulunur.

Örnek 3.1: 10 haftada kiralanan oda sayısı ve şikayetlere ilişkin aşağıdaki tablo ele alınsın.

Tablo 3.1 Kiralanan Oda Sayısı ve Şikayetlere İlişkin Değerler

Hafta	Kiralanan oda sayısı (X)	Şikayetler (Y)
1	30	0
2	23	3
3	15	2
4	18	1
5	32	1
6	25	0
7	20	3
8	19	2
9	20	2
10	19	0

10 haftada kiralanan ortalama oda sayısı ve ortalama şikayetler sırasıyla;

$$\overline{X} = \frac{30 + 23 + 15 + 18 + 32 + 25 + 20 + 19 + 20 + 19}{10} = 22{,}1$$

ve

$$\overline{Y} = \frac{0 + 3 + 2 + 1 + 1 + 0 + 3 + 2 + 2 + 0}{10} = 1{,}4$$

olacaktır. 10 haftada ortalama kiralanan oda sayısı 22'dir. Yaklaşık olarak ortalama 1,4 şikayet söz konusu olmuştur.

b) Sınıflandırılmış (Sıklık Dağılımı Düzenlenmiş) Verilerde Aritmetik Ortalama

Sıklık dağılımı düzenlenmiş veriler için aritmetik ortalama aşağıdaki biçimde hesaplanır;

$$\overline{X} = \frac{\sum_{i=1}^{k} f_i \times S_i}{n} \tag{3.3}$$

Burada f_i, i. sınıfın sıklığı; S_i, i.sınıfın sınıf değeri (sınıf orta değeri) dir.

Örnek 3.2: 30 öğrencinin Matematiksel İstatistik I dersinden aldıkları notlara ilişkin değerler aşağıda verilmiştir. Değerlere ilişkin sıklık dağılımını 7 sınıf için oluşturunuz ve aritmetik ortalamayı bulunuz.

75, 36, 73, 52, 79, 65, 56, 53, 48, 53, 59, 60, 84, 49, 58,
43, 52, 50, 82, 60, 60, 64, 74, 55, 51, 74, 66, 84, 55, 69

Çözüm: Öncelikle veriler küçükten büyüğe sıralanır ve sıklıkların daha kolay elde edilmesi sağlanır.

36, 43, 48, 49, 50, 51, 52, 52, 53, 53, 55, 55, 56, 58, 59,
60, 60, 60, 64, 65, 66, 69, 73, 74, 74, 75, 79, 82, 84, 84

Sınıf aralığı,

w=(84-36)/7=6,85~7

olmak üzere verilere ilişkin sıklık dağılımı aşağıdaki biçimde elde edilir.

Tablo 3.2 Öğrencilerin Notlarına İlişkin Sıklık Dağılımı

Alt Sınır-Üst Sınır	f_i	S_i	$f_i \times S_i$
36-42	1	(36+42)/2=39	1×39=39
43-49	3	46	3×46=138
50-56	9	53	9×53=477
57-63	5	60	5×60=300
64-70	4	67	4×67=268
71-77	4	74	4×74=296
78-84	4	81	4×81=324
			$\sum_{i=1}^{7}(f_i \times S_i) = 1842$

$$\overline{X} = \frac{1842}{30} = 61,4$$

30 öğrencinin notlarına ilişkin aritmetik ortalama değeri 61,4 olarak bulunmuştur.

Tepe Değeri

a)Sınıflandırılmamış verilerde tepe değeri

Sınıflandırılmamış verilerde tepe değeri (mode), en çok tekrarlanan gözlem değeri, yani en büyük sıklığa sahip olan değerdir. Kitap boyunca tepe değeri $\hat{X}$ ile gösterilecektir.

Örnek 3.3: Aşağıdaki verilere ilişkin tepe değerini bulunuz

144000, 98000, 204000, 177000, 155000, 316000, 177000, 177000, 100000

Burada en sık rastlanan değer 177000 (3 kez tekrarlanır) olduğu için tepe değeri; $\hat{X}$=177000'dır.

Örnek 3.4: Örnek 3.2'de verilen, öğrencilerin notlarına ilişkin verilerin tepe değeri, en çok tekrarlanan $\hat{X}$=60'dır.

Not: Bir örneklem dağılımında bazen en çok tekrarlanan değer iki ya da daha fazla olabilir. Bu durumda, iki ya da daha fazla tepe değeri elde edilebilir ve dağılım iki tepeli ya da çok tepeli olarak adlandırılır.

Örnek 3.5: Aşağıda verilen örnekleme ilişkin tepe değerini bulunuz.

2 2 2 7 7 7 11 14 15 16 6 8

En çok tekrarlanan değerler 2 ve 7'dir. Her ikisi de 3 sıklığına sahiptir. Dolayısıyla, $\hat{X}_1 = 2$, $\hat{X}_2 = 7$ olmak üzere, iki tepe değeri vardır.

b) Sınıflandırılmış verilerde tepe değeri

Sınıflandırılmış verilerde tepe değeri aşağıda verilen formül ile elde edilir:

$$\hat{X} = A_S + w\left(\frac{F_1}{F_1 + F_2}\right) \tag{3.4}$$

Burada,

A_S: En büyük sıklığın bulunduğu sınıfın alt sınırı

w: Sınıf aralığı

F_1: En büyük sıklık ile bir önceki sınıfın sıklığı arasındaki fark

F_2: En büyük sıklık ile bir sonraki sınıfın sıklığı arasındaki fark

Örnek 3.6: Aşağıda sıklık dağılımı düzenmiş veriler için tepe değerini bulunuz.

Tablo 3.3 Bir Sınava Giren Öğrencilerin Yaptıkları Hata Sayıları

Alt Sınır-Üst Sınır	fi
8-10	22
11-13	15
14-16	25
17-19	24
20-22	6
23-25	5
26-28	3

A_S=14; w=3; F_1= 25-15=10; F_2=25-24=1 olmak üzere tepe değeri,

$$\hat{X} = 14 + 3\left(\frac{10}{10+1}\right) = 16{,}73$$

olur.

Ortanca

Sınıflandırılmamış verilerde ortanca (median), küçükten büyüğe sıralanmış veriler için, en ortadaki yani verileri iki eşit parçaya bölen değerdir ve $\overline{X}'$ ile gösterilecektir. Toplam gözlem sayısının tek ya da çift sayı olmasına göre farklı hesaplama yapılır.

i) Gözlem sayısı n tek ise; $\frac{n+1}{2}$'inci değer ortancadır.

Örnek 3.7: Elimizde aşağıdaki gibi örneklem verisi olsun.

1 3 4 5 6 6 6 9 10 n=9

(9+1)/2=5'inci değer ortancadır. Yani, $\overline{X}'$=6'dır.

ii) Gözlem sayısı n çift ise; $\frac{n}{2}$'inci ve $(\frac{n}{2}+1)$'inci değerin ortalaması ortancadır.

Örnek 3.8: Elimizde aşağıdaki gibi örneklem verileri olsun.

15,6 15,6 18,9 18,9 20,7 20,7 23,4 23,4 25,4 25,4 25,4 25,4 n=12

$$\frac{n}{2} = \frac{12}{2} = 6 \quad \text{ve} \quad \frac{n}{2} + 1 = 6 + 1 = 7$$

6. ve 7. değerin ortalaması, verilere ilişkin ortancayı verir. Yani, 6. değer 20,7 ve 7. değer 23,4 olmak üzere ortalama,

(20,7+23,4)/2=22,05 olup, ortanca $\overline{X}'$=22,05'dır.

Örnek 3.2'de verilen, öğrencilerin notlarına ilişkin verilerin ortanca değerini bulalım.

$\frac{n}{2}=\frac{30}{2}=15$ 'inci değer ile $\frac{n}{2}+1=15+1=16$ 'ıncı değer ortalaması alınır.

15'inci değer=59 ve 16'ıncı değer=60 olmak üzere, ortanca $\overline{X}'$=59,5'dir.

Sınıflandırılmış verilerde ortanca birikimli sıklık değerleri kullanılarak, yüzde ellilik noktanın değeri olarak verilir. Tablo 3.6'da verilen sıklık dağılımından yararlanarak verilere ilişkin ortancayı bulalım.

Alt Sınır-Üst Sınır	fi	...Ve Az
8-10	22	22
11-13	15	37
14-16	25	62
17-19	24	86
20-22	6	92
23-25	5	97
26-28	3	100

50. gözlem değeri ortancadır. 50. gözlem değeri, 37 ve 62 arasına düşer. Doğrusal öteleme kullanılarak ortanca,

13.............. 37

$\overline{X}'$............. 50

16.............. 62

$$\frac{50-37}{62-37}=\frac{\overline{X}'-13}{16-13} \Rightarrow \overline{X}'=14,56$$

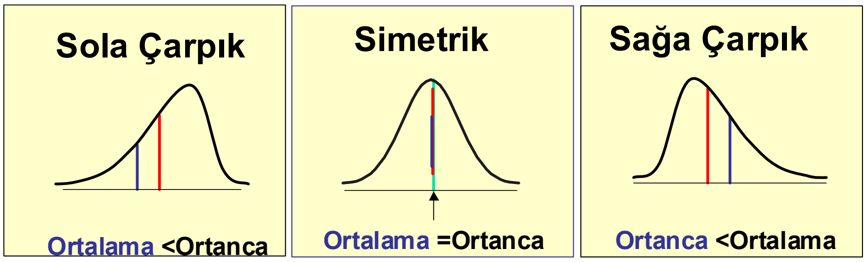
olarak bulunur.

Tek Tepeli Dağılımlarda Ortalama, Tepe Değeri ve Ortanca Arasındaki İlişki

1) Ortalama<Ortanca<Tepe değeri ise, dağılım sola-çarpık yani (-) yöne eğilimli dağılımdır.
2) Ortalama=Tepe Değeri=Ortanca ise, dağılım simetrik dağılımdır.
3) Tepe değeri<Ortanca<Ortalama ise, dağılım sağa-çarpık yani (+) yöne eğilimli dağılımdır.

Not: Dağılım simetrik dağılıma yakın ise tepe değeri ile ortanca yer değiştirebilir.

Aşağıda sırasıyla sola çarpık (a), simetrik (b) ve sağa çarpık (c) dağılımlara ilişkin grafikler verilmiştir.



(a) **(b)** **(c)**

Şekil 3.1 Çarpık ve Simetrik Dağılımlar

Örnek 3.9: Örnek 3.2'de verilen 30 öğrencinin notlarına ilişkin ortalama 61,4, tepe değeri 60 ve ortanca 59,5 olarak bulunmuştur. Bu değerleri karşılaştırdığımızda; ortanca<tepe değeri<ortalama eşitsizliği vardır. O halde, bu örneklem dağılımının sağa çarpık bir dağılım olduğu söylenebilir. Başka bir deyişle, 30 öğrencinin notları sağa çarpık bir dağılım göstermektedir.

Çeyrek Değerler

Yüzdelik kavramı, kendisinden önce ve sonra belirli oranlarda değerler bulunan noktanın değerini açıklar. Örneğin 25'inci yüzdelik (birinci çeyrek, Q_1) demek, gözlemlerin % 25'i kendisinden küçük, % 75'i kendisinden büyük değerli olan noktadır. Aynı şekilde 75'inci yüzdelik (üçüncü çeyrek, Q_3), gözlemlerin % 75'i kendisinden küçük, % 25'i kendisinden büyük değerli olan noktadır. İkinci çeyrek ise (Q_2) % 50'si kendisinden küçük % 50'si kendisinden büyük değerli olan noktadır ki, bu ortancaya karşılık gelir.

Örnek 3.10: Küçükten büyüğe sıralanmış veri dizisi,

2 5 8 10 11 14 17 20

olsun. Birinci ve üçüncü çeyrek değerleri hesaplayalım.

Veri kümesi için Q_1, $\frac{1}{4}(n+1)$'inci değere karşılık gelir. $\frac{1}{4}(n+1) = \frac{9}{4} = 2{,}25$

2,25’inci değer, 2’inci ile 3’üncü değer arasındadır. 2’inci değer 5, 3’üncü değer 8’dir. 5 ile 8 arasındaki uzaklık dört eşit parçaya ayrılır ve buradan Q_1 aşağıdaki biçimde elde edilir.

$$Q_1 = 5 + \frac{1}{4}(8-5) = 5{,}75$$

Benzer şekilde üçüncü çeyrek Q_3 için, $\frac{3}{4}(n+1) = \frac{3}{4}(8+1) = 6{,}75$’inci değer Q_3 olacaktır. 6,75 verinin 6’ıncı ve 7’inci değerleri arasında olup, 6’ıncı değer 14 ve 7’inci değer 17’dir. Aradaki farkın ¾’ü (%75), 6’ıncı gözlemin aldığı değerin (14’ün) üzerine eklenerek aşağıdaki biçimde elde edilir.

$$Q_3 = 14 + \frac{3}{4}(17-14) = 16{,}25$$

Sınıflandırılmış verilerde, birinci ve üçüncü çeyrek değerler birikimli sıklık değerleri kullanılarak sırasıyla $\frac{1}{4}(n+1)$. ile $\frac{3}{4}(n+1)$. noktanın değeri olarak verilir. Tablo 3.6’da verilen sıklık dağılımından yararlanarak verilere ilişkin birinci ve üçüncü çeyrek değerleri bulalım.

Alt Sınır-Üst Sınır	fi	...Ve Az
8-10	22	22
11-13	15	37
14-16	25	62
17-19	24	86
20-22	6	92
23-25	5	97
26-28	3	100

$\frac{1}{4}(n+1) = \frac{1}{4}(100+1) = 25{,}25$. gözlem değeri birinci çeyrek değere karşılık gelir. 25,25. gözlem değeri, 22 ve 37 arasına düşer. Doğrusal öteleme kullanılarak birinci çeyrek değer,

10.............. 22

Q_1............. 25,25

13.............. 37

$$\frac{25{,}25-22}{37-22} = \frac{Q_1-10}{13-10} \Rightarrow Q_1=10{,}65$$

olarak bulunur.

$\frac{3}{4}(n+1)=\frac{3}{4}(100+1)=75,75$. gözlem değeri üçüncü çeyrek değere karşılık gelir.

75,75. gözlem değeri, 62 ve 86 arasına düşer. Doğrusal öteleme kullanılarak üçüncü çeyrek değer,

$$\begin{array}{ll} 16 \ldots\ldots\ldots\ldots\ 62 & \\ Q_3 \ldots\ldots\ldots\ldots\ 75,75 & \dfrac{75,75-62}{86-62}=\dfrac{Q_3-16}{19-16} \Rightarrow Q_3=17,71 \\ 19 \ldots\ldots\ldots\ldots\ 86 & \end{array}$$

olarak bulunur.

Ağırlıklı Ortalama

Bazı verilerde önem derecesi bakımından farklar olabilir. Ortalama hesaplanmasında bu farkların etkili olması isteniyorsa, önem derecesi ile orantılı olarak bir ağırlık verilir. Bu ağırlıklar kullanılarak hesaplanan ortalamaya 'Ağırlıklı Ortalama' denir ve

$$\overline{X}_A=\sum W_i X_i \quad , \quad W_i=\frac{f_i}{n} \quad , \quad \sum W_i=1 \qquad (3.5)$$

ile elde edilir. Burada, X_i, i. denek değeri; f_i, X_i değerini alan gözlem sayısıdır.

Ağırlıklı ortalamanın kullanım alanı çok geniştir. Bileşik indekslerde, ortalama maliyet ve satış fiyatlarının saptanmasında, mevduat ve kredilerde ortalama sürenin hesabında, faiz haddinin yıllık ortalamasının hesaplanmasında kullanılır.

Örnek 3.11: Değişik yaş grubundaki 186 kişinin bir yılda ortalama sinemaya gidiş sayısı tablodadır. 186 kişi için yıllık ortalama sinemaya gidiş sayısı nedir?

Tablo 3.4 Yaş Gruplarına Göre Ortalama Sinemaya Gidiş Sayıları

Yaş Grupları	Sıklık	Ortalama Sinemaya Gidiş Sayısı	W_i
15-20	50	6	50/186
21-35	68	7	68/186
36-45	40	5	40/186
46 ve üstü	28	5	28/186

$$\overline{X}_A=\frac{50}{186}6+\frac{68}{186}7+\frac{40}{186}5+\frac{28}{186}5=5,9998$$

Bir yılda 186 kişinin ortalama olarak 6 kez sinemaya gittiği söylenebilir.

Geometrik Ortalama

Geometrik ortalama gelişme ve büyüme hızı, bileşik faiz, endeks hesaplamalarında kullanılan ortalamadır. Biyoloji ve tıpta herhangi bir olayın zaman akışı içersinde, nispeten sabit bir oranda değiştiği durumlarda da kullanılır. Dağılım sınırları, geometrik ortalama sonucunu aritmetik ortalamadan daha az etkiler. Herhangi bir gözlem değeri sıfır olduğunda geometrik ortalama hesaplanamaz. Geometrik ortalama aşağıda verilen eşitlik ile elde edilir.

$$GO = \sqrt[n]{X_1 \times X_2 ... \times X_n} \tag{3.6}$$

Gözlem sayısı büyük olduğunda, bu formül ile hesaplamaların yapılması güç olduğundan geometrik ortalama Eşitlik (3.7) yardımıyla hesaplanır:

$$\log GO = \frac{1}{n}(\log X_1 + \log X_2 + ... + \log X_n) \tag{3.7}$$

Örnek 3.12: 2, 4, 5, 8 sayılarının geometrik ortalamasını bulunuz.

$$GO = \sqrt[4]{2 \times 4 \times 5 \times 8} = 4{,}23$$

Verilere ilişkin geometrik ortalama yaklaşık olarak 4,23'dür.

Örnek 3.13: Temmuz 2005 yılında bir manavda meyvelerin perakende fiyatları aşağıda verilmiştir. Verilere göre ortalama meyve fiyatını bulunuz.

Tablo 3.5 Meyve Perakende Fiyatları

Meyve Adı	kg fiyatı (YTL)	logX
Elma	1,500	0,1761
Armut	1,250	0,0969
Üzüm	0,950	-0,022
Şeftali	1,000	00000
Portakal	0,950	-0,022
Muz	2,450	0,3891

$$\log GO = \frac{1}{6}(0{,}1761 + 0{,}0969 - 0{,}022 + 0 - 0{,}022 + 0{,}3891) = 0{,}1030$$

GO=1,2677'dır. Manavda satılan meyvelerin ortalama fiyatı yaklaşık olarak 1,2677 YTL'dir.

Örnek 3.14: Bir malın toptan satış fiyatı birinci yıldan ikinci yıla %35, ikinci yıldan üçüncü yıla %48 oranında artmıştır. İncelenen iki yıl içinde ortalama artış ne kadardır?

İkinci yıl fiyatları birinci yılın %135, üçüncü yıl fiyatları da ikinci yıl fiyatının %148'idir. İki yıl içinde gözlenen ortalama artış geometrik ortalama ile bulunur.

$$GO = \sqrt{1{,}35 \times 1{,}48} = 1{,}4135$$

Gözlenen iki yıl içinde malın toptan satış fiyatındaki artışı % 41,35'dir.

Harmonik Ortalama

Harmonik ortalama; ortalama hız, ortalama fiyat, üretim ve verim hesaplamalarında kullanılır. Genel olarak, doğrudan doğruya ifade ettiği anlam yerine, tersine çevrildiğinde taşıyacağı anlama önem verilen, oran niteliğindeki değerlerin ortalamasının hesaplanmasında harmonik ortalamadan yararlanılır. Herhangi bir gözlem değeri sıfır olduğunda hesaplanamaz. Harmonik ortalamaya ilişkin eşitlik;

$$HO = \frac{n}{\sum_{j=1}^{n} \frac{1}{X_j}} \tag{3.8}$$

ile verilir.

Örnek 3.15: Ayşe bir işi 15 dakika, Ayla aynı işi 17 dakika, Melek 15 dakika ve Esin 18 dakika da bitiriyor. Bu işin bitirilme süresi ortalama kaç dakikadır?

$$HO = \frac{4}{\frac{1}{15} + \frac{1}{17} + \frac{1}{15} + \frac{1}{18}} = 16{,}26 \text{ dak}$$

Örnek 3.16: Bir fabrika ürettiği arabaların belli bir uzaklık için yapabilecekleri hızı test etmek istiyor. Arabalardan birini seçiyor. Araba ilk 50 km'yi 90 km/saat hızla, 60 km'yi 80 km/saat hızla ve son 40 km'yi 70 km/saat hızla gidiyor. Arabanın ortalama hızını bulunuz.

$$HO = \frac{150}{\frac{50}{90} + \frac{60}{80} + \frac{40}{70}} = 80{,}21 \text{ km / saat}$$

Arabalar ortalama olarak yaklaşık 80 km/saat hız yapabilmektedir.

Not: Ortalamalar arasında, **aritmetik ortalama>geometrik ortalama>harmonik ortalama** bağıntısı vardır.

3.3. NİTEL VERİLERDE KONUM ÖLÇÜLERİ

Nitel verilerde ortalama yerine yüzdelikler elde edilir. Yüzdelik, örneklemi belirten istatistiklerden biridir. Sınıflardan biri ele alınarak, alınan sınıf sıklığının sıklık toplamına yüzdesi hesaplanır. Bu ifade daha önce Eşitlik (2.4) ile verilmiş olan göreli sıklıktır.

Tablo 2.7'de verilen örnekte Bodrum tatil bölgesini tercih eden turistlerin yüzdesi Eşitlik (2.4)'den,

$$p = \frac{280}{1000} = 0,28$$

olarak hesaplanır.

3.4. NİCEL VERİLERDE DEĞİŞİM ÖLÇÜLERİ

İstatistikte aldığımız bilgiler, iki farklı kaynaktan gelen verilerin, dağılımını karşılaştırmamızda yardımcı olur. Bir önceki konuda gördüğümüz konum ölçüleri bu karşılaştırma için yeterli olmaz. Örneğin iki farklı hisse senedinin yedi aylık fiyatları her ay için ortalama olmak üzere aşağıdaki gibi verilsin.

Tablo 3.6 A ve B Hisse Senetlerinin Aylar İtibarıyla Fiyatları

Aylar	A Hisse Senedi YTL	B Hisse Senedi YTL
Ocak	1,7	3,4
Şubat	3,5	3,6
Mart	4,6	3,5
Nisan	3,1	3,4
Mayıs	3,5	3,5
Haziran	4,1	3,6
Temmuz	4,0	3,5
Ortalama	**3,5**	**3,5**
Ortanca	**3,5**	**3,5**

Bu iki hisse senedinin ortalamaları ve ortancaları aynıdır, ancak "A ve B hisse senetleri aynı özelliktedir, hangisini isterseniz tercih edin" diyebilir misiniz? Bu hisse senetlerinden hangisine yatırım yapmayı tercih edersiniz? Niçin? Bu sorulara verilecek cevap, A hisse senedinin değişkenliğinin daha fazla olması, B'nin daha tutarlı olması sebebiyle, B'nin tercih edileceğidir. Burada yeni bir ölçüm dikkate alınmıştır. Bu ölçüm "veriler arasındaki değişimden kaynaklanan farklılıkların ölçümü" olacaktır. İstatistikte bu ölçümlere değişim ölçüleri denir. Sık kullanılan değişim ölçüleri aşağıda incelenecektir.

Değişim Genişliği

En basit değişim ölçüsü, değişim genişliğidir (range). Değişim genişliği (DG), bir dağılımda en büyük değerli veri ile en küçük değerli verilerin değerleri arasındaki farktır. Ortalamaları eşit olan iki dağılımın karşılaştırılmasında kullanılabilir. Ancak, dağılımda bulunan uç değerlerden çok etkilendiği için eşit gözlem sayısı olmayan iki dağılımın karşılaştırılmasında yanıltıcı olabilir.

Örnek 3.17 : Tablo 3.6'da verilen A ve B hisse senetleri için sırasıyla değişim genişlikleri,

$DG_A = 4,6–1,7=2,9$

$DG_B = 3,6–3,4=0,2$

olacaktır. B hisse senedinin değişimi daha az olduğundan bu hisse senedi diğerine tercih edilir.

Örnek 3.18: A ve B planları ile üretilen işin ard arda 5 günde kaç birim üretildiğine ilişkin bir tablo aşağıda verilmiştir:

Tablo 3.7 A ve B Planlarına İlişkin Üretim Birimleri

Gün	A Planı	B Planı
1	15	23
2	25	26
3	35	25
4	20	24
5	30	27

Burada değişim genişlikleri,

A planı için = 35 – 15 = 20

B Planı için = 27 – 23 = 4

olup, B planının değişimi A planına göre küçük olduğundan tercih edilir.

Çeyrek Sapma

Değişim genişliğinin uç değerlerden etkilendiği belirtildi. Çeyrek sapma, bu sakıncayı ortadan kaldırmak için kullanılan bir ölçüdür. Bu ölçü, üçüncü çeyrek değer (Q_3) ile birinci çeyrek değer (Q_1) arasındaki farkın yarısıdır ve Q ile gösterilir.

$$Q = \frac{Q_3 - Q_1}{2} \tag{3.9}$$

Çeyrek sapma hesaplamasında tüm değerler kullanılmadığı için, yeterli bir değişim ölçüsü olarak kabul edilmez.

Örnek 3.19: Bir dişli fabrikasında günlük üretim, ortalama 1000'dir. Aşağıda ard arda 30 gün içinde hatalı üretim sayıları verilmiştir.

1, 0, 0, 2, 1, 5, 4, 0, 1, 2, 2, 5, 3, 2, 4, 10, 2, 6, 5, 3, 8, 9, 10, 5, 4, 3, 5, 3, 2, 6

Verilere ilişkin sıralanmış dağılım;

0, 0, 0, 1, 1, 1, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 4, 4, 4, 5, 5, 5, 5, 5, 6, 6, 8, 9, 10, 10

olur. Bu durumda, $Q_1 = 2$ ve $Q_3 = 5$ olmak üzere çeyrek sapma,

$$Q = \frac{Q_3 - Q_1}{2} = \frac{5-2}{2} = 1{,}5$$

olacaktır.

Ortalama Mutlak Sapma

Mutlak sapma bir veri kümesinde, değerlerin ortalamadan farklarının mutlak değerlerine ait ortalamadır. MS ile gösterilir ve aşağıdaki eşitlik ile verilir:

$$MS = \frac{\sum_{j=1}^{n} \left| X_j - \overline{X} \right|}{n} \tag{3.10}$$

Mutlak değer almak hesaplamalarda zorluk çıkardığı için, sık kullanılan bir değişim ölçüsü değildir.

Örnek 3.20: Tablo 3.6'da verilen A, B hisse senetlerine ait mutlak sapma ölçülerini bulunuz ve karşılaştırınız.

A Senedi	B Senedi
|1,7-3,5|=1,8	|3,4-3,5|=0,1
|3,5-3,5= 0	|3,6-3,5|=0,1
|4,6-3,5|=1,1	|3,5-3,5|=0
|3,1-3,5|=0,4	|3,4-3,5|=0,1
|3,5-3,5|=0	|3,5-3,5|=0
|4,1-3,5|=0,6	|3,6-3,5|=0,1
|4,0-3,5|=0,5	|3,5-3,5|=0
MS= 4,4/7=0,6286	**MS= 0,4/7=0,0571**

MS değeri A Senedinde daha büyük olduğundan, B hisse senedi tercih edilmelidir.

Kitle Varyansı, Örneklem Varyansı ve Standart Sapma

a) Kitle Varyansı ve Standart Sapması

Tüm verilerin alındığı bir çalışmada, verilerin değişim ölçüsü için varyans, en sık kullanılan değişim ölçüdür. Her bir verinin kitle ortalamasından farklarının karesinin toplamının, kitle büyüklüğüne bölümüdür. Farkların kareleri alındığı için, birimi de kare olacaktır. Standart sapma ise, varyansın kareköküdür. Buradaki birim verilere ait birimdir. Verilerin ortalamalarından uzaklıklarının karelerine ait ortalama olan kitle varyansı (σ^2) ve varyansın karekökü olan standart sapma (σ) sırasıyla aşağıda verilmiştir:

$$\sigma^2 = \frac{\sum_{j=1}^{N}(X_j - \mu)^2}{N} = \frac{\sum X_i^2}{N} - \mu^2 \tag{3.11}$$

$$\sigma = \sqrt{\frac{\sum_{j=1}^{N}(X_j - \mu)^2}{N}} \tag{3.12}$$

Burada,

X_J: j. verilerin değeri;

μ: Kitle ortalaması;

N: Kitle gözlem sayısı

Örnek 3.21: Aşağıdaki verilere ilişkin varyansı ve standart sapmayı bulunuz.

$X_1 = 15$, $X_2 = 25$, $X_3 = 35$, $X_4 = 20$, $X_5 = 30$

$$\mu = \frac{\sum_{j=1}^{N} X_j}{N} = \frac{15+25+35+20+30}{5} = \frac{125}{5} = 25$$

X	**X-μ**	**(X-μ)²**
15	15-25 =-10	$(15-25)^2 = (-10)^2 = 100$
25	25-25 = 0	$(0)^2 = 0$
35	35-25 = 10	$(10)^2 = 100$
20	20-25 = -5	$(-5)^2 = 25$
30	30-25 = 5	$(5)^2 = 25$

Kitle varyansı,

$$\sigma^2 = \frac{\sum_{j=1}^{N} (X_J - \mu)^2}{N} = \frac{250}{5} = 50 \ \text{(birim)}^2$$

kitle standart sapması,

$$\sigma = \sqrt{50} = 7{,}07 \text{ birim}$$

olarak elde edilir.

b) Örneklem Varyansı ve Standart Sapması

Sınıflandırılmamış (Sıklık Dağılımı Düzenlenmemiş) Verilerde Varyans ve Standart Sapmanın Hesaplanması

Örneklem, kitleden örnekleme yöntemleriyle seçilen, kitleyi temsil eden küçük topluluktur. Kitleye ait bilgiler, örneklemden yararlanarak tahmin edilir. Eşitlik (3.11) ile kitle için verilen varyans formülü, örneklemden hesaplanıyor ise bu formül bir tahmin edici formülü olacaktır. Tahmin edicinin, kitle bilgisini doğru verebilmesi için bazı özelliklere sahip olması gerekir. Bu özellikler daha sonraki konularda ele alınacaktır. Örneklemden hesaplanan varyans (S^2),

$$S^2 = \frac{\sum_{j=1}^{n} (X_j - \overline{X})^2}{n-1} = \frac{\sum X_i^2 - n\overline{X}^2}{n-1} \qquad (3.13)$$

eşitliği ile bulunur. Bu eşitlikte;

X_j: j. verinin değeri;

$\overline{X}$: örneklem ortalaması;

n: örneklem büyüklüğüdür.

Bu eşitliğin kitle varyansından farkı, paydasında N yerine, (n-1) olmasıdır. Örnekleme ilişkin standart sapma (S),

$$S=\sqrt{\frac{\sum_{j=1}^{n}\left(X_j-\overline{X}\right)^2}{n-1}} \qquad (3.14)$$

eşitliği ile bulunur.

Örnek 3.22: Tablo 3.6'da verilen A ve B hisse senetleri için, örnekleme ait varyansı ve standart sapmayı hesaplayınız.

$$S_A^2=\frac{(1{,}7-3{,}5)^2+(3{,}5-3{,}5)^2+\cdots\cdots+(4-3{,}5)^2}{7-1}=\frac{5{,}22}{6}=0{,}87$$

$$S_B^2=\frac{(3{,}4-3{,}5)^2+(3{,}6-3{,}5)^2+\cdots\cdots+(3{,}5-3{,}5)^2}{7-1}=0{,}0067$$

$$S_A=\sqrt{0{,}87}=0{,}9327 \quad \text{ve} \qquad S_B=\sqrt{0{,}0067}=0{,}0819$$

B hisse senedinin varyansı, A hisse senedine göre daha küçük olduğundan B hisse senedine yönelmek daha akıllıca olabilir.

Sınıflandırılmış (Sıklık Dağılımı Düzenlenmiş) Verilerde Varyans ve Standart Sapmanın Hesaplanması

Sınıflandırılmış verilerde varyans ve standart sapma hesaplayabilmek için, önceden oluşturulmuş sıklık dağılımı kullanılır. Bu durumda örneklem varyansı için,

$$S^2=\frac{\sum_{i=1}^{k}\left(S_i-\overline{X}\right)^2\times f_i}{n-1}=\frac{\sum f_i S_i^2-n\overline{X}^2}{n-1} \qquad (3.15)$$

eşitliği kullanılır. Bu eşitlikte;

S_i: i. sınıfın sınıf değeri,

f_i: i. sınıfın sıklığı,

$\overline{X}$: örneklem ortalaması,

n: örneklem büyüklüğüdür.

Sıklık dağılımı düzenlenmiş verilerde örneklem için standart sapma aşağıda verilmiştir.

$$S=\sqrt{\frac{\sum_{i=1}^{k}\left(S_i-\overline{X}\right)^2\times f_i}{n-1}} \qquad (3.16)$$

Örnek 3.23: Aşağıda sıklık dağılımı düzenlenmiş veriler için varyans ve standart sapmayı hesaplayınız.

Alt Sınır-Üst Sınır	**S_i**	**f_i**	**$f_i \times S_i$**	**$S_i - \overline{X}$**	**$(S_i - \overline{X})^2 \times f_i$**
5-7	6	1	6	(6-13,8)= -7,8	60,84
8-10	9	2	18	(9-13,8)= -4,8	46,08
11-13	12	5	60	(12-13,8)= -1,8	16,2
14-16	15	6	90	(15-13,8)= 1,2	8,64
17-19	18	3	54	(18-13,8)= 4,2	52,92
20-22	21	1	21	(21-13,8)= 7,2	51,84
		$\sum_{i=1}^{k} f_i = 18$	$\sum_{i=1}^{k} f_i \times S_i = 249$		$\sum_{i=1}^{k} (S_i - \overline{X})^2 \times f_i = 236,52$

Verilere ilişkin ortalama, varyans ve standart sapma sırasıyla aşağıda verilmiştir.

$\overline{X}$ =249/18=13,8 , S^2=236,52/(18-1)=13,91 ve S=3,73

Örnek 3.24: Örnek 3.2'de verilen notlara ilişkin, standart sapmayı sıklık dağılımından yararlanarak bulunuz.

Alt Sınır-Üst Sınır	**f_i**	**Si**	**$(S_i - \overline{X})^2$**	**$(S_i - \overline{X})^2 \times f_i$**
36-42	1	39	501,76	501,76
43-49	3	46	237,16	711,48
50-56	9	53	70,56	635,04
57-63	5	60	1,96	9,8
64-70	4	67	31,36	125,44
71-77	4	74	158,76	635,04
78-84	4	81	384,16	1536,64

$$\sum_{i=1}^{k} (S_i - \overline{X})^2 \times f_i = 4155,2$$

$$S = \sqrt{\frac{\sum_{i=1}^{k} \left(S_i - \overline{X}\right)^2 \times f_i}{n-1}} = \sqrt{\frac{4155,2}{30-1}} = 11,97$$

Örnekleme ilişkin standart sapma 11,97'dir.

Nicel Verilerde Standart Hata

Örneklem ortalamasının oluşturduğu dağılımın standart sapması, örneklem ortalamalarından her birinin standart hatası sayılır. Buna göre bir örneklemin ortalamasının standart hatası ($S_{\bar{x}}$)

$$S_{\bar{x}} = \frac{S}{\sqrt{n}} \tag{3.17}$$

eşitliği ile verilir.

Değişim Katsayısı

Farklı birimlerde ölçülmüş verilerin değişimlerini (yayılımlarını, dağılımlarını) karşılaştırmak için kullanılan değişim katsayısı, birimden bağımsız değişim ölçüsüdür. Standart sapmanın ortalamaya olan yüzdesidir ve yüzde ile gösterilir. Değişim katsayısı (CV, V) kitle için,

$$CV = V = \frac{\sigma}{\mu} \times 100 \tag{3.18}$$

ile örneklem için,

$$CV = V = \frac{S}{\overline{X}} \times 100 \tag{3.19}$$

ile elde edilir.

Örnek 3.25: İki ayrı sınıfta Matematiksel İstatistik I sınav sonuçları sırası ile 10 ve 100 üzerinden değerlendirilmiş ve bu iki sınıfa ait bilgiler aşağıda verilmiştir:

	(A Sınıfı)	(B Sınıfı)
$\overline{X}$	6	75
S	1	5

Bu iki sınıfın değişim katsayılarını karşılaştırınız.

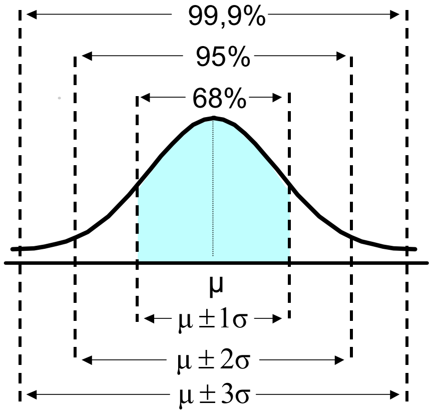
$$CV_A = \frac{1}{6} \times 100 = 16{,}7$$

$$CV_B = \frac{5}{75} \times 100 = 6{,}7$$

A sınıfının değişim katsayısı %16,7, B sınıfının değişim katsayısı %6,7’dir. $CV_A > CV_B$ olduğundan, B sınıfının daha az değişim gösterdiği, daha istikrarlı olduğu söylenebilir.

Nicel Verilerde Örneklem Dağılımı İle Standart Normal Dağılımın Standart Sapmalar İle Karşılaşması

Nicel veriler genellikle çana benzeyen bir dağılım gösterir. Bu dağılıma normal dağılım adı verilir. Normal dağılım ortalama, tepe değeri ve ortanca değeri eşit olduğunda simetrik bir dağılımdır. (Bakınız Şekil 3.1b). Normal dağılım ND($\mu;\sigma^2$) ile gösterilir. Burada μ, örneklemin çekildiği kitlenin ortalamasını, σ^2 örneklemin çekildiği kitlenin varyansını gösterir. Örneklemden elde edilen her değer için, $Z = \frac{X - \overline{X}}{S}$ dönüşümü uygulandığında, standart normal dağılım denilen ortalaması 0, varyansı 1 olan bir normal dağılım elde edilir. Standart normal dağılımda, verilerin % 68'i $\mu \pm 1\sigma$ sınırları içerisinde, %95'i $\mu \pm 2\sigma$ sınırları içerisinde, %99,9'u $\mu \pm 3\sigma$ sınırları içerisinde ve % 100'ü $\mu \pm 4\sigma$ sınırları içerisinde bulunur (Şekil 3.2).



Şekil 3.2 Ortalamaya Göre Belirli Standart Sapma Sınırları Arasında Kalan Normal Eğri Alanları

Dağılımın simetrikliğinin ve yüksekliğinin bozulması bazı katsayılar ile bulunabilir. Bu katsayılar aşağıda verilmiştir.

Çarpıklık ve Basıklık

Dağılım hakkında bilgi almanın diğer bir yolu da çarpıklık katsayısına bakmaktır. Dağılımlar arasında çarpıklık karşılaştırmaları yapabilmek için değişim katsayısında olduğu gibi, ölçü birimlerinin etkisinden bağımsız bir katsayıya, çarpıklık katsayısına gereksinim vardır. Örneklem için çarpıklık katsayısı,

$$ÇK = \frac{\sum_{j=1}^{n}\left(X_j - \overline{X}\right)^3 / n}{S^3} \qquad (3.20)$$

ile elde edilir. Burada ,

$\overline{X}$: Aritmetik ortalama,

S :Standart sapmadır.

ÇK=0 ise, dağılımın ortalamaya göre simetrik; ÇK>0 ise, dağılımın sağa doğru çarpık ya da (+) yöne eğilimli; ÇK<0 ise, dağılımın sola doğru çarpık ya da (-) yöne eğilimli olduğu söylenebilir (Bakınız Şekil 3.1).

Bir dağılımın normal dağılıma uygunluğu sadece simetrik olması ile ölçülemez. Ayrıca dağılımın yüksekliğinin de normal dağılıma uygun olması gerekmektedir. Bir dağılımın normalden dik ya da basık dağılım olarak nitelendirebilmesi için aşağıda verilen örneklem basıklık katsayısının hesaplanması gerekmektedir.

$$BK = \frac{\sum_{j=1}^{n}\left(X_j - \overline{X}\right)^4 / n}{S^4} \qquad (3.21)$$

BK=3 ise, dağılımın yüksekliği standart normal dağılıma uygun; BK>3 ise dağılımın standart normal dağılımdan daha dik ya da daha sivri; BK<3 ise dağılımın standart normal dağılımdan daha basık olduğu söylenebilir.

Örnek 3.26: Aşağıdaki verilere ilişkin çarpıklık ve basıklık katsayılarını bulunuz ve yorumlayınız.

3 4 4 5 6 8

Verilere ilişkin ortalama,

$$\overline{X} = \frac{\sum_{j=1}^{n} X_j}{n} = \frac{3+4+4+5+6+8}{6} = \frac{30}{6} = 5$$

standart sapma,

$$S = \sqrt{\frac{\sum_{j=1}^{n}\left(X_j - \overline{X}\right)^2}{n-1}} = \sqrt{\frac{(3-5)^2 + (4-5)^2 + (4-5)^2 + (5-5)^2 + (6-5)^2 + (8-5)^2}{6-1}} = \sqrt{3,2} \cong 1,79$$

olarak elde edilir. Çarpıklık katsayısı,

$$ÇK=\frac{\sum_{j=1}^{n}\left(X_j-\overline{X}\right)^3/n}{S^3}=\frac{\left[(3-5)^3+(4-5)^3+(4-5)^3+(5-5)^3+(6-5)^3+(8-5)^3\right]/6}{1{,}79^3}\cong 0{,}5$$

ve basıklık katsayısı,

$$BK=\frac{\sum_{j=1}^{n}\left(X_j-\overline{X}\right)^4/n}{S^4}=\frac{\left[(3-5)^4+(4-5)^4+(4-5)^4+(5-5)^4+(6-5)^4+(8-5)^4\right]/6}{1{,}79^4}\cong 1{,}6$$

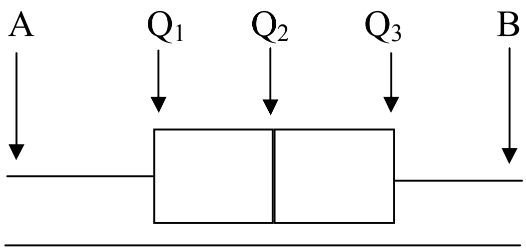
dır.

Yorum: ÇK>0 olduğu için, veriler sağa doğru çarpık ya da (+) yöne eğilimlidir. BK<3 olduğundan verilere ilişkin dağılımın standart normal dağılımdan daha basık olduğu söylenebilir.

Sapan Değer ve Kutu Çizimi

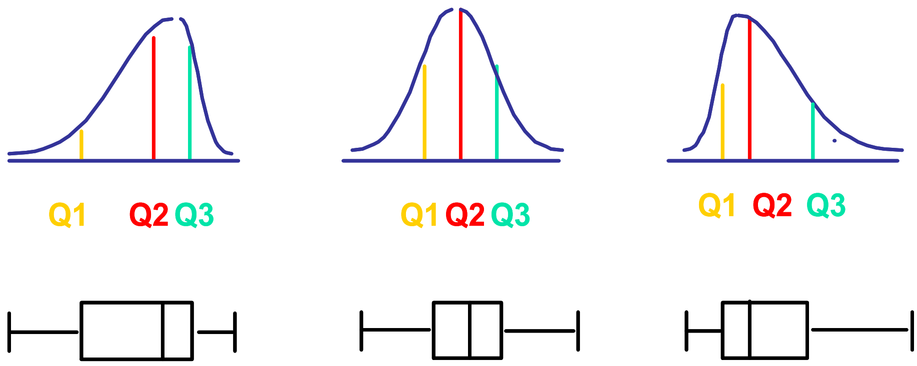
Sapan değer, dağılımdaki gözlem değerlerinden farklılık gösteren uç değer ya da uç değerlerdir. Sapan değerin hesaplamaya alınıp alınmamasına göre farklı ortalamalar bulunur. Kitlede uç değerlerin örnekleme katılma olasılığı çok düşüktür. Örneklemde görülen sapan değer, bir rastlantı sonucunda örnekleme katılmış değer olabileceği gibi, ölçüm hatasından kaynaklanan, yazım hatasından kaynaklanan ya da başka bir kitleden yanlışlıkla katılan bir değer de olabilir. Bu farkı ortaya çıkarmak için kullanılan çeşitli yöntemler vardır. Bu yöntemlerden biri de kutu çizimidir. Kutu çizimleri, verilerin çeyrek değerlerine dayalı grafiklerdir. Sapan değeri bulabilmek için öncelikle Q_1, Q_2, (ortanca) ve Q_3 sırasıyla hesaplanır. Q_3 ve Q_1 arasındaki mesafe dikkate alınarak yatay bir kutu çizilir. Bu kutunun genişliği d ile gösterilir. Kutunun içinde, ortanca dikey bir çizgi ile belirtilir. Kutunun her iki ucundan $A=(Q_1-1{,}5d)$ ve $B=(Q_3+1{,}5d)$ uzunlukları şüpheli gözlem sınırları olarak belirlenir. A'dan daha küçük ve B'den daha büyük değerler sapan değerlerdir.

Kutu çizimleri gözlemlerin etkinlikleri ve dağılımlar için de kullanılır. Çeyrek değerlerin farkı olarak tanımlanan kutu uzunluğu, verilerin yayılım genişliğini göstermektedir. Bunun yanında dağılımın simetrikliği de görülebilir. Şekil 3.3'de görüldüğü gibi eğer dağılım simetrik ise ortanca, kutuyu tam olarak ikiye böler, kutunun her iki yanındaki doğrular eşit uzunluktadır ve her iki yanda kalan şüpheli değerler eşit sayıda olup simetrik görünümlüdür.



Şekil 3.3 Simetrik Kutu Çizimi

Kutu içinde işaretlenen ortancanın birinci ve üçüncü çeyrek değerlere göre yakınlığından, çarpıklığı görmek mümkündür. Ortanca birinci çeyrek değere daha yakın ise, veri kümesinin sağa doğru çarpık (+) yöne eğilimli, ortanca üçüncü çeyrek değere daha yakın ise, veri kümesinin sola doğru çarpık (-) yöne eğilimli olduğu söylenebilir. Şekil 3.4'de sırasıyla sola (a), simetrik (b) ve sağa (c) çarpık dağılımın grafikleri verilmiştir.



a) Sola çarpık b) Simetrik c) Sağa çarpık

Şekil 3.4 Çeşitli Kutu Çizimleri; Sola çarpık (a), Simetrik (b), Sağa çarpık (c)

Örnek 3.27: Örnek 3.10'daki veriler için,

a- Sapan değer olup olmadığını bulunuz.

b- Dağılım hakkında bilgi veriniz.

a-

Q_1=5,75 , Q_2=10,5 , Q_3=16,25 , d=16,25-5,75=10,5

A=(Q_1-1,5d)

=5,75-1,5.10,5=-10

$B=(Q_3+1,5d)$

$=16,25-1,5.10,5=32$

Yorum: Tüm değerler –10 ile 32 değerlerinin arasına düştüğü için dağılımda sapan değer yoktur.

b- Q_2-Q_1=10,5-5,75=4,75

Q_3-Q_2=16,25-10,5=5,75

Yorum: Ortanca değeri, Q_1 değerine daha yakın olduğundan verilerin sağa çarpık bir dağılım gösterdiği söylenebilir.

3.5. NİTEL VERİLERDE DEĞİŞİM ÖLÇÜSÜ

Nitel Verilerde Standart Sapma ve Standart Hata

Nitel verilerde standart sapma, yüzde değerlerinden yararlanılarak aşağıdaki eşitlik ile bulunur.

$$S_i = \sqrt{p_i\, q_i} \tag{3.22}$$

Burada S_i. i. sınıfın standart sapması, p_i, i. sınıfa düşen göreli sıklık değeri ve $q_i=1-p_i$'dir.

Nitel verilerde standart hata,

$$S_{\overline{X}} = \sqrt{\frac{pq}{n}} \tag{3.23}$$

eşitliği ile elde edilir.

Örnek 3.28: 2005-2006 Güz yarıyılında Matematiksel İstatistik I dersini alan 150 öğrencinin 93'ü kız, 57'si erkektir. Kız öğrencilere ilişkin standart sapmayı bulunuz.

Sınıflar	f_i	p_i
Kız	93	0,62
Erkek	57	0,38

$$S_1 = \sqrt{0,62 \times 0,38} = 0,48$$

PROBLEMLER

1- 4 4 6 7 7 9 10 10

Yukarıdaki verilere ilişkin

a- Aritmetik ortalama, tepe değeri ve ortancayı bulunuz.

b- Dağılım hakkında bilgi veriniz.

c- Birinci, ikinci ve üçüncü çeyrek değerleri elde ediniz.

2-

4	4	6	7	7	7	10	12

a- Aritmetik ortalama, tepe değeri ve ortancayı bulunuz.

b- Dağılım hakkında bilgi veriniz.

c- Çarpıklık ve basıklık katsayılarını bulunuz ve yorumlayınız.

3- Bir ekmek fabrikası hamur tartan aletini kontrol etmek için bir günlük üretimde alet ile tartılmış ekmek hamurlarından rasgele 40 ekmek seçip yeniden ağırlıklar belirlemiştir. 350 gr.lık ekmek elde etmek için hamurlar 400 gr. tartılmaktadır. İşlemlerde kolaylık sağlamak için verilerden 400 gr. çıkartılarak elde edilen bilgiler aşağıda verilmiştir.

-6	-3	-1	0	1	1	3	3	4	6
-6	-3	-1	0	1	1	3	3	4	6
-4	-2	0	1	1	1	3	3	4	6
-4	-2	0	1	1	3	3	3	5	6

a- Sınıflandırılmamış verilerden ortalama, ortanca ve tepe değerini bulunuz.

b- Yukarıda verilen verileri kullanarak 5 ve 8 sınıflı sıklık dağılımlarını oluşturunuz.

c- Farklı sınıflar için oluşturduğunuz sıklık dağılımlarını kullanarak aritmetik ortalama ve varyansları bulunuz.

d- Farklı sınıflandırma yapılan sıklık dağılımlardan hesaplanan aritmetik ortalama ve varyanslar arasındaki farkların nedenlerini tartışınız.

4- 30 öğrencisi olan bir sınıfın Matematiksel İstatistik I dersinden 30 üzerinden aldıkları notlar aşağıda verilmiştir. Bu sınıfı bir kitle olarak düşününüz ve ortalamayı, tepe değerini, ortancayı bulunuz.

2	2	3	4	6	7	7	8	8	9
10	10	10	10	11	11	11	12	12	12
13	13	14	16	18	20	25	25	28	28

a- Bu kitleden rasgele örnekleme ile 15 öğrenci seçiniz, yukarıda hesapladığınız konum ölçülerini yeniden hesaplayarak karşılaştırma yapınız.

b- 7 sınıflı sıklık dağılımını düzenleyiniz. Aritmetik ortalamayı ve varyansı bulunuz.

c- Sınıflandırılmamış verilerden hesapladığınız ortalama ile b şıkkından bulunan sonucu karşılaştırınız.

5- Aşağıdaki cümlelerin istatistik bilgileriniz doğrultusunda doğru ya da yanlış olduklarını söyleyiniz.

a- Sağa çarpık bir dağılımda $\overline{X}' < \overline{X}$ eşitsizliği vardır.

b- Bir örneklem verileri bir kez 8 sınıflı, bir kez de 10 sınıflı olacak şekilde özetlenmiştir. 8 sınıflı sıklık dağılımından bulunan ortalama daha gerçeğe yakındır.

c- Nicel veriler de konum ölçüleri ortalama, ortanca tepe değeri, çeyrek değerlerdir.

d- Bir örneklem dağılımında en çok tekrarlanan değer ortalamadır.

6- Daktilo sekreterliğine başvuran 80 kişinin yazmaları için verilen bir metni bitirme süreleri dakika olarak kaydedilmiştir. Sınıf değerleri ve sıklıkları verilen sıklık dağılımında, aşağıdaki sorulara cevap veriniz.

S_i:	5	10	15	20	25	30	35
f_i:	4	5	15	28	21	6	1

a- Alt ve üst sınır kolonunu, göreli sıklık kolonunu, ...ve az ile ...ve çok kolonlarını bulunuz.

b- Tepe değerinin hangi sınırlar arasında olmasını beklersiniz?

c- 20 kişi işe alınabileceğine göre, kaç dakikadan az sürede metni tam bitirenler alınacaktır?

d- Ortancayı hesaplayınız.

e- Aritmetik ortalamayı ve varyansı bulunuz.

7- 50 öğrencinin bulunduğu sınıfta öğrencilerin deneyi bitirme süreleri dakika olarak kaydedilmiş ve aşağıdaki sıklık dağılımı verilmiştir.

Deneyi bitirme süresi	**Öğrenci sayısı**
As-Üs	fi
0-1	5
2-3	10
4-5	20
6-7	12
8-9	3

a- Verilere ilişkin ortalama, tepe değeri, varyans ve standart sapmayı bulunuz.

b- Öğrencilerin % 50'si kaç dakikadan az sürede deneyi tamamlamaktadır. Bu sonuç konum ölçülerinden hangisine karşılık gelmektedir.

8- 30 kişinin bir günde içtikleri sigara miktarları aşağıda verilmiştir. Sıklık dağılımını düzenleyiniz. Sınıflandırılmış ve sınıflandırılmamış verilerden bulunan değerler arasındaki farkın nedenini tartışınız.

5	15	20	8	20	12	10	10	12	15
18	20	3	2	6	10	20	3	15	8
20	15	15	20	5	3	8	20	2	21

9- Bir benzin istasyonu rasgele seçtiği 30 müşterisinin aldığı benzin miktarlarını litre olarak aşağıdaki gibi kaydetmiştir.

2	2	3	4	6	7	7	8	8	9
10	10	10	11	11	11	11	12	12	12
13	13	14	16	18	20	25	25	28	28

a- 7 sınıflı sıklık dağılımını düzenleyiniz.

b- Aritmetik ortalamayı ve standart sapmayı bulunuz.

c- Tepe değerini ve ortancayı bularak dağılımın yapısını tartışınız. Bu bilgi ile istasyondaki benzin satışları için ne söyleyebilirsiniz?

10- İki farklı deponun özellikleri karşılaştırmak isteniyor. A deposunun sıcaklık ortalaması 10^0 C ve standart sapması $0{,}7^0$ C'dir. B deposunun sıcaklık ortalaması 33^0 F ve standart sapması $0{,}3^0$ F'dir. Bu iki depo, sıcaklık değişimleri yönünden aynı kabul edilebilir mi?

11- 20 kişilik bir sınıfta öğrencilerin istatistik sınav notlarına ilişkin değerler aşağıda verilmiştir.

35	36	37	39	40	40	41	45	47	47
48	50	52	52	52	54	55	60	61	64

a- Yukarıdaki verilere ilişkin k= 6 olacak biçimde sıklık dağılımını düzenleyiniz.

b- Düzenlediğiniz sıklık dağılımından yararlanarak, aritmetik ortalama ve standart sapmayı bulunuz.

c- Sapan değer var mıdır? İnceleyiniz.

12- Aşağıdaki bilgiler göre,

a- Boş yerleri doldurunuz.

b- Ortalama, ortanca, tepe değeri ve varyansı bulunuz.

Alt sınır	Üst sınır	Si	fi	...Ve az	...Ve çok
	1,6		3		
	1,8		5		
	2,0		7		
	2,2		18		
	2,4		7		
	2,6		5		
	2,8		5		

13-

10	11	11	13	14	14	16	17	17	20
23	24	24	26	29	30	30	31	34	34

a- Yukarıdaki verilere ilişkin dağılım genişliğini bulunuz.

b- Mutlak sapma ve varyansı bulunuz.

c- Bulduğunuz değişim ölçülerini karşılaştırınız.

14- Bir sıklık dağılımında çarpıklık katsayısı -1,6, basıklık katsayısı sıfır olarak bulunmuştur. Dağılımın şekli için ne söyleyebilirsiniz?

15- Aşağıdaki verilere göre iki örneklemi karşılaştırınız.

	I Örneklem	II. Örneklem
Ortalama	4,7	6,3
Tepe Değeri	5,2	6,3
Varyans	0,25	1,3

16- Bir dağılımda ortalama 5, ortanca 6 ve tepe değeri 8 ise dağılımın şekli hakkında bilgi veriniz.

17- Bir örneklemde yer alan deneklerin ortalamadan ayrılış değerleri aşağıda verildiğine göre mutlak sapma ve varyansı hesaplayınız.

-1 4 -3 1 1 -2 1 2 2 -3

18- Tedavi olmak isteyen 200 hastanın 70'i göz polikliniğine, 30'u kulak burun boğaza, 80'i dahiliye ve geri kalan hastalar ortopediye başvurmuştur. Uygun sıklık dağılımını hazırlayıp nitel verilerde konum ölçüsünü soru üzerinde gösteriniz.

19- Farklı iki yerleşim bölgesinden rasgele seçilen 1 aylık, 30 ve 35 kız çocuğuna ait ağırlık dağılımları ile ilgili bilgiler aşağıda verilmiştir. Bu bilgilerden yararlanarak iki örneklemi karşılaştırınız.

	I. Bölge	II. Bölge
Ortalama	4200 gr	3,9 kg
Varyans	16 gr^2	0,49 kg^2

20- T.C. Devlet İstatistik Enstitüsü web sitesini kullanarak 2000 yılı nüfus sayımı bilgilerine ulaşınız. Buradan yaş dağılımı ile ilgili tabloyu bulunuz, tabloyu dikkate alarak en uygun konum ölçüsü ve değişim ölçüsünü hesaplayınız.

4. DÖRDÜNCÜ BÖLÜM

4.1. GİRİŞ

Örneklem, kitle hakkında bilgi elde etmek için kitleden çekilir, ancak kesin bir bilgi vermez. Örneğin müşteri eğilimini ölçmek ve talep hakkında bilgi elde etmek için bir anket uygulanırsa, sonuç talebe ilişkin bir bilgi oluşturur. Ancak kitle ile ilgili kesin bir sonuç olamaz. Kitleye ilişkin kesin bilgiler çıkarsamak olanaksız olmakla birlikte örneklemi temel alarak, belirsizliğin doğası konusunda doğru söylemler kullanılabilir. Bu tür söylemler, istatistikte temel önem taşıyan olasılık diliyle ifade edilir. Hava durumu tahminleri, satış tahminleri, yatırım kararları, yatırımcının gelecekteki olası getirileri olasılığa dayanır. Örneğin at yarışında oyuncular, yarışa giren atların eski performanslarını değerlendirerek tahmin yaparlar. Konunun uygulanma alanı oldukça geniş olmasıyla birlikte daha kolay anlaşılabilmesi için basit şans oyunları örnek olarak kullanılacaktır.

Olasılık belirsizliğin tartışıldığı dil olarak düşünülebilir. Olasılığı anlayabilmek için bu dilin kurallarını, tanımlarını öğrenmek gerekmektedir. Bu nedenle öncelikle Bölüm 4.2'de bu tanımlar verilecektir.

4.2. OLASILIK İÇİN GEREKLİ BAZI KAVRAMLAR

Rasgele Deney

Hangisinin gerçekleşeceği konusunda belirsizlik bulunan en az iki sonuca yol açan bir süreçtir. Para atılması, zar atılması, loto şans sayısının çekimi, pay senetleri fiyat indeksinin günlük değişimi, bir hesaplar kümesinde bir kalem hesabın denetlenmesi gibi olayların her birine rasgele deney denir.

Rasgele Sonuç

Rasgele deneylerde ortak özellik, hangi sonucun gerçekleşeceğinin önceden belirsiz olmasıdır. Bir para atılma deneyinde ya yazı ya da tura gelir. 6 yüzlü bir zarın fırlatılmasında 1,2,3,4,5,6 sonuçlarından biri gelir. Bir tüketici, ürünlerden birini beğenir ya da hiç birini beğenmeyebilir. Bunlardan her birine rasgele (olası) sonuç denir.

Örneklem Uzayı

Rasgele bir deneyin olası bütün sonuçlarına temel sonuçlar denir. Tüm olası sonuçları küme de örneklem uzayı (olay uzayı) adını alır. Örneklem uzayı S harfi ile gösterilir.

Olay

Bir olay, örneklem uzayında temel sonuçların bir alt kümesidir. Rasgele deneyde kendisini oluşturan olası sonuçlardan biri ortaya çıkarsa olay gerçekleşmiştir denir.

Bir zar atımında 1,2,3,4,5,6 sayılarının gelmesi birer olası sonuçtur. Bu olası sonuçların oluşturduğu küme

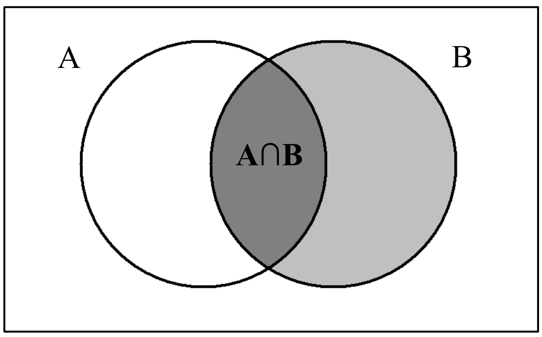
$$S:\{1,2,3,4,5,6\}$$

olan bir örneklem uzayıdır. Burada her denemede aynı anda birden fazla sonuç gerçekleşemez ancak biri zorunlu olarak gerçekleşir. Burada zarın çift gelmesi ile ilgileniliyor ise 2,4,6 sonuçları ele alınır. Olası sonuçların bir alt kümesi olan bu kümeye olay denir. Bir örneklem uzayında birden çok olay tanımı yapılabilir. Bazen bu olaylar aynı anda oluşabilir. Örneğin, bir rasgele deneyde, çift gelen zarlar ile en az 4 gelen temel sonuçlar bizi ilgilendirebilir. Burada iki olay vardır. Biri temel sonuçlardan çift gelme, diğeri zarın en az 4 gelmesidir. Bu iki olayın ortak olası sonucu olabilir. Bu sonuca arakesit denir.

Arakesit

S örneklem uzayında iki olay A ile B olsun. Bunların $A \cap B$ ile gösterilen arakesiti, S örneklem uzayında hem A'da, hem de B'de yer alan olası sonuçların alt kümesidir. Şekil 4.1'de ortak taralı bölge $A \cap B$'yi göstermektedir.

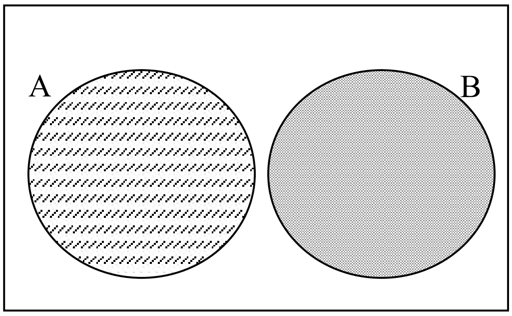
Daha genel olarak $E_1, E_2, \ldots, E_k$ gibi k tane olayın arakesiti $E_1 \cap E_2 \cap E_3 \ldots \cap E_k$ olur. Bu olay her bir E_i'de yer alan tüm olası sonuçların kümesidir. Bu duruma olayların çarpımı da denir.



Şekil 4.1 $A \cap B$'nin Gösterimi (Olayların Çarpımı)

Ayrık Olay

A ve B olaylarında ortak olan sonuçlar yoksa, A ve B'ye ayrık olaylar denir. A∩B arakesitinin, **boş küme (null set)** olduğu söylenir. Dolayısıyla A∩B gerçekleşmez. Boş küme ∅ ile gösterilir. Şekil 4.2, iki ayrık olayın gösterimidir.



Şekil 4.2 A∩B=∅'nin Gösterimi

Örnek 4.1: Bir banka denetlemesinde,

A olayı: {%5'inden az hata olan hesap kalemleri}

B olayı: {%10'undan çok hata olanları}

olsun. Bu durumda A ve B'nin ortak elemanı olmadığı için A∩B ayrıktır denebilir.

Örnek 4.2: Bir müşteri topluluğunda,

A olayı: {15-20 yaş grubu}

B olayı: {40-60 yaş grubu}

olarak tanımlanıyorsa, A∩B ayrıktır.

Birleşim

Bazen tanımlanan olaylardan en az birinin olması düşünülebilir. Örneğin,

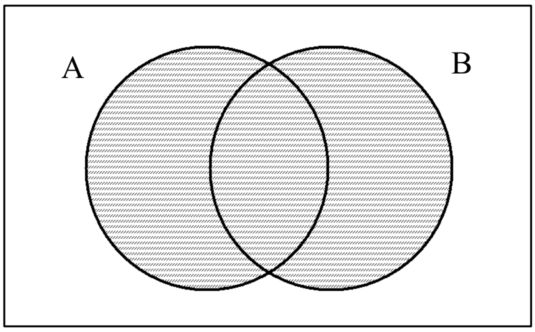
A olayı: {6 yüzlü zar atışında çift gelenler; 2,4,6}

B olayı: {6 yüzlü zar atışında en az 4 gelenler; 4,5,6}

olarak gösterilsin. A veya B olayı: çift olanlar veya en az 4 gelenler kümesidir. Bu durumda A ile B olayının birleşimi söz konusu olur. Bu iki olayın birleşimi A∪B ile gösterilir.

S örneklem uzayında iki olay A ve B olsun. Bunların A∪B ile gösterilen birleşimi (toplamı), A ve B olaylarından en az birinin ortaya çıktığı durumdur. Bu durumda AUB ancak ve ancak ya A, ya da B (ya da ikisi birden) gerçekleştiğinde gerçekleşir. Şekil 4.3, AUB'yi verir.

Daha genel olarak, $E_1, E_2,...,E_k$ gibi k tane olay verilsin. $E_1 \cup E_2 \cup E_3 ... \cup E_k$ birleşimi, bu k tane olaydan hiç olmazsa birinde yer alan bütün olası sonuçlar kümesidir.



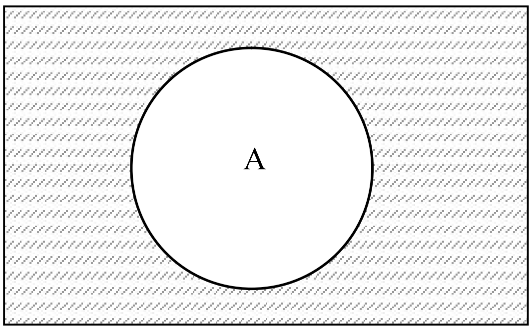
Şekil 4.3 A∪B'nin Gösterimi

Tam Sistem (Bütünü Kapsayıcı Sistem)

S örneklem uzayında $E_1, E_2,...,E_k$ gibi k tane olayda, $E_i \neq \varnothing$ (i=1,2,...,k) olmak üzere $E_1 \cup E_2 \cup E_3 ... \cup E_k = S$ ise ve $E_i \cap E_j = \varnothing$ ($i \neq j$) için sağlanıyorsa, $E_1, E_2,...,E_k$ bir tam sistem (bütünü kapsayıcı sistem) oluşturur.

Tümleyen Olay

S örneklem uzayındaki bir olay A olsun. S'de yer alan ancak A'da yer almayan olası sonuçlar kümesine A'nın tümleyeni denir. $\overline{A}$ ile gösterilir. İlgili gösterim Şekil 4.4'de verilmiştir.



Şekil 4.4. ▨ $\overline{A}$ 'nın Gösterimi

Örnek 4.3: Altı yüzlü bir zarın atılmasında

A:{Gelen yüzün çift olması}

B: {Gelen yüzün en az 4 olması}

olsun. Aşağıda verilen tanımları gösteriniz.

a) $\overline{A}$, $\overline{B}$ b) $A \cap B$ c) $A \cup B$

Çözüm: A: {2,4,6} ve B: {4,5,6}

a) $\overline{A}$: {1,3,5} ; $\overline{B}$: {1,2,3}

b) $A \cap B$: {4,6}

c) $A \cup B$: {2,4,5,6}

Örnek 4.4: İstanbul altın borsasındaki değerleri ard arda iki gün izlediğimizi düşünelim. İncelenen borsa için temel olası sonuçlar aşağıda verilmiştir:

	1.Gün	2.Gün
O_1	Yükseliyor	Yükseliyor
O_2	Yükseliyor	Aynı Kalıyor
O_3	Yükseliyor	Düşüyor
O_4	Aynı Kalıyor	Yükseliyor
O_5	Aynı Kalıyor	Aynı Kalıyor
O_6	Aynı Kalıyor	Düşüyor
O_7	Düşüyor	Yükseliyor
O_8	Düşüyor	Aynı Kalıyor
O_9	Düşüyor	Düşüyor

Sonuçlar incelendiğinde, yukarıda verilen 9 sonuçtan birinin gerçekleşmek zorunda olduğu görülür. Ancak aynı anda ikisi gerçekleşemez. Bunlar temel sonuçlar olup örneklem uzayı;

$$S: \{O_1, O_2, O_3, O_4, O_5, O_6, O_7, O_8, O_9\}$$

şeklindedir. Bu örneklem uzayında aşağıda verilen iki olay tanımlansın.

A:{ İstanbul altın borsasındaki değerin ilk gün yükselmesi}

B:{ İstanbul altın borsasındaki değerin ikinci gün yükselmesi}

Bu durumda,

A:{ O_1, O_2, O_3}

B: {O_1, O_4, O_7}

olacaktır. A ve B olaylarına ilişkin $A \cap B$:{O_1} olur. Birinci ya da ikinci günde yükselmesi ve B'deki bütün temel sonuçların kümesidir ve aşağıda verilmiştir:

$$A \cup B: \{O_1, O_2, O_3, O_4, O_7\}$$

$\overline{A}$: İstanbul altın borsasındaki değerin ilk gün yükselmeyen durumlarını içerir. Bu da A'da bulunmayan ancak S'de bulunan temel olası sonuçlar olup.

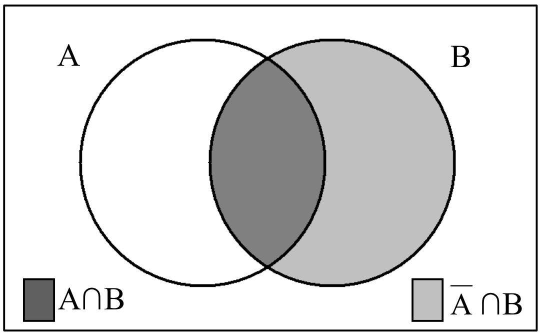
$$\overline{A}:\{O_4, O_5, O_6, O_7, O_8, O_9\}$$

olarak verilir.

A ve B Olayları İçin İlgili Bağıntılar

i) A ile B iki olay olsun.

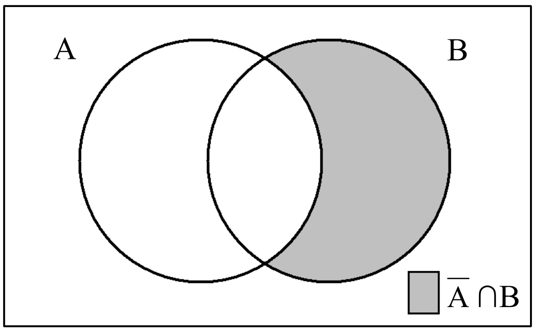
$A \cap B$ ile $\overline{A} \cap B$ olayları ayrık olaylardır ve Şekil 4.5'de gösterilmiştir.



Şekil 4.5 $A \cap B$ ve $\overline{A} \cap B$ 'nin Gösterimi

Burada, $(A \cap B) \cup (\overline{A} \cap B) = B$ olacaktır.

ii) A ile B iki olay olsun. A ile $\overline{A} \cap B$ olayları ayrık olaylar olup, birleşimi $A \cup B$ 'dir (Şekil 4.6).



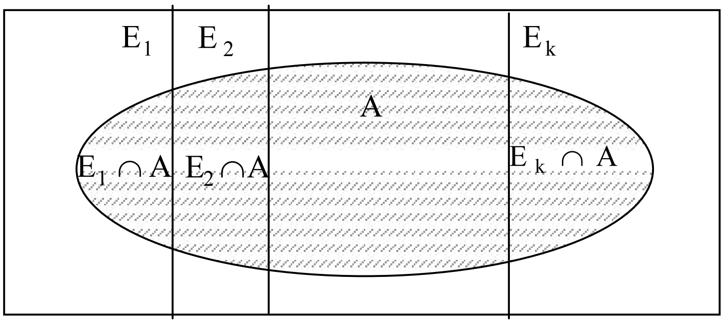
Şekil 4.6 $\overline{A} \cap B$ 'nin Gösterimi

Yani,

$$A \cup (\overline{A} \cap B) = A \cup B$$

dir.

iii) $E_1, E_2, ..., E_k$ tam sistem oluştursunlar. Bu durumda $E_1 \cap A, E_2 \cap A, ..., E_k \cap A$ olaylarının birleşimi A'dır. Bu sonucun doğruluğu Şekil 4.7'de görülmektedir.



Şekil 4.7 $(E_1 \cap A) \cup (E_2 \cap A) \cup ... \cup (E_k \cap A) = A$

Örnek 4.5: Zar atma deneyinde

A: $\{2,4,6\}$, $E_1 : \{1,2\}$ $E_2 : \{3,4\}$ $E_3 : \{5,6\}$

olarak tanımlandığında;

E_1, E_2, E_3 tam sistem oluştururlar.

$E_1 \cap A : \{2\}$; $E_2 \cap A : \{4\}$; $E_3 \cap A : \{6\}$

$(E_1 \cap A) \cup (E_2 \cap A) \cup (E_3 \cap A) = A$ 'dır.

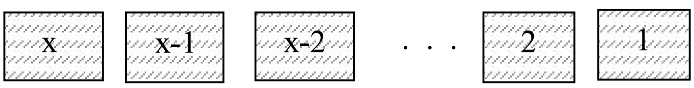
$\{2\} \cup \{4\} \cup \{6\} = \{2,4,6\}$=A'dır.

Permütasyon

Bir olayın olasılığı hesaplanırken, örneklem uzayında ilgilenilen olayın temel sonuçlarının sayılması sorun olur. Bunun için permütasyon ve kombinasyon olasılık hesapları için yararlı yöntemlerdir. Bu bölümde bu iki konu ele alınacaktır.

x tane farklı nesne ile her biri bir kez kullanılmak üzere kaç değişik sıralama yapılabileceğine cevap bulmak bu konuya ilişkindir.

Aşağıdaki şekilde verilen kutuları düşünelim. İlk kutuya x nesne, x farklı şekilde yerleştirilebilir. Geriye kalan (x-1) tane nesne, sonraki kutuya (x-1) farklı şekilde yerleşir. Bu uygulamaya aynı şekilde devam edilirse,

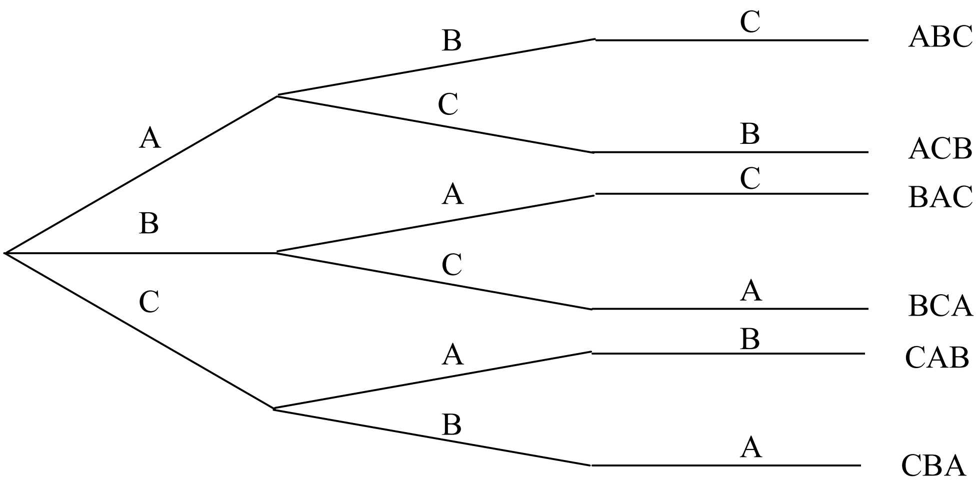


olur. İlk iki kutuya x(x-1) tane farklı şekilde nesne konulabilir. Üçüncü kutu da düşünüldüğünde, x.(x-1).(x-2) tane farklı yolda yerleştirme yapılabilir. En son kutuya gelindiğinde, yerleştirilecek bir nesne kalır. Tüm sıralamanın farklı sayısı ise,

$$x(x-1)(x-2)(x-3)...2.1=x! \qquad (4.1)$$

kadardır.

Örnek 4.6: A,B,C harflerinin kaç değişik biçimde sıralanabileceğini ağaç çizimi ile görelim.



Şekil 4.8 Ağaç Çizimi

Yukarıda verilen ağaç çiziminden görüldüğü gibi 3!=6 değişik biçimde sıralama gerçekleşmiştir.

Örnek 4.7: Bir pazarlamacı 5 müşterisini ziyaret edecektir. Bu pazarlamacı kaç farklı şekilde ziyaret sırasını düzenleyebilir?

5!=120 farklı şekilde ziyaret sırasını düzenleyebilir.

n tane nesnenin x tane kutuya konulacağı düşünülsün. (x < n). Her bir nesne yalnız ve yalnız bir kez kullanılabilsin. Olanaklar içinde sıralanabilme sayısına, n nesneden çekilmiş x nesnenin permütasyon sayısı denir. Permütasyon sayısı ${}_nP_x$ sembolü ile gösterilir.

n	n-1	n-2	. . .	n-x+2	n-x+1

$${}_nP_x = n.(n\text{-}1).(n\text{-}2)...(n\text{-}x+2).(n\text{-}x+1)$$

kalan (n-x) nesne, (n-x)(n-x-1)...(2).(1) sayıdadır. Bu sayının

$$n.(n\text{-}1).(n\text{-}2)...(n\text{-}x+2).(n\text{-}x+1)$$

ile çarpılması n! kadardır. n!, (n-x).(n-x-1)...(2).(1) çarpımına bölünerek, ${}_nP_x$ elde edilir. Böylece permütasyon sayısı,

$${}_nP_x = \frac{n!}{(n-x)!} \tag{4.2}$$

dir.

Örnek 4.8: A,B,C,D,E harfleri arasından iki harf seçilip bunların sıralanacağı düşünülsün. n=5, x=2 permütasyon sayısı

$${}_5P_2 = \frac{5!}{(5-2)!} = 20$$

olacaktır. Bunun anlamı; 20 farklı şekilde sıralama yapılabileceğidir.

Örnek 4.9: Örnek 4,7'de verilen örnekte, pazarlamacı bir günde iki müşteri ziyaret edebilecek ise 5 müşteriden bu iki kişi kaç değişik şekilde ziyaret edilebilir?

$${}_5P_2 = \frac{5!}{(5-2)!} = 20$$

20 farklı şekilde ziyaret söz konusu olabilir.

Örnek 4.10: 8 kişi 3 kişilik yere kaç farklı şekilde oturtulabilir?

$${}_8P_3 = \frac{8!}{(8-3)!} = 336$$

336 farklı şekilde oturtulabilir.

Permütasyona İlişkin Diğer Durumlar:

1. n tane nesnenin içinde n_1 tane birbirinin aynı, n_2 tane birbirinin aynı ... n_k tane birbirinin aynı nesne varsa $n_1 + n_2 + ... + n_k = n$ koşulu ile bunların farklı sıralama sayıları,

$${}_nP_{n_1,n_2,...,n_k} = \frac{n!}{n_1!n_2!...,n_k!} \tag{4.3}$$

olacaktır.

Örnek 4.11: MATEMATİK sözcüğünden kaç farklı sıralama yapılabilir?

Burada n=9, n_1=2 M, n_2=2 A, $n_3 = 2$ T, $n_4 = 1$ E, $n_5 = 1$ K, $n_6 = 1$ i'dir. 2+2+2+1+1+1=9 olmak üzere, sıralama sayısı aşağıda verilmiştir.

$$ {}_nP_{n_1,n_2,\ldots,n_k} = \frac{9!}{2!2!.2!1!1!1!} = 45360 $$

2. n farklı nesneden yerine iade edilerek, r tane nesne çekildiğinde, sıralanma sayısı n^r tanedir.

Örnek 4.12: 1,2,3 rakamları ile 2 basamaklı (yerine iade edilerek) $3^2 = 9$ tane rakam elde edilebilir. Bu sıralamalar;

11 12 13 21 22 23 31 32 33

biçimindedir.

* Yerine iade edilerek sıralamada n nesnenin tümü kullanılıyor ise farklı sıralama sayısı n^n kadardır.

* Milli piyango çekiliş biletleri, 0,1,2,3,4,5,6,7,8,9 rakamlarından 6 basamaklı sayılardan oluşur. Buradaki farklı numaralı bilet sayısı 10^6 kadar olacaktır.

Kombinasyon

Burada n tane nesneden x tanesinin kaç değişik şekilde seçileceği ile ilgilenilir. Kombinasyonda yerine koyma ve sıralama yapıldığı düşünülmez.

$$ {}_nC_x = \binom{n}{x} = \frac{n!}{x!(n-x)!} \tag{4.4} $$

n faklı nesneden x tanesinin kaç farklı şekilde seçilebileceğini verir.

Örnek 4.13: Bir lokantada boş 4 garson kadrosu bulunmaktadır. Lokantaya yedisi erkek, üçü bayan olmak üzere 10 aday başvurmuştur.

a) Adaylardan 4 tanesinin tamamen tesadüfi olarak seçilmesi kaydı ile, kadrolar kaç farklı şekilde doldurulur?
b) Eğer hiç bayan işe alınmayacak ise, kadrolar kaç farklı şekilde doldurulur?

Bu problemin çözümü sıralamanın önemli olmadığı farklı seçilme işlemi olan kombinasyon ile çözümlenir.

a) 10 kişiden hiçbir ayırım yapılmadan 4 kişi seçilecektir. O halde farklı seçim sayısı,

$$_{10}C_4 = \binom{10}{4} = \frac{10!}{4!(10-4)!} = 210$$

dur.

b) Eğer hiçbir bayan seçilmeyecek ise, seçimler yalnız erkekler içinden yapılır. Bu durumda

$$_{7}C_4 = \binom{7}{4} = \frac{7!}{4!(7-4)!} = 35$$

farklı seçim yapılabilir.

Örnek 4.14: 5 kız ve 7 erkek öğrenciden 6 kişilik bir grup seçilecektir. Grup 3 kız ve 3 erkek öğrenciden oluşacağına göre kaç farklı seçim yapılabilir?

Gruba girecek 3 kız için $\binom{5}{3}$ farklı seçim, gruba girecek 3 erkek için $\binom{7}{3}$ farklı seçim yapılır. Buna göre 6 kişilik değişik grup sayısı,

$$\binom{5}{3}\binom{7}{3}=350$$

olacaktır.

4.3. OLASILIK

Bir olayın göreli sıklığı ikinci bölümde de görüldüğü gibi bir A olayının gerçekleşme sayısının tüm denemeler sayısına (n) oranıdır (A'nın göreli sıklığı=A'nın gerçekleşme sayısı/n).

Tanımı verilen göreli sıklık bir olayın gerçekleşme yüzdesi olarak tanımlanan olasılık tanımı ile yakından ilgilidir. Çünkü bir olayın olasılığı, olası temel sonuçlar içinde ilgilenilen sonuç yüzdesidir. Örnek 4.4'de, A olayının olasılığı 9 temel sonuç içinde A olayının olma yüzdesi olan %3,3, yani 3/9=1/3'dür. Bir para atma deneyinde yazı ve tura olmak üzere iki temel sonuç vardır. Temel sonuç olan yazı 1, tura 1 tanedir. Burada yazı gelme olasılığı ½ olacaktır. Bu açıklamadan sonra klasik olasılık ölçümü tanımlanabilir.

Olasılık, bir olayın gerçekleşebilirliğinin sayısal bir ölçüsünü verir. Olasılık 0 ile 1 arasında bir ölçekle ölçülür. Olayın gerçekleşmesi olanaksız ise, 0; olayın gerçekleşmesi kesin ise 1 değerini alır. Belirsiz olaylarda olayın gerçekleşebilirliği ne kadar fazla ise olasılık değeri de 1'e o kadar yakın olacaktır. Gerçekleşebilirliği ne kadar düşük ise, olasılık değeri de sıfıra o kadar yakın olacaktır.

Tanım 4.1: Bir A olayının olasılığı P(A) ile gösterilir.

P(A)=(Olası temel sonuçlar içinde A olayının gerçekleşme sayısı)/(Olası temel sonuçlar sayısı) (4.5)

Örneğin bir zar atma olayında A olayı "Gelen sayı çifttir" ise 6 temel sonuçtan 3 tanesi A'dadır. O halde P(A)=3/6=1/2'dir.

Eğer olasılık, deney sonucundan bulunacak ise aşağıda verilecek olan deneysel olasılık tanımı kullanılır.

Tanım 4.2: n kez yinelenen denemelerde A olayının gerçekleşme sayısı n_A ise, n sonsuza giderken, n_A/n oranına deneysel olasılık denir.

Örnek 4.15: Büyük bir şirkette eğitim ihtiyacı belirlemek için 500 kişilik bir örneklemeden elde edilen tablolardan biri aşağıda verilmiştir.

Tablo 4.1 Kalite Çemberi Eğitimi İsteme Sonuçları

K.Ç. Eğitimi	Beyaz Yakalılar	Mavi Yakalılar	Toplam
İsteyenler	100	250	350
İstemeyenler	100	50	150
Toplam	200	300	500

Tablo 4.1'e göre, bu firmada kalite çember eğitimini isteme olasılığı deneysel olasılık tanımından,

P(Kalite çember isteme)=350/500=0,70

olarak bulunur.

Örnek 4.16: Örnek 4.13'de, hiçbir bayanın işe alınmama olasılığı nedir?

$P(\overline{K}) =$ Hiçbir bayanın bulunmadığı durum sayısı/Oluşabilecek tüm gruplama sayısı,

$$P(\overline{K}) = 35/210 = 1/6$$

olur.

Olasılık Önermeleri

Bir rasgele deneyin örneklem uzayı S, temel sonuçları O_i, olay ise A ile gösterilsin. "A olayının gerçekleşme olasılığı" için P(A) gösterimi kullanılmak üzere aşağıdaki önermeler söz konusudur.

1. A, örneklem uzayı S içinde yer alan bir olaysa, A'nın gerçekleşme olasılığı 0 ile 1 arasında bir değerdir.

   $0 \leq P(A) \leq 1$

2. A, S içinde yer alan bir olay, O_i de temel sonuçlar olsun. Bu durumda,

   $$P(A) = \sum_{O_i \in A} P(O_i)$$

   olacaktır. Burada toplama işlemi A içindeki bütün temel sonuçları kapsar. Bu önerme, göreli sıklık bağlamında kullanılabilir. Bir rasgele denemenin n kez tekrarlandığı düşünülsün. O_i temel sonucunun gerçekleşme sayısı n_i, A olayının gerçekleşme sayısı da n_A olsun. Bu durumda, temel sonuçlar bağdaşmaz olduğundan, A'daki bütün temel sonuçlar için n_i'lerin toplamı, n_A'yı verir, yani

   $$n_A = \sum_A n_i$$

   ve

   $$P(A) = \frac{n_A}{n} = \frac{n_1}{n} + \frac{n_2}{n} + \ldots + \frac{n_i}{n} = P(O_1) + P(O_2) + \ldots + P(O_i)$$

   olur.

3. P(S)=1'dir. Bu önerme, örneklem uzayındaki bütün temel sonuçların olasılıkları toplamının 1 olduğunu ifade eder.

Önermelerin Doğurduğu Sonuçlar

i) Örneklem uzayı S, her birinin gerçekleşmesi aynı olan $O_1, O_2, \ldots, O_n$ gibi n tane temel sonuçtan oluşuyorsa, bunların her birinin olasılığı 1/n'dir. Yani;

$$P(O_i) = \frac{1}{n} \qquad (i = 1,2,\ldots,n)$$

dir. Örneğin bir zar atıldığında temel sonucun her birinin gerçekleşme olasılığı 1/6'dır.

ii) A ile B bağdaşmaz olaylar (ayrık olay) ise A ile B olaylarının birleşiminin olasılığı, her bir olasılığın toplamıdır, yani

$$P(A \cup B) = P(A) + P(B) \tag{4.6}$$

olur. Daha genel olarak $E_1, E_2, ..., E_k$ bağdaşmaz olaylar ise,

$$P(E_1 \cup E_2 \cup ... \cup E_k) = P(E_1) + P(E_2) + ... + P(E_k)$$

yazılabilir.

Bu bilgi ikinci önermeden gelmektedir. A ile B'nin birleşiminin olasılığı,

$$P(A \cup B) = \sum_{O_i \in (A \cup B)} P(O_i) \tag{4.7}$$

dır. Burada toplama işlemi P(A∪B) içindeki bütün temel sonuçları kapsamaktadır. Ama A ile B bağdaşmaz olaylar olduklarından hiçbir temel sonuç her ikisi içinde bulunamaz. Ayrıca A ile A'nın tümleyeni olan $\overline{A}$ ayrık olay olduklarından,

$$P(\overline{A})=1-P(A)$$

olur.

iii) $E_1, E_2, ..., E_k$ tam sistem oluşturan olaylar ise, bunların birleşimlerinin olasılığı 1'dir. Bu durumda,

$$P(E_1 \cup E_2 \cup ... \cup E_k) = 1$$

yazılabilir. Bu olaylar tam sistem oluşturduklarından birleşimleri bütün örneklem uzayı S'dir.

Herhangi İki Olay İçin Toplama Kuralı

Herhangi iki olay A ve B olsun. A ya da B'nin gerçekleşmesi olasılığı;

$$P(A \cup B)=P(A)+P(B)-P(A \cap B) \tag{4.8}$$

dir.

Örnek 4.17: Bir süpermarket hediye çekilişi düzenliyor. 1000 müşteri bu çekilişe katılıyor. Her bir müşteriye sadece bir form doldurma şansı veriliyor. Çekilişte 10 tane büyük, 100 tane de küçük hediye vardır. Her bir formun büyük hediyelerden birini kazanma şansı birbirinin aynı, küçük hediyelerden birini kazanma şansı da birbirinin aynıdır. Hiçbir form birden çok ikramiye kazanamaz. Tek bir müşterinin büyük hediye kazanma olasılığı kaçtır? Bir hediye kazanma olasılığı kaçtır?

A:{Seçilen formun büyük hediye kazanması}

B:{Seçilen formun küçük hediye kazanması}

P(A)=10/1000=0,01

P(B)=100/1000=0,1

Seçilen formun herhangi bir hediye kazanması olayı, A ile B'nin birleşimidir. Ayrıca bunlar ayrık olaylar olduğundan istenen olasılık,

$$P(A \cup B) = P(A) + P(B) = 0{,}01 + 0{,}1 = 0{,}11$$

olacaktır.

4.4. İKİ DEĞİŞKENLİ OLASILIK

Bu bölümde, yapılan bir deneyin iki değişkene bağlı olan sonuçları dikkate alınacaktır. Örneğin, bir gazete yöneticisi politikasını belirlemek için ekonomi sayfasını okuyan müşterilerin yaşları ile ilgilensin. Burada A olayı ekonomi sayfasının okunmasını, B olayı da yaşlar ile ilgili bilgiyi tanımlamaktadır. B olayının B_1 (25 yaş ve küçük), B_2 (26 – 40), B_3 (41 yaş ve büyük) olarak ayrık alt olaylardan, A olayının da A_1 (hiç okumuyor), A_2 (ara sıra okuyor), A_3 (devamlı okuyor) şeklinde ayrık alt olaylardan oluştuğu düşünülsün. Bu rasgele deneyde,

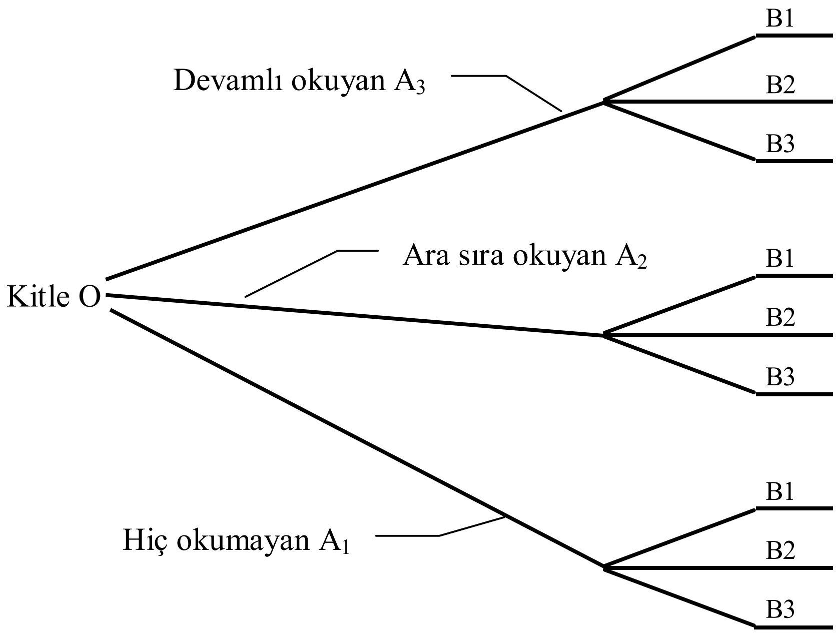
$$A_1 \cap A_2 = \varnothing,\ A_1 \cap A_3 = \varnothing,\ A_2 \cap A_3 = \varnothing,\ B_1 \cap B_2 = \varnothing,\ B_1 \cap B_3 = \varnothing, \ldots, B_5 \cap B_6 = \varnothing$$

olmakla birlikte A olayı ile B olaylarının ortak sonuçları olabilir ($A_i \cap B_j \neq \varnothing$). Burada $A_1 \cup A_2 \cup \ldots \cup A_k = A$, $B_1 \cup B_2 \cup \ldots \cup B_r = B$'dir. Böyle olaylara iki değişkenli olay, olasılıklarına da iki değişkenli olasılık denilmektedir. İki değişkenli olaylara ilişkin tablo, Tablo 4.2'de verilmiştir.

Tablo 4.2 İki Değişkenli Olay Gösterimi

	B_1	B_2	. . .	B_r
A_1	$A_1 \cap B_1$	$A_1 \cap B_2$	. . .	$A_1 \cap B_r$
A_2	$A_2 \cap B_1$	$A_2 \cap B_2$	. . .	$A_2 \cap B_r$
. . .				
A_k	$A_k \cap B_1$	$A_k \cap B_2$	. . .	$A_k \cap B_r$

Ekonomi sayfasını okuma ve yaşlar ile ilgili örnek, ağaç çizimi üzerinde aşağıdaki biçimde gösterilebilir.



Şekil 4.9 Ekonomi Sayfasını Okuyan Kişilerin Yaşlarına İlişkin Ağaç Grafiği

Şekil 4.9'da görüldüğü gibi, A ve B olaylarının sonuçları birleşerek iki değişkenli sonuçları oluştururlar. Toplam sonuç sayısı, A'nın sonuç sayısı ile B'nin sonuç sayıları çarpımı kadardır (3 x 3 = 9). Bu sonuçlara olasılıklar da bağlanabilir. Örneğe ilişkin alan araştırması sonuçları Tablo 4.3'de verilmiştir.

Tablo 4.3 Ekonomi Sayfasını Okuyan Kişilerin Yaşlarına İlişkin Çapraz Olasılık Tablosu

(A) Ekonomi Sayfasını Okuma	(B) Yaşlar			
	(B_1) (25 yaş ve küçük)	(B_2) (26– 40)	(B_3) (41 yaş ve büyük)	Toplam $P(A_i)$
(A_1) Hiç okumuyor	0,13	0,02	0,03	0,18
(A_2) Ara sıra okuyor	0,11	0,13	0,21	0,45
(A_3) Düzenli okuyor	0,03	0,21	0,13	0,37
Toplam $P(B_j)$	0,27	0,36	0,37	1

Burada deneysel olasılık kullanıldığı görülmektedir. Eğer araştırmanın geniş bir örneklem de yapıldığı düşünülürse, bu sonuçların olasılıklara yaklaştığı düşünülebilir. Böylece gazete okuyucu kitlesinin davranışları hakkında bilgi alınabilir. Örneğin; $P(A_1 \cap B_2) = 0{,}02$ olan sonucunun, 26–40 yaş grubunda, ekonomi sayfasını hiç okumayanların olasılığının 0,02 olduğu söylenebilir.

Kenar (Marjinal) Olasılıklar: İki değişkenli olasılıklarda (çapraz tablo oluşturuluyor) tekil olay olasılıklarına kenar olasılıklar denir. Tablo 4.2'de verilen A_i ve B_j olayları için, $P(A_i)$ ve $P(B_j)$ kenar olasılıkları gösterir.

Ortak Olasılıklar: İki değişkenli olasılıklarda arakesit olasılıkları $P(A_i \cap B_j)$ olarak yazılır. Bu olasılıklara ortak olasılıklar denir.

Burada kenar olasılıklar, A_i olayını oluşturan ve birbirlerinden ayrık olan $A_i \cap B_1; A_i \cap B_2; ...; A_i \cap B_r$ olasılıklarının toplamına eşittir.

$$P(A_i) = P(A_i \cap B_1) + P(A_i \cap B_2) + ... + P(A_i \cap B_r) \tag{4.9}$$

Bu sonuç $P(B_j)$ için de geçerlidir. Yani

$$P(B_j) = P(A_1 \cap B_j) + P(A_2 \cap B_j) + ... + P(A_k \cap B_j) \qquad (4.10)$$

olur.

Ekonomi sayfasını okuma örneğinde, rasgele seçilmiş bir kimsenin bu sayfayı ara sıra okuyor olma olasılığı aşağıda verilmiştir:

$$\begin{aligned} P(\text{Ara sıra okuyor}) &= P[\text{Ara sıra okuyor} \cap (\text{25 yaş ve küçük})] + \\ &\quad P[\text{Ara sıra okuyor} \cap (26\text{-}40] + \\ &\quad P[\text{Ara sıra okuyor} \cap (\text{41 yaş ve büyük})] \\ &= 0,11 + 0,13 + 0,21 = 0,45 \end{aligned}$$

Diğer olasılıklar da benzer biçimde aşağıdaki gibi elde edilir.

P(Düzenli okuyor)=0,37

P(Hiç okumuyor)=0,18

P(25 yaş ve küçük)=0,27

P(26-40)=0,36

P(41 yaş ve büyük)=0,37

A_1, A_2, ..., A_k ayrık ve bütünü tamamlayan olaylar olduğundan olasılıklar toplamı 1 olacaktır. Aynı durum hem B olayları hem de arakesit olayları olan $(A_i \cap B_j)$'lere ait ortak olasılıkları için geçerlidir. Yani ,

$$\sum_j^r \sum_i^k P(A_i \cap B_j) = 1$$

dir.

Koşullu Olasılık

Yukarıda verilen araştırmada (26-40) yaş grubunda ekonomi sayfasını okuyanlar bizi daha fazla ilgilendirebilir. Ya da, bir sigorta şirketinde sağlık sigortası yaptıran müşterilerden herhangi birinin bir yıl içinde sigortacısına sağlık masraflarının karşılanması için başvurma olasılığı yerine, (50-60) yaş grubunda olan müşterilerden birinin başvurma olasılığı bizi daha çok ilgilendirir. Bu iki örnekte görüldüğü gibi istenen olasılık bir ek koşul altında belirlenmektedir. Burada belli bir olayın gerçekleşme olasılığı, bir başka olayın gerçekleşmesi verilmişken ilgilenmemizi gerektirmektedir. Bu duruma koşullu olasılık denir.

Tanım 4.3: A ile B iki olay olsun. B olayının gerçekleştiği biliniyorken, A olayının gerçekleşme olasılığına koşullu olasılık denir ve P(A/B) ile gösterilir. P(B)>0 olma şartı ile bu olasılık,

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \tag{4.11}$$

olarak tanımlanır. Benzer şekilde koşul A olayı ise P(A)>0 şartı ile

$$P(B/A) = \frac{P(A \cap B)}{P(A)} \tag{4.12}$$

olur.

Ekonomi sayfası okuma ile ilgili araştırmada, rasgele seçilmiş bir kişinin (26-40) yaş grubunda olduğu bilindiğine göre, bu sayfayı ara sıra okuma olasılığı:

P[Ara sıra okuyor/(26-40)]=P[Ara sıra okuyor$\cap$(26-40)]/P(26-40)

$$= \frac{0,13}{0,36} = \frac{13}{36}$$

olacaktır. (25 ve daha küçük) yaş grubunda olduğu bilinen kişinin, bu sayfayı hiç okumama olasılığı aşağıda verilmiştir:

P[Hiç okumuyor/(25 yaş ve küçük)]=P[Hiç okumuyor$\cap$(25 yaş ve küçük)]/P(25 yaş ve küçük)

$$= \frac{0,13}{0,27} = \frac{13}{27}$$

Örnek 4.18: Büyük bir şehirde araba kazalarının nedenleri incelenmiş ve kazaların 0,04'ü kaygan zemine, 0,70'i şoförün dikkatsizliğine, 0,01'i de her iki nedene bağlı bulunmuştur. Sonuçlara göre aşağıdaki olasılıkları hesaplayınız?

A:{Zeminin kaygan olması}

B:{Şoförün dikkatsiz olması}

a) Şoförün dikkatsizliği görülmüşken, zeminin de kaygan olması olasılığı nedir?

$$P(A/B)=\frac{P(A\cap B)}{P(B)}=\frac{0,01}{0,70}=\frac{1}{70}$$

b) Zeminin kaygan olması görülmüşken, şoförün de dikkatsiz olması olasılığı nedir?

$$P(B/A)=\frac{P(A\cap B)}{P(A)}=\frac{0,01}{0,04}=\frac{1}{4}$$

c) Şoförün dikkatsiz olmadığı ve zeminin kaygan olması olasılığı nedir?

$$P(A\cap \overline{B})=P(A)-P(A\cap B)=0,04-0,01=0,03$$

d) Şoförün dikkatsiz olmadığı görülmüşken, zeminin kaygan olması olasılığı nedir?

$$P(A/\overline{B})=\frac{P(A\cap \overline{B})}{P(\overline{B})}=\frac{P(A\cap \overline{B})}{1-P(B)}=\frac{0,03}{1-0,70}=\frac{3}{30}=\frac{1}{10}$$

e) Zeminin kaygan olması veya şoförün dikkatsiz olması olasılığı nedir?

$$P(A\cup B)=P(A)+P(B)-P(A\cap B)$$
$$=0,04+0,70-0,01=0,73$$

Bağımsız Olay

İki ya da daha çok olayın ortaya çıkması birbirlerine bağlı değilse bu olaylara bağımsız olaylar denir. Örneğin iki zar birlikte atılırsa, zarlardan birinin 3, diğerinin 2 gelmesi bağımsız iki olaydır.

A ve B olaylarının bağımsız olması için gerek ve yeter koşul,

$$P(A\cap B)=P(A)P(B) \qquad (4.13)$$

dir.

A ve B bağımsız iki olay ise,

$$P(A/B)=P(A)$$

ve

$$P(B/A)=P(B)$$

olur.

Tanıt: Eşitlik (4.13)'ün tanıtı için, Eşitlik (4.11) ve Eşitlik (4.12) koşullu olasılık formüllerinin payları aynı olduğundan

$P(A \cap B) = P(A/B)P(B) = P(B/A)P(A)$

yazılabilir. A ve B olayları bağımsız ise;

$P(A/B) = P(A)$ ve $P(B/A) = P(B)$

olacağı için

$P(A \cap B) = P(A)P(B)$

olur.

Örnek 4.19: Bir para ve altı yüzlü zar birlikte atılıyor. A ve B olayları sırasıyla,

A:{Parada tura gelmesi}

B:{Zarda 3 gelmesi}

olarak tanımlandığına göre, bu olaylar bağımsız mıdır?

Para atma deneyinde elemanlar Y_1, Y_2; zar atma deneyinde elemanlar $T_1, T_2, T_3, T_4, T_5, T_6$ olmak üzere örneklem uzayı $S:\{T_1Y_1,\ T_2Y_1, \ldots\ T_6Y_1,\ T_1Y_2, \ldots,\ T_6Y_2\}$ olarak yazılabilir. Yukarıda verilen olaylara ilişkin olasılıklar,

$$P(A \cap B) = \frac{1}{12},\ P(A) = \frac{1}{2},\ P(B) = \frac{1}{6}$$

olduğuna göre,

$$P(A \cap B) = P(A)P(B)$$

$$\frac{1}{12} = \frac{1}{2} \times \frac{1}{6}$$

olup, Eşitlik (4.13) sağlandığı için A ve B olayları bağımsızdır.

Örnek 4.20: Ekonomi sayfasını okuma ile ilgili araştırmada, (26-40) yaş grubu ile (Ara sıra okuma) olayları bağımsız mıdır?

P(26-40)=0,36, P(Ara sıra okuyor)=0,45, P[(26-40)$\cap$(Ara sıra okuyor)=0,13

$0{,}36 \times 0{,}45 = 0{,}16$

$0{,}13 \neq 0{,}16$ olduğundan, bu iki olay bağımsız olaylar değildir.

BAYES Teoremi

Öncelikle toplam olasılık formülü anlatılacaktır. Toplam olasılık formülünde, S örneklem uzayının $E_1, E_2, \ldots, E_k$ gibi birbirinden ayrık ve toplamları S örneklem uzayını oluşturacak alt uzaylara ayrıldığı düşünülür. Örnek 4.21 üzerinde bu kavramlar açıklanacaktır.

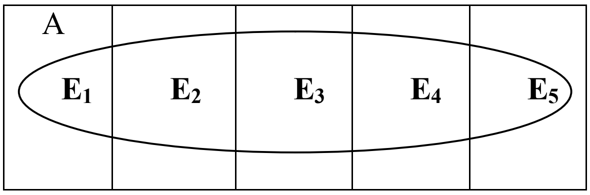
Örnek 4.21: Bir sigorta şirketi bir bölgede sağlık sigortaları için faaliyet gösterecektir. Ön bilgi için o bölgede yaşayan kitlenin yaş dağılımı ile ilgili bir araştırma yapılmış ve bilgiler Tablo 4.4'de verilmiştir.

Tablo 4.4 Yaşlara İlişkin Göreli Sıklıklar

Yaşlar	0-15	16-25	26-40	41-55	56 ve üzeri
Göreli sıklık (olasılık)	0,16	0,25	0,22	0,25	0,12

Tablo 4.4'de olaylar, E_1:{0-15}, E_2:{16-25}, E_3:{26-40}, E_4:{41-55}, E_5:{56 ve üzeri} olarak ayrılmıştır. Bu yaş grupları ayrıktır ve toplamları S örneklem uzayını vermektedir.

Bu bölgede bir anket uygulandığı düşünülsün. Burada sağlık sigortası yaptırmayı düşünenler için bir A olayı tanımlansın.



Şekil 4.10 A Olayının $E_1,...,E_5$ İle Gösterimi

Şekil 4.10'da A olayının olasılığı,

$$P(A)=P[(A\cap E_1)\cup (A\cap E_2)\cup (A\cap E_3)\cup (A\cap E_4) \cup(A\cap E_5)]$$

dır. $A\cap E_i$'ler ayrık olduğu için

$$P(A)=P(A\cap E_1)+P(A\cap E_2)+P(A\cap E_3)+P(A\cap E_4)+P(A\cap E_5)$$

yazılır. Buradan;

$$P(A)=\sum_{i=1}^{5} P(A\cap E_i) \tag{4.14}$$

olur.

Koşullu olasılık için verilen Eşitlik (4.11)'den

$$P(A\cap E_i)=P(E_i)P(A/E_i) \tag{4.15}$$

yazılabilir. Bu durumda Eşitlik (4.15), Eşitlik (4.14)'de yerine konularak,

$$P(A)=\sum_{i=1}^{5} P(E_i)P(A/E_i)$$

elde edilir. E_i'lerin n tane olduğunu düşünülürse, P(A) olasılığı

$$P(A)=\sum_{i=1}^{n} P(E_i)P(A/E_i) \tag{4.16}$$

olarak yazılabilir. Bu eşitlik toplam olasılık formülü olarak adlandırılır.

Örnek 4.21'de, birinci yaş grubunda olup, sağlık sigortasını yapmayı düşünenlerin olasılığı,

$P(A/E_1)=0{,}01$

ve ikinci, üçüncü, dördüncü, beşinci yaş grubunda olup sağlık sigortasını yaptırmayı düşünenlerin olasılıkları sırasıyla,

$P(A/E_2)=0{,}10$ $P(A/E_3)=0{,}09$ $P(A/E_4)=0{,}12$ $P(A/E_5)=0{,}20$

olmak üzere bu bölge için P(A) olasılığını bulunuz.

Toplam olasılık formülü kullanılarak sigorta yaptırmayı düşünenlerin olasılığı;

$P(A)=(0{,}16x0{,}01)+(0{,}25x0{,}10)+(0{,}22x0{,}09)+(0{,}25x0{,}12)+(0{,}12x0{,}20)=0{,}1004$

olarak bulunur.

Bayes Formülü:

Örnek 4.21'de araştırma yapılan bölgeden herhangi bir kişi belirlenmiş ve sağlık sigortasına baş vurmayı düşündüğü anlaşılmıştır. Bu kişinin hangi yaş grubundan olabileceği araştırılmıştır. Böyle bir problemin çözümü için Eşitlik (4.11)'de verilen koşullu olasılık formülü kullanılır. Koşul, sağlık sigortası yaptırmayı düşünenler olup, Eşitlik (4.11)'den

$$P(E_i/A)=\frac{P(E_i\cap A)}{P(A)} \tag{4.17}$$

dır. $P(E_i \cap A)=P(E_i)P(A/E_i)$ yazılabileceği daha önce görülmüştü. P(A) içinde Eşitlik (4.16) toplam olasılık formülü kullanılarak;

$P(A)=P(E_1)P(A/E_1)+ P(E_2)P(A/E_2)+ P(E_3)P(A/E_3)+ P(E_4)P(A/E_4)+ P(E_5)P(A/E_5)$

olur. $P(E_i/A)$ eşitliğinde bu bilgiler yerlerine konulur ise,

$$P(E_i/A)=\frac{P(E_i)P(A/E_i)}{\sum_{i=1}^{n} P(E_i)P(A/E_i)} \tag{4.18}$$

elde edilir. Eşitlik (4.18)'e Bayes formülü denir. Eşitlik (4.18)'de ki, $P(E_i)$ (i=1,2,...,n) olasılıklarına **önsel olasılıklar** denir. Önsel olasılıklar konu ile ilgili verilebilecek önceki olasılıklardır. $P(E_i/A)$ (i=1, ..., n) olasılıklarına da **sonsal olasılıklar** (ardıl) denir.

Örnek 4.22: Örnek 4.21'de verilen sigorta şirketi örneğinde; Bir kişi sigorta yaptırmayı düşündüğüne göre, bu kişinin 56 ve üzeri yaş grubunda olma olasılığı nedir?

P(56 ve üzeri/A)=P[(56 ve üzeri)$\cap$A]/P(A)

P[(56 ve üzeri)$\cap$A]=$P(E_i)\times P(A/E_i)$

$=0{,}12 \text{ x } 0{,}20 = 0{,}024$

Daha önce hesaplanan A olasılığına ait 0,1004 değeri yukarıdaki eşitlikte yerine yazılırsa,

$$P(E_5 / A) = \frac{0{,}024}{0{,}1004} = 0{,}239$$

elde edilir. Diğer olasılıklarda benzer biçimde bulunabilir.

Örnek 4.23: Bir üniversitenin öğrenci işlerinde, öğrenci dosyalarını inceleyen bir denetmen dosyaların %18'inin hata içerdiğini görmüştür. Tüm dosyaların %25'i dikey geçiş yapan öğrencilere aittir. Hatalı dosyaların %56'sı dikey geçiş yapan öğrencilere aittir. Rasgele seçilen bir dosya dikey geçiş yapan öğrenciye ait ise, hatalı çıkma olasılığı nedir?

P(hata) = 0,18 ; P(dikey geçiş yapanlar)=0,25 ; P(dikey geçiş yapan/hata)=0,56

Bayes teoreminden yaralanılarak istenen olasılık aşağıdaki biçimde bulunabilir.

P(hata/dikey geçiş yapan)=[P(dikey geçiş yapan /hata)×P(hata)]/P(dikey geçiş yapan)

$$= \frac{0{,}56 \times 0{,}18}{0{,}25} = 0{,}40$$

Örnek 4.24: Bir şirketin risk yöneticisi çok sayıda fonun geleceğini değerlendirmiştir. Ertesi yıl, bu fonların başarısı incelenmiş ve aşağıda verilen sonuçlar elde edilmiştir:

E_1:{Fon pazar ortalamasının çok üstünde başarı gösteriyor}

E_2:{Fon pazar ortalamasıyla aynı başarıyı gösteriyor}

E_3:{Fon pazar ortalamasının çok altında başarı gösteriyor}

A:{Risk yöneticisi tarafından fonun alınmaya değer bulunması}

$P(E_1)=0{,}35$; $P(E_2)=0{,}50$; $P(E_3)=0{,}15$

ve fonların risk yöneticisi tarafından alınmaya değer bulunma olasılıkları,

$P(A/E_1)=0,40$; $P(A/E_2)=0,30$; $P(A/E_3)=0,15$

olarak belirlenmiştir. Risk yöneticisi tarafından fonun alınmaya değer bulunması kararı verilmişken, pazar ortalamasının çok üstünde başarılı olmasının olasılığı nedir?

$$P(E_1/A)=\frac{P(E_1)P(A/E_1)}{\sum_{i=1}^{3}P(E_i)P(A/E_i)}$$

$$=\frac{0,35\times 0,40}{0,35\times 0,40+0,50\times 0,30+0,15\times 0,15}=0,448$$

4.5. RASLANTI DEĞİŞKENİ TANIMI VE ÖZELLİKLERİ

Olasılık, kitle ile örneklem arasındaki bağıntıyı kurmaya yardımcıdır. Bu amaçla önceki konularda, rastlantıya bağlı olaylardan ve bu olayların gerçekleşme olasılıklarından söz edilmişti. Konunun daha kolay anlaşılabilmesi için basit örnekler verilmişti. Örneğin, para atma olayında yazı ya da tura gelmesi, iskambil kağıtlarından kart çekilmesi, altı yüzlü zar deneyi gibi. Bu deneyler birer rasgele deneydir. Rasgele deneylerin tüm gerçekleşebilecek sonuçlarına örneklem uzayı (S), bu sonuçlar içinde tanımlanan özel sonuçların oluşturduğu alt kümeye de olay denildi.

Rasgele deney sonucunda tanımlanan bir özellik sayısal bir değerle ifade edilebiliyorsa, bu özellik raslantı değişkeni olarak isimlendirilecektir. Örneğin:

1- Bir basketbol takımının bir sezon boyunca her maçta kazandığı sayılar
2- İki paranın birlikte atılması deneyinde yazı gelme sayısı
3- Bir metro istasyonunda, belirlenen bir ay içinde belli bir zaman aralığında (sabah 9 ile 12 arasında) gelen müşterilerin günlük sayıları
4- Bir paranın yazı gelinceye kadar atılma sayısı
5- Bir denetçinin hatalı dosya buluncaya kadar inceleyeceği dosya sayısı
6- Bir gıda maddesinde koruyucu madde oranı
7- Paketlenmiş olarak satılan bir ürünün ağırlığı

Yukarıda verilen birinci ve üçüncü örneklerde raslantı değişkenlerinin sayısal değerleri sayı olarak belirlenmektedir. Birinci örnekte bir sezonda 34 maç olup, birinci maçta 82, ikinci maçta 70 ve...34. maçta 65 sayı yapılmış olabilir. İkinci örnekte gelen yüzde yazı sayısı 0,1,2 tane olabilir. Üçüncü örnekte belirlenen bir ay içinde farklı günlerde belli bir zaman aralığında gelen müşteri sayısı 102, 90, 95,...,110 olabilir. Bu üç örnekte raslantı değişkeni sayılabilir sonlu özelliktedir. Dördüncü ve beşinci örneklerde raslantı değişkeni yine sayılabilirdir ancak nerede durulacağı belli değildir. Örneğin dördüncü örnekte paranın yazı gelinceye kadar atılmasında örneklem uzayı;

S:{Y, TY, TTY, TTTY,...}

ve yazı gelinceye kadar atılma sayısı 1, 2, 3, 4 ... olacaktır. Beşinci örnekte de durum benzerdir. Denetçi, ilk dosyada da hata bulabilir, ikincide de ya da 30. dosyada da ve bu böylece devam edebilir. Bu örneklerde raslantı değişkeni sayılabilir sonsuzluktadır. Altıncı örnekte verilen raslantı değişkeni katkı maddesinin bulunma oranı olup, % 0,1, % 0,12 ... gibi değerler alabilir. Yedinci örnekte paketlerin ağırlıkları çok hassas ölçülürse belli iki değer arasında her ağırlığı alması mümkündür. Bu iki örnekte raslantı değişkeni süreklidir. Raslantı değişkenleri; sonlu ya da sonsuz sayılabilir özellikte ise kesikli raslantı değişkeni, iki değer arasındaki tüm reel değerleri alabiliyorsa, sürekli raslantı değişkeni olarak adlandırılır. Verilen örneklerde ilk 5 örnek kesikli, 6. ve 7. örnekler sürekli raslantı değişkenidir.

Raslantı değişkenleri X, Y, Z gibi büyük harflerle, bu değişkenlerin aldığı sayısal değerler de x, y, z gibi küçük harflerle gösterilir. Örneğin yukarıda verilen birinci örnekte sezon boyunca her maçta kazanılan sayılar X, alınan değerler ise x_1=82, x_2=70,...,x_{34}=65 olacaktır.

Kesikli Raslantı Değişkenin Dağılımı

X, sonlu sayıda $x_1, x_2, ..., x_n$ değerlerini

$$f(x_i)=P(X=x_i) \qquad i=1,2,...,n$$

olasılıkları ile alabilen kesikli raslantı değişkeni olsun. Bu raslantı değişkeninin alabileceği değerlerin, o değeri alma olasılıkları ile birlikte belirtilmesine X'in olasılık dağılımı ya da olasılık fonksiyonu (kesikli olasılık fonksiyonu) denir.

X'in olasılık dağılımı ya da olasılık fonksiyonu aşağıdaki koşulları sağlar:

1- $P(X=x_i) \geq 0$ tüm x_i'ler için

2- $\sum_{i=1}^{n} P(X = x_i) = 1$

Örneklem uzayı sonsuz ise ikinci koşul $\sum_{i=1}^{\infty} P(X = x_i) = 1$ olur.

Örnek 4.25: Bir paranın iki kez atılmasında örneklem uzayı $S:\{YT, TY, TT, YY\}$ idi. X tura gelme sayısı olsun. Burada x_1=0 (hiç tura gelmemesi), x_2=1 (bir kez tura gelmesi), x_3=2 (iki kez tura gelmesi) olabilir. Bu örnek için olasılık dağılımı,

f(0)=P(X=0)=1/4 (hiç tura olmayan bir tane sonuç vardır: YY)

f(1)=P(X=1)=2/4 (bir tura olan sonuç sayısı iki: YT, TY)

f(2)=P(X=2)=1/4 (ikisi de tura olan bir tane sonuç vardır: TT)

Yukarıdaki verilerin daha özet bir gösterimi aşağıda verilmiştir:

X=x	0	1	2
P(X=x)	¼	½	¼

Burada olasılıklar toplamı 1'e eşitir.

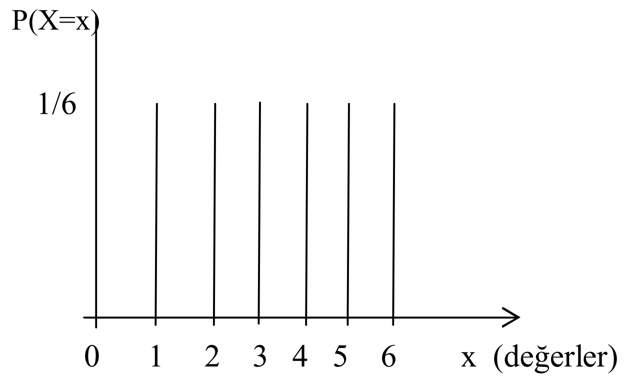
P(X=0)+P(X=1)+P(X=2)=1/4+1/2+1/4=1

Örnek 4.26: Zar atma deneyine ilişkin değerler ve değerlere ilişkin olasılıklar,

X=x	1	2	3	4	5	6
P(X=x)	1/6	1/6	1/6	1/6	1/6	1/6

şeklindedir. Olasılıklar toplamı $\sum_{i=1}^{6} P(X = x_i) = 1$'dir.

Olasılık dağılımının çizimi Şekil 4.11'de verilmiştir.



Şekil 4.11 Zar Deneyi İçin Olasılık Dağılımının Çizimi

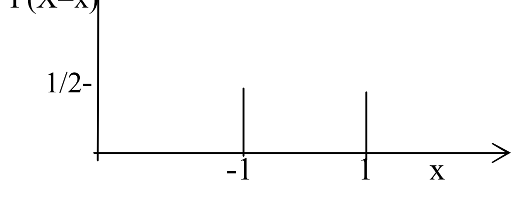
Örnek 4.27: Bir bilgisayar oyununda, 0, 1, 2, ..., 37'ye kadar numaralanmış kutular vardır. Oyuncu, tek sayılı kutuyu düşürebilmesi halinde bir puan kazanacak, çift sayılı kutuyu düşürürse bir puan kaybedecektir. Oyuncunun, tek sayılı kutuyu yakalayabilme olasılığı,

$$\text{P(Tek sayılı kutuyu düşürebilme)}=\frac{19}{38}=\frac{1}{2}$$

dir. Oyuncunun, tek sayı geldiğinde 1 puan kazandığı, gelmemesi durumunda da 1 puan kaybettiği yukarıda belirtildi. Oyuncunun kazancı X rastlantı değişkeni ile gösterilirse, X rastlantı değişkeninin olasılık fonksiyonu,

$$P(X=x)=\begin{cases} \dfrac{1}{2} & , \quad x=-1 \quad \text{için} \\ \dfrac{1}{2} & , \quad x=1 \quad \text{için} \\ 0 & , \quad \text{ö.d.} \quad \text{için} \end{cases}$$

olur. Bu fonksiyonun olasılık grafiği Şekil 4.12'de verilmiştir:



Şekil 4.12 Bilgisayar Oyunu İçin Olasılık Grafiği

Örnek 4.28: Bir bilgisayar oyununda 4 şifre vardır. Bu şifrelerden sadece birisi oyunu başlatacaktır. X, kişinin oyunu başlatmak için denediği şifre sayısı olduğuna göre, X'in olasılık fonksiyonunu bulunuz.

X'in alacağı değerler 1, 2, 3, 4 olacaktır. Çünkü, denenen şifre sayısı 1, 2, 3, 4 olabilir. Burada $A:\{$Oyunun başlaması$\}$, $\overline{A}:\{$Oyunun başlamaması$\}$ olarak tanımlanırsa, aşağıdaki olasılıklar elde edilir.

$$P(X=1)=P(A)=\frac{1}{4}$$

$$P(X=2)=P(\overline{A}\cap A)=P(\overline{A}).P(A)=\frac{3}{4}\times\frac{1}{3}=\frac{1}{4}$$

$$P(X=3)=P(\overline{A}\cap\overline{A}\cap A)=\frac{3}{4}\times\frac{2}{3}\times\frac{1}{2}=\frac{1}{4}$$

$$P(X=4)=P(\overline{A}\cap\overline{A}\cap\overline{A}\cap A)=\frac{3}{4}\times\frac{2}{3}\times\frac{1}{2}\times 1=\frac{1}{4}$$

Bu olasılıklara göre, X'in olasılık fonksiyonunu,

$$P(X=x)=\begin{cases}\frac{1}{4}, & x=1,2,3,4\\ 0, & \text{ö.d.}\end{cases}$$

dır.

Bu problem k şifre için genellenebilir. Yani,

$$P(X=j)=\begin{cases}\frac{1}{k}, & j=1,2,...,k\\ 0, & \text{ö.d.}\end{cases}$$

olur. İlk denemede veya 2.denemede,...,k. denemede başarıya ulaşma olasılığı aynıdır.

Kesikli Olasılık Dağılımından Yararlanarak Olasılık Hesaplama

Olasılık dağılımları raslantı değişkeni için bir fonksiyon vereceğinden, tüm örneklem uzayını tek tek yazmadan da yukarıda verilen fonksiyon yardımıyla istenilen olasılıklar elde edilebilir. Örneğin; P(X=a), P(X≤a), P(X≥a), P(a≤X≤b) olasılıkları toplam alınarak bulunabilir. Aşağıda ilgili olasılık formülleri verilmiştir:

$$P(X=a)=f(a) \tag{4.19}$$

$$P(X \geq a) = \sum_{x_k=a}^{\infty} P(X = x_k) \tag{4.20}$$

$$P(X \leq a) = \sum_{x_k=-\infty}^{a} P(X = x_k) \tag{4.21}$$

$$P(a \leq X \leq b) = \sum_{x_k=a}^{b} P(X = x_k) \tag{4.22}$$

Örnek 4.29: Bir zarın iki kez atılmasında üste gelen yüzlerdeki sayıların toplamı raslantı değişkeni ise, bu raslantı değişkeninin olasılık dağılımını bulunuz ve bu dağılımdan yararlanarak aşağıdaki olasılıkları hesaplayınız.

a- P(10≤X)=? b- P(4>X)=? c- P(9<X≤12)=?

Bu problemde olasılık dağılımını bulmak için S örneklem uzayı oluşturulmalıdır.

Tablo 4.5 İki Kez Zar Atmada Üst Yüze Gelen Sayıların Toplamını Veren Tablo

	Birinci atışta gelen değerler					
İkinci atış Değ.	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Tablo 4.5'deki toplam değerleri alma olasılıkları aşağıda verilmiştir:

X=x	2	3	4	5	6	7	8	9	10	11	12
P(X=x)	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

$P(X=x_i)$'nin aldığı değerleri verebilecek bir fonksiyon P(X=x)'den yararlanarak,

$$P(X=x_i)=(6-|x_i-7|)/36$$

dır. Bu fonksiyondan yararlanarak istenen olasılıklar aşağıdaki biçimde bulunabilir:

a- $$P(10 \le X) = \sum_{x_i=10}^{12} P(X=x_i) = \sum_{x_i=10}^{12}\left(\frac{6-|x_i-7|}{36}\right)$$

$$=\frac{3}{36}+\frac{2}{36}+\frac{1}{36}=\frac{1}{6}$$

b- $$P(4 > X) = \sum_{x_i=-\infty}^{3} P(X=x_i) = \sum_{x_i=-\infty}^{1} P(X=x_i) + \sum_{x_i=2}^{3} P(X=x_i)$$

$$=0+\left(\frac{6-|2-7|}{36}+\frac{6-|3-7|}{36}\right)=\frac{1}{36}+\frac{2}{36}=\frac{1}{12}$$

c- $$P(9 < X \le 12) = \sum_{x_i=10}^{12} P(X=x_i) = \sum_{x_i=10}^{12}\left(\frac{6-|x_i-7|}{36}\right)$$

$$=\frac{3}{36}+\frac{2}{36}+\frac{1}{36}=\frac{1}{6}$$

Kesikli Raslantı Değişkeninin Beklenen Değeri ve Varyansı

Kitle ortalaması μ'nün bir diğer gösterimi E(X)'dir. E(X), beklenen değer olarak adlandırılır.

$x_1,x_2,...,x_n$, X raslantı değişkeninin değerleri ve $P(X=x_1)$, $P(X=x_2)$, ...,$P(X=x_n)$ de bu raslantı değişkenlerini alma olasılıkları ise E(X);

$$E(X)=\mu=x_1P(X=x_1)+x_2P(X=x_2)+...+x_nP(X=x_n)$$

$$=\sum_x xP(X=x) \qquad (4.23)$$

eşitliği ile bulunur. μ, X'in aldığı değerlerin ağırlıklı ortalaması olup, ağırlıklar P(X=x) olasılıklarıdır. Beklenen değer, raslantı değişkeninin çok sayıda denemede alacağı değerlerin uzun dönem ortalaması olarak da açıklanabilir.

Örnek 4.30: Bir kitap sayfalarındaki yanlış sözcük sayıları belirlenmiştir. Sayfaların %79'unda hiç yanlış bulunmamıştır. %19'unda 1, %1'inde 2, %0,8'inde 3, %0,2'sinde 4 yanlış sözcük bulunmuştur. Burada X raslantı değişkeni yanlış sözcük sayıları ise, kitaptaki ortalama yanlış sözcük sayısı nedir?

Olasılıklar,

$$P(X=0)=0{,}79;\ \ P(X=1)=0{,}19;\ \ P(X=2)=0{,}01\ ;\ P(X=3)=0{,}008;\ \ P(X=4)=0{,}002$$

olmak üzere, kitaptaki ortalama yanlış sözcük sayısı,

$$E(X)=\mu=0\times0{,}79+1\times0{,}19+2\times0{,}01+3\times0{,}008+4\times0{,}002=0{,}242$$

dır.

Örnek 4.31: Bir beyaz eşya servis istasyonunda günlük verilen hizmetler için olasılık dağılımı aşağıdaki tabloda verilmiştir. Verilere göre günlük ortalama servis sayısı ne kadardır?

Günlük Verilen Servis Sayısı (X)	Olasılık P(X=x)	x×P(X=x)
0	0,075	0×0,0175=0
1	0,100	1×0,100=0,1
2	0,250	2×0,250=0,5
3	0,200	3×0,200=0,6
4	0,175	4×0,175=0,7
5	0,150	5×0,150=0,75
6	0,050	6×0,050=0,3

$E(X)=\mu=0+0,1+0,5+0,6+0,7+0,75+0,3=2,95$

Tanım 4.4: X raslantı değişkeni ister kesikli ister sürekli olsun, bu değişkenlerin kitlede birbirinden farklı olabileceği açıktır. Raslantı değişkenlerinin ortalamadan ayrılış ölçüsü varyans ile belirlenir. X raslantı değişkeninin varyansı V(X) ile ya da σ^2 ile gösterilir. Bu tanımdan,

$$\sigma^2 = E\left[(X-\mu)^2\right] = \sum_x (x-\mu)^2 P(X=x) \qquad (4.24)$$

olarak hesaplanır. Bu Eşitlik (4.24) daha basit olarak aşağıdaki biçimde de yazılabilir:

$$\sigma^2 = E(X^2) - \mu^2 = \sum_x x^2 P(X=x) - \mu^2 \qquad (4.25)$$

Varyans'ın kare kökü standart sapmadır ve X raslantı değişkeninin standart sapma, σ ile gösterilir.

Örnek 4.32: Örnek 4.30'daki yazım hataları ile ilgili problem için varyans ve standart sapmayı bulunuz.

Sayfalardaki yanlış sayısı	Olasılık	$x^2P(X=x)$
0	0,79	$0^2 \times 0,79 = 0$
1	0,19	$1^2 \times 0,19 = 0,19$
2	0,01	$2^2 \times 0,01 = 0,04$
3	0,008	$3^2 \times 0,008 = 0,072$
4	0,002	$4^2 \times 0,002 = 0,032$

Verilere ilişkin varyans,

$$\sigma^2 = 0,334 - 0,242^2 \cong 0,28$$

ve standart sapma,

$$\sigma = \sqrt{0,28} \cong 0,53$$

olarak elde edilir.

Örnek 4.33: Servis istasyonunun günlük hizmeti için varyans ve standart sapmayı bulunuz.

Günlük verilen servis sayısı(X)	Olasılık P(X=x)	x^2 P(X=x)
0	0,075	$0^2 \times 0{,}0175=0$
1	0,100	$1^2 \times 0{,}100=0{,}1$
2	0,250	$2^2 \times 0{,}250=1$
3	0,200	$3^2 \times 0{,}200=1{,}8$
4	0,175	$4^2 \times 0{,}175=2{,}8$
5	0,150	$5^2 \times 0{,}150=3{,}75$
6	0,050	$6^2 \times 0{,}050=1{,}8$

E(X)= μ=2,95

$$\sigma^2 = \sum_x x^2 P(X = x) - \mu^2$$

$$=11{,}75-2{,}95^2 =2{,}5475$$

olarak, standart sapma, $\sigma \cong 1{,}596$ olarak elde edilir.

Sürekli Raslantı Değişkenleri ve Olasılık Dağılımından Yararlanarak Olasılık Hesaplama

Tanım 4.5: X raslantı değişkeni, a, b aralığında her gerçel değeri alıyor ise, X sürekli raslantı değişkenidir. Örneğin boy uzunluğu, ağırlık, alınan notlar vb.

X sürekli raslantı değişkeninin belli değerleri alma olasılıklarını hesaplamak için kullanılan fonksiyona **olasılık yoğunluk fonksiyonu** denir. Kesikli raslantı değişkenlerinde olasılık fonksiyonunun gördüğü tüm işlevleri, sürekli raslantı değişkenlerinde olasılık yoğunluk fonksiyonu üslenir.

X sürekli bir raslantı değişkeni olduğunda, X'e ilişkin olasılık yoğunluk fonksiyonu, f(x) ile gösterilir. f(x), sürekli bir olasılık yoğunluk fonksiyonu olması için aşağıdaki koşulları sağlamalıdır:

1- $\int_a^b f(x)dx \geq 0$; $a \leq x \leq b$ için

2- $\int_{-\infty}^{\infty} f(x)dx = 1$

Kesikli raslantı değişkenlerinde olduğu gibi, sürekli raslantı değişkenleri içinde P(X=a), P(X≤a), P(X≥a), P(a≤X≤b) olasılıkları hesaplanabilir. Aşağıda ilgili olasılıklar verilmiştir:

$$P(X=a)=0 \tag{4.26}$$

$$P(X \geq a) = \int_a^{\infty} f(x)dx \tag{4.27}$$

$$P(X \leq a) = \int_{-\infty}^{a} f(x)dx \tag{4.28}$$

$$P(a \leq X \leq b) = \int_a^b f(x)dx \tag{4.29}$$

Örnek 4.34: f(x) fonksiyonu aşağıdaki gibi tanımlanmış olsun.

$$f(x) = \begin{cases} \frac{1}{4} & , \ 0 \leq x \leq 4 \\ 0 & , \quad \text{d.d.} \end{cases}$$

a- Bu fonksiyon bir olasılık yoğunluk fonksiyonu mudur?

b- P(2<X<3) olasılığını bulunuz.

c- P(X<2,8) olasılığını bulunuz.

d- P(X>3) olasılığını bulunuz.

a- f(x)≥0 ve $\int_0^4 \frac{1}{4} dx = 1$ olduğundan, olasılık yoğunluk fonksiyonudur.

b- $P(2 < X < 3) = \int_2^3 \frac{1}{4} dx = \frac{1}{4} x \Big|_2^3 = 0,25$

c- $P(X < 2,8) = \int_0^{2,8} \frac{1}{4} dx = \frac{1}{4} x \Big|_0^{2,8} = 0,7$

d- $P(X > 3) = \int_3^4 \frac{1}{4} dx = \frac{1}{4} x \Big|_3^4 = 0,25$

Tanım 4.6: Sürekli bir raslantı değişkeninin beklenen değeri (ortalaması),

$$E(X) = \mu = \int_{-\infty}^{\infty} x f(x) dx \tag{4.30}$$

ve varyansı

$$\sigma^2 = E\left[(X-\mu)^2\right] = \int_{-\infty}^{\infty} (X-\mu)^2 f(x) dx \tag{4.31}$$

olarak hesaplanır. Bu formül daha basit olarak aşağıdaki biçimde de verilebilir:

$$\sigma^2 = E(X^2) - \mu^2 = \int_{-\infty}^{\infty} x^2 f(x) dx - \mu^2 \tag{4.32}$$

Örnek 4.35: Örnek 4.34'de verilen f(x) fonksiyonun beklenen değerini ve varyansını bulunuz.

$$E(X) = \mu = \int_{-\infty}^{\infty} x f(x) dx$$

$$= \int_0^4 x \frac{1}{4} dx = \int_0^4 \frac{x}{4} dx = \left.\frac{x^2}{8}\right|_0^4 = 2$$

$$\sigma^2 = \int_0^4 x^2 \frac{1}{4} dx - (2)^2 = \left.\frac{x^3}{12}\right|_0^4 - 4 = \frac{64}{12} - 4 = \frac{16}{12} = \frac{4}{3}$$

- **Beklenen değer ile ilgili özellikler:**

1- Sabit sayıların beklenen değeri kendisine eşittir. k herhangi bir sabit sayı olmak üzere;

$$E(k) = k \tag{4.33}$$

dır.

2- k herhangi bir sabit sayı ve X raslantı değişkeni olmak üzere k ile X'in çarpımının beklenen değeri, k tane X'in beklenen değerinin çarpımına eşittir. Yani

$$E(kX)=kE(X) \tag{4.34}$$

dır.

3- k sabit bir sayı, X bir raslantı değişkeni ve U(X), X raslantı değişkeninin bir fonksiyonu olsun. k sabit sayı ile U(X) fonksiyonunun çarpımının beklenen değeri, U(X) fonksiyonun beklenen değeri ile k sabit sayısının çarpımına eşittir. Yani,

$$E[kU(X)]=kE[U(X)] \tag{4.35}$$

olup, burada $E[U(X)]=\begin{cases} \sum\limits_{x} u(x)P(X=x) & \text{(kesikli)} \\ \int\limits_{-\infty}^{+\infty} u(x)f(x)dx & \text{(sürekli)} \end{cases}$

dir.

- **Varyans ile ilgili özellikler:**

1- Sabit bir sayının varyansı sıfıra eşittir. k sabit bir sayı olmak üzere,

$$V(k)=0 \tag{4.36}$$

dir.

2- k sabit bir sayı ve X raslantı değişkeni olmak üzere, k ile X'in çarpımının varyansı, X raslantı değişkeninin varyansı ile k sabit sayısının karesinin çarpımına eşittir.

$$V(kX)=k^2\,V(X) \tag{4.37}$$

3- X raslantı değişkeni ile k sabit sayısının toplamının varyansı, X'in varyansına eşittir. Yani,

$$V(X+k)=V(X) \tag{4.38}$$

olur.

4- a, b sabit birer sayı ve X raslantı değişkeni olmak üzere aşağıdaki ifade doğrudur:

$$V(aX+b)=a^2V(X) \tag{4.39}$$

PROBLEMLER

1- Eğer yineleme yapılmaz ise 2, 5, 6, 7 ve 9 rakamlarından;

a- 3 basamaklı kaç farklı sayı yapılır?

b- Bunlardan kaç tanesi 400 den küçüktür?

c- Bunlardan kaç tanesi çifttir?

d- Bunlardan kaç tanesi tektir?

e- Bunlardan kaç tanesi 5'in katlarıdır?

2- 7 kişi;

a- Bir sıradaki 7 iskemlede kaç farklı düzende oturulabilir?

b- Yuvarlak bir masa etrafında kaç farklı düzende oturulabilir?

3- Bir grupta bulunan 3 erkek, 2 kız arkadaş,

a- Bir sırada kaç değişik biçimde oturabilirler?

b- Kızlar bir tarafta erkekler diğer tarafta olacak biçimde kaç değişik biçimde oturabilirler?

c- Sadece kızlar yan yana olma koşulu ile kaç değişik biçimde oturabilirler?

4- A ve B arasında 6 yol, B ve C arasında 4 yol vardır.

a- Bir kimse A'dan C'ye B yoluyla kaç farklı biçimde gidebilir?

b- Bir kimse A'dan C'ye B yoluyla kaç farklı gidiş geliş yapar?

c- Bir kimse aynı yoldan birden fazla geçmeme koşulu ile A'dan C'ye kaç farklı gidiş geliş yapabilir?

5- Bir öğrenci bir sınavda 30 sorudan 20'sini yanıtlamak zorundadır.

a- Öğrencinin kaç seçeneği vardır?

b- İlk iki soruyu yanıtlaması halinde kaç seçeneği vardır?

c- İlk 10 sorudan 3'ünü yanıtlama koşulu ile kaç seçeneği vardır?

d- İlk 10 sorudan en az 3'ünü yanıtlama koşulu ile kaç seçeneği vardır?

e- Ya birinci ya da ikinci soruyu yanıtlama koşulu ile kaç seçeneği vardır?

6- 5 A partisinden, 6 B partisinden ve 4 C partisinden oluşan bir gruptan, 7 kişilik bir komite ikisi A'dan, ikisi B'den, üçü C'den olacak şekilde oluşturulacaktır. Kaç değişik komite oluşturulabilir?

7- Bir torbada bulunan 9'u beyaz, 6'sı kırmızı 15 tane top arasından ard arda iadesiz olarak 3 top çekiliyor.

a- Toplardan üçünün de kırmızı olma olasılığı nedir?

b- Üç topunda aynı renkte olması olasılığı nedir?

8- Bir binada bağımsız olarak çalışan iki asansör bulunmaktadır. Bunlardan birincisi zamanın %18'inde, ikincisi de zamanın %25'inde birinci katta beklemektedir.

a- Her ikisinin de birinci katta bekleme olasılığı nedir?

b- Her ikisinin de birinci katta beklememe olasılığı nedir?

c- En fazla bir asansörün birinci katta bekleme olasılığı nedir?

9- A ve B isimli iki satranç oyuncusu 12 el satranç oynamışlar, 6 elde A, 4 elde B kazanmış ve 2 elde berabere bitmiştir. Yeniden 3 el oynamaya karar vermişlerdir.

a- 3 oyunda da A'nın kazanması,

b- 2 oyunun beraberlikle sonuçlanması,

c- A ve B'nin sırayla kazanması

olasılıklarını bulunuz.

10- Bir sigorta elemanının hazırladığı 20 adet sele karşı ev sigortasından 7'si tam riskli bölgededir. Bu poliçelerden bir yıl içinde 4 tanesine ödeme yapılacağı düşünüldüğüne göre,

a- Bu dört ödemeden tümünün tam riskli bölgeden olma olasılığı nedir?

b- Bu dört ödemeden birinin risksiz bölgeden üçünün riskli bölgeden olma olasılığı nedir?

11- Bir sokakta bulunan 260 aileden 120 tanesi oturdukları evlerin sahibi, 140 tanesi ise kiracıdır. Ev sahiplerinin 105 tanesi ve kiracıların 120 tanesi serbest meslek sahibidir. Diğerleri ücretle çalışmaktadır. Bu sokaktan rasgele seçilen bir ailenin,

a- Ev sahibi ve serbest meslek sahibi olması,

b- Ev sahibi veya serbest meslek sahibi olması,

c- Ücretli olması olasılıklarını bulunuz?

12- A, B, C gibi üç atlet yarışıyor. A'nın kazanma olasılığı B'nin kazanma olasılığına eşit, C'nin iki katıdır. A, B ve C'nin kazanma olasılıklarını bulunuz.

13- Ticari Bilimler Fakültesinde okuyan öğrencilerin 0,30'u hukuk dersinden, 0,20'si ekonomiden ve 0,15'i hem hukuk dersinden hem de ekonomi dersinden başarısız olmuştur. Rasgele seçilen bir öğrenci hukuk dersinden başarısız ise, ekonomi dersinden de başarısız olma olasılığı nedir?

14- Bir fabrikada üretilen malların % 40'ı birinci makineden, % 25'i ikinci makineden ve %35'i üçüncü makineden üretilmektedir. Bu makinelerden elde edilen malların arızalı çıkma olasılıkları sırası ile %2, %4 ve % 5'dir.

a- Üretilen mallardan rasgele olarak alınan bir tanesinin arızalı olma olasılığı nedir?

b- Rasgele olarak seçilen bir malın arızalı olarak çıktığı bilindiğine göre, bu malın üçüncü makineden seçilme olasılığı nedir?

15- Aşağıda verilen X kesikli olasılık dağılımına göre:

a- X'in beklenen değerini bulunuz.

b- X'in standart sapmasını bulunuz.

X	P(x)
200	0,40
250	0,35
260	0,25

16- Bir TV'nin sürekli olarak yapmış olduğu hava durumları üzerindeki tahmini kayıtları aşağıdaki tabloda verilmiştir.

a- Havanın güneşli olduğu tahmin edilmişse, o gün yağmur yağma olasılığı nedir?

b- O gün havanın güneşli olduğu bilindiğine göre, tahminin doğru olması olasılığı nedir?

	Tahmini Durum		
Gerçek durum	Güneşli	Bulutlu	Yağmurlu
Güneşli	0,30	0,05	0,05
Bulutlu	0,04	0,20	0,02
Yağmurlu	0,10	0,04	0,20

17- Bursa Uludağ kayak alanında bulunan bir otelde profesyonel 3 kayak öğretmeni kurs vermektedir. Bunlardan biri Bursa, biri Erzurum ve üçüncüsü de Erzincanlıdır. Otelin politikasına göre, öğretmenler öğrencilere rasgele atanıyor. Dolayısıyla, bir müşteri öğretmen aradığında, rasgele bir seçim yapılıyor ve müşteriyle seçilen öğretmen ders saatleri belirleniyor.

a- Belirli bir günde, dört müşteri ders almak için aradığında, bu müşterilerden üçünün Bursalı birinin de Erzurumlu öğretmene atanması olasılığı nedir?

b- Bir başka günde, üç müşteri ders almak için aradığında üç'ünün de Bursalı öğretmenden ders alması olasılığı nedir?

18- f(x) fonksiyonu aşağıdaki gibi tanımlanmış olsun.

$$f(x)=\begin{cases}\frac{2}{3}x & , \quad 0\le x\le 1\\ \frac{1}{3} & , \quad 1<x\le 3\\ 0 & , \quad dd\end{cases}$$

a- Bu fonksiyon bir olasılık yoğunluk fonksiyonu mudur?

b- P(1<X<2) olasılığını bulunuz.

c- P(X<2,4) olasılığını bulunuz.

d-P(X>2) olasılığını bulunuz.

e- X raslantı değişkeninin ortalamasını ve varyansını bulunuz.

19- Bir kafeteryaya gelen öğrencilerin içecek siparişi verme olasılığı 0,85, tost siparişi verme olasılığı 0,55 ve hamburger siparişi verme olasılığı 0,40 olarak belirlenmiştir.

a- İçecek ve hamburger siparişi verme olayları bağımsız olaylar ise, kafeteryaya gelen bir öğrencinin verdiği siparişin hamburger içermeme olasılığı nedir?

b- Tost siparişi veren bir öğrencinin aynı zamanda içecek siparişi verme olasılığı 0,85 ise, siparişin hem içecek hem de tost içermesi olasılığı nedir?

20- Başkent Üniversitesi acil bölümü, 365 gün süreyle her gün alınan acil çağrıları kaydetmiş ve aşağıdaki sıklık dağılımını elde etmiştir.

a- Verilen sıklık dağılımını temel alarak olasılık dağılımını bulunuz.

b- Olasılık dağılımının ortalamasını ve standart sapmasını bulunuz

c- Her acil çağrıya üç kişilik bir ekibin yanıt verdiği düşünülsün. Acil çağrılarının en azından %75'ine yanıt verebilmek için, Başkent Üniversitesinde kaç kişi çalıştırılmalıdır?

ÇAĞRI	GÜN SAYISI
0	40
1	36
2	72
3	99
4	51
5	36
6	10
7	21
	365

5. BEŞİNCİ BÖLÜM

5.1. GİRİŞ

5.2. BAZI KESİKLİ OLASILIK DAĞILIMLARI

5.3. SÜREKLİ RASLANTI DEĞİŞKENLERİNİN OLASILIK DAĞILIMLARI

PROBLEMLER

5.1. GİRİŞ

Önceki bölümde kesikli ve sürekli raslantı değişkenlerine ait özellikler verilmişdi. Bu bölümde de bazı önemli kesikli ve sürekli olasılık dağılımları verilecektir.

5.2. BAZI KESİKLİ OLASILIK DAĞILIMLARI

Bu bölümde kesikli raslantı değişkenlerinde sık raslanan bernoilli, binom, poisson olasılık dağılımlarına değinilecektir. Örnek problemler ile bu dağılımlardan nasıl yararlanılacağı gösterilecektir.

Bernoulli Dağılımı

Olasılık konularında verilen örneklerde, tek bir deneme için ortaya çıkacak sonuçlar yalnızca iki durum içeriyorsa, Bernoulli dağılımı söz konusudur. Örneğin, paranın tek atışında yazı ya da tura gelmesi, tek bir oyunda kazanma ya da kaybetme vb.

Tanım 5.1: X raslantı değişkeni başarı için 1, başarısızlık için 0 değerini alsın. X'in olasılık fonksiyonu;

$$P(X=1)=p$$

$$P(X=0)=1-p=q$$

olur. Ya da

$$P(X=x)=p^{x}(1-p)^{1-x} \quad ; \quad x=0,1 \qquad (5.1)$$

ise bu dağılıma bernoulli dağılımı denir.

Bernoulli dağılımının beklenen değeri,

$$E(X)=\mu=p \qquad (5.2)$$

ve varyansı,

$$\sigma^2 = E(X^2)-[E(X)]^2 = 1^2p-p^2=p(1-p)= pq \qquad (5.3)$$

dir.

Binom Dağılımı (İki Terimli Dağılım)

Bernoulli denemelerinin n kez tekrarlandığı düşünülsün. Bu denemelerde başarılı sonuçların toplam sayısı X raslantı değişkeni olarak gösterilsin. X raslantı değişkeni aşağıdaki koşulları sağlıyorsa, bu raslantı değişkeni binom raslantı değişkenidir:

1- Deneyde iki sonuç vardır. Başarılı olma olasılığı p, başarısız olma olasılığı (1-p)=q olarak tanımlanır.

2- Deney boyunca yapılan n deneme, aynı koşullar altında gerçekleştirilir.

3- Bir tek deneme için başarılı olma olasılığı p ve başarısızlık olasılığı q her deneme için aynıdır.

4- Denemeler birbirinden bağımsızdır.

5- Deney boyunca n sabit kalır.

X raslantı değişkeni bir binom raslantı değişkeni ise dağılım, binom dağılımı olarak adlandırılır.

Binom raslantı değişkeni için aşağıdaki örnekler verilebilir:

1- 3 çocuklu ailelerde kız çocuk sayısı

2- Bir paranın 4 kez atılmasında yazıların sayısı

3- Kusurlu oranı 0,03 olan bir üretimde, 10'arlık ürün içeren paketlerde kusurlu parça sayısı.

Dağılımın Elde Edilmesi

Örnek 5.1: Bir üretimde kusurlu oranı 0,10'dur. 4'er ürün bulunan paketlerde kusurlu ürünlerin dağılımını oluşturunuz.

Bu olayda karşılaşılacak olan sonuçlar, X raslantı değişkeninin değerleri ve olasılıkları aşağıda verilmiştir:

Tablo 5.1 Dört Ürün İçin Kutularda Kusurlu Ürün Dağılımı

Sonuçlar	X Raslantı Değerleri	X Raslantı Değerlerini Alma Sayısı	Olasılık
KKKK	4	$\binom{4}{4}=1$	$1[0{,}10^4\times0{,}90^0]=0{,}0001$
KKKS KKSK KSKK SKKK	3	$\binom{4}{3}=4$	$4[0{,}10^3\times0{,}90^1]=0{,}0036$
KKSS KSKS KSSK SKKS SKSK SSKK	2	$\binom{4}{2}=6$	$6[0{,}10^2\times0{,}90^2]=0{,}0486$
SSSK SSKS SKSS KSSS	1	$\binom{4}{1}=4$	$4[0{,}10^1\times0{,}90^3]=0{,}2916$
SSSS	0	$\binom{4}{0}=1$	$1[0{,}10^0\times0{,}90^4]=0{,}6561$

K: Kusurlu ürünü; S: sağlam ürünü göstermektedir.

Tablo 5.1 incelendiğinde, X'in olasılıklar toplamının 1 ve bir olasılık fonksiyonu olduğu görülür.

Paketlerde bulunan ürün sayısına n, kusurlu ürün sayısı X'in aldığı değerlere de x denilirse, X'in olasılık fonksiyonu;

$$P(X=x)=\binom{n}{x}p^x(1-p)^{n-x} \quad , \qquad x=1,2,3,...,n \tag{5.4}$$

olarak yazılabilir. Ya da Eşitlik (5.4),

$$P(X=x)=\frac{n!}{x!(n-x)!}p^x q^{n-x} \quad , \quad x=1,2,3,...,n \tag{5.5}$$

olarak da verilebilir. Eşitlik (5.4) ve Eşitlik (5.5) binom dağılımının olasılık fonksiyonudur. Burada, örneklem büyüklüğü n, ilgilenilen olayın olasılığı p, p+q=1'dir. Binom olasılık fonksiyonu kullanılarak bazı olasılıklar bulunabilir. Örneğin, kutularda 2 kusurlu ürün bulunma olasılığı,

$$P(X=2)=\frac{4!}{2!(4-2)!}0{,}10^2 0{,}90^{4-2}=0{,}0486$$

dır.

Binom Tablosunu Kullanma

Binom dağılımında olasılıkları kolay hesaplayabilmek için tablolar oluşturulmuştur. Bu tablolardan bir bölüm Ek 1'de verilmiştir. Binom tablosunda her n değeri için ayrı ayrı olasılıklar bulunmaktadır. Tabloda X binom raslantı değişkeninin alacağı x değerleri ile belirlenen olayın gerçekleşme olasılığı olan p'ye ait değerler vardır. Tablonun içindeki değerler, x ve p'ye karşılık gelen olasılıklardır.

Örnek olarak, Ek 1 ile verilen binom olasılıkları tablosundan, n=4 ve bazı p'ler için değerler aşağıda verilmiştir.

		n=4	
X	p=0,01	p=0,10	p=0,50
0	0,9606	0,6561	0,0625
1	0,0388	0,2916	0,2500
2	0,0006	0,0486	0,3750
3	0,0000	0,0036	0,2500
4	0,0000	0,0001	0,0625

Örnek 5.1 için bazı olasılık değerlerini tablo kullanarak bulalım. Örnek 5.1'de n=4 ve p=0,10 olduğu için binom olasılık tablosundan n=4, p=0,10 olan sütundan yararlanırız. Bu sütun Tablo 5.1'de verilen olasılıkları, hesaplama gerekmeksizin vermektedir. Örneğin n=4 ve p=0,10 için $P(X=2)=0{,}0486$; $P(X=4)=0{,}0001$'dir.

Binom Dağılımının Ortalaması, Varyansı ve Standart Sapması

Binom dağılımının ortalaması, varyansı ve standart sapması sırasıyla aşağıda verilmiştir:

$$\mu = E(X) = np \tag{5.6}$$

$$\sigma^2 = npq \tag{5.7}$$

$$\sigma = \sqrt{npq} \tag{5.8}$$

dır. Burada, n=örneklem büyüklüğü; p= ilgilenilen olayın olasılığı; q=1-p' dir.

Örnek 5.2: 5 çocuklu ailelerde erkek çocuk sayısına ilişkin dağılımı oluşturunuz ve aşağıdaki soruları cevaplayınız. (X, erkek çocuk sayısı, ailede erkek çocuğu olma olasılığı p=1/2'dir.)

a- 3 ve daha az erkek çocuk olma olasılığı nedir?

b- 2 den daha çok erkek çocuk olma olasılığı nedir?

Erkek Çocuk Sayısı (X)	$P(X=x)=\frac{n!}{x!(n-x)!}p^x q^{n-x}$
0	$\binom{5}{0}\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^5=0{,}0313$
1	$\binom{5}{1}\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^4=0{,}1563$
2	$\binom{5}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^3=0{,}3125$
3	$\binom{5}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^2=0{,}3125$
4	$\binom{5}{4}\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^1=0{,}1563$
5	$\binom{5}{5}\left(\frac{1}{2}\right)^5\left(\frac{1}{2}\right)^0=0{,}0313$

a-$P(X\leq 3)=P(X=0)+P(X=1)+P(X=2)+P(X=3)=0{,}0313+0{,}1563+0{,}3125+0{,}3125=0{,}8126$

b-$P(X>2)=P(X=3)+P(X=4)+P(X=5)=0{,}3125+0{,}1563+0{,}0313=0{,}5001$

Örnek 5.3: Bir futbolcunun penaltı atışını gole çevirmesi olasılığı 0,40'dır. Her atışın bir diğerinden bağımsız olduğu varsayımı altında, yapılan bir maçta bu futbolcu 5 atış yaparsa,

a- Tamı tamına 2 gol yapması olasılığı nedir? (Formül kullanarak bulunuz)

$$P(X=2)=\binom{5}{2}(0{,}40)^2(0{,}60)^3=0{,}3456$$

b- En çok 2 gol yapması olasılığı nedir? (Tablo kullanarak bulunuz.)

Ek 1'de verilen tabloda n=5'de; X=0, X=1 ve X=2 için verilen olasılıklar toplanır.

$$P(X\leq 2)=P(X=0)+P(X=1)+P(X=2)=0{,}0778+0{,}2592+0{,}3456=0{,}6826$$

c- En az 3 tanesinde gol yapması olasılığı nedir?(Formül kullanarak bulunuz).

$$P(X \geq 3) = P(X = 3) + P(X = 4) + P(X = 5)$$

$$= \binom{5}{3}(0,40)^3 (0,60)^2 + \binom{5}{4}(0,40)^4 (0,60)^1 + \binom{5}{5}(0,40)^5 (0,60)^0 = 0,2304 + 0,0768 + 0,0102 = 0,3174$$

Poisson Dağılımı

Belli bir zaman aralığında, belli bir alanda ya da hacimde nadir rastlanan olayların olasılık dağılımları Poisson dağılımı ile modellenebilir. Aşağıda Poisson dağılımı ile modellenebilecek örnekler verilmiştir:

1- Bir kavşakta bir ay içinde meydana gelen ölümcül trafik kazalarının sayısı
2- Bir iş kolunda belli bir sözleşme döneminde gerçekleşen grev sayısı
3- Bir dakikada bir kasaya gelen müşterilerin sayısı
4- Bir bölgede yapılan taramada, kanser hastalığı yaşanmış bireylerin sayısı

Poisson dağılımının ortalaması ve varyansı aynı olup, tek bir parametresi vardır ve bu parametre λ ile gösterilir.

X Poisson raslantı değişkeninin olasılık fonksiyonu,

$$P(X = x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!} \tag{5.9}$$

olarak tanımlanır. Burada,

e=2,71828

x=t birim zaman içinde ilgilenilen olay sayısı,

λt=t birim zaman içinde ilgilenilen olayın ortalama oluş sayısı,

dır. Genellikle t= 1 alınır. Bu durumda Poisson dağılımının olasılık fonksiyonu;

$$P(X = x) = \frac{(\lambda)^x e^{-\lambda}}{x!} \tag{5.10}$$

olur. Dağılıma ilişkin ortalama,

$$E(X)=\mu=\lambda \tag{5.11}$$

varyans,

$$\sigma^2=\lambda \tag{5.12}$$

ve standart sapma,

$$\sigma = \sqrt{\lambda} \tag{5.13}$$

dır.

Örnek 5.4: Bir bankaya bir saatte gelen müşterilerin ortalama sayısı 20 olsun. Burada X raslantı değişkeni bir Poisson dağılımı göstersin. 15 dakikada gelecek müşteri sayısı ortalama ne olur? 15 dakikada 1 müşteri gelme olasılığı nedir?

Burada 1 saatlik zaman diliminin [t=60×(1/4)=15] yani, ¼'ü kullanılmıştır.

t=1 saat iken λ=20,

t=1/4 saat iken λt=20×1/4=5 olur.

15 dakikada 1 müşteri gelme olasılığı,

$$P(X=1)=\frac{(5)^1 e^{-5}}{1!}=0,0337$$

olur.

Örnek 5.5: Yılda 1000 defter tutan bir muhasebe şirketinde hatalı hesap içeren defterlerin ortalama hata sayısının λ=0,4 olduğunu ve Poisson dağılımına uyduğunu bildirmektedir. X raslantı değişkeni hata sayısı olup, X raslantı değişkeninin olasılık fonksiyonu;

$$P(X=x)=\frac{(0,4)^x e^{-0,4}}{x!} \qquad , \qquad x=0,1,2,3...$$

dir. Bir yıl içinde tutulan bir defterde hiç hata olmama, 1 hata olma, 2 hata olma, 3 hata olma olasılıklarını ve 1000 defterde ortalama kaç tanesinde 0,1,2,3 hata bulunacağını hesaplayınız.

$$P(\text{Hiç hata olmama})=\frac{(0,4)^0 e^{-0,4}}{0!}=0,6703 \Rightarrow 0,6703\times1000\cong670 \text{ adet}$$

$$P(1 \text{ hata olma})=\frac{(0,4)^1 e^{-0,4}}{1!}=0,2681 \Rightarrow 0,2681\times1000\cong268 \text{ adet}$$

$$P(2 \text{ hata olma})=\frac{(0,4)^2 e^{-0,4}}{2!}=0,0536 \Rightarrow 0,0536\times1000\cong54 \text{ adet}$$

$$P(3 \text{ hata olma})=\frac{(0,4)^3 e^{-0,4}}{3!}=0,0072 \Rightarrow 0,0072\times1000\cong7 \text{ adet}$$

Binom dağılımında olduğu gibi Poisson dağılımı için de, farklı λ ve X'ler için tablolar hazırlanmıştır. Aşağıda buna ilişkin örnek verilmiştir.

Örnek 5.6: Bir telefon santralinde 9.00-9.10 saatleri arasında gelen telefonların ortalama sayısı $\lambda=3$ olan Poisson dağılımı göstermektedir. Santralde oturan bir operatörün 9.00-9.10 saatleri arasında,

a- Hiç konuşma yapmaması,

b- 2'den az konuşma yapması,

c- 2 ve daha fazla konuşma yapması olasılıklarını bulunuz.

Ek 2 ile verilen Poisson olasılıkları tablosundan, $\lambda=3$ için verilen değerler alınır ve istenen X'e göre olasılıklar elde edilir.

X	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
P(X=x)	0,0498	0,1494	0,2240	0,2240	0,1680	0,1008	0,0504	0,0216	0,0081

a- $P(X=0)=\frac{3^0 e^{-3}}{0!}=0{,}0498$

b- $P(X<2)=P(X=0)+P(X=1)=0{,}0498+0{,}1494=0{,}1992$

c- $P(X\geq 2)=1-P(X<2)$

$=1-0{,}1992$

$=0{,}8008$

5.3. SÜREKLİ RASLANTI DEĞİŞKENLERİNİN OLASILIK DAĞILIMLARI (NORMAL DAĞILIM)

Sürekli raslantı değişkenlerine ait bazı özel dağılımlar vardır. Bu dağılımlar genelde üstel fonksiyon biçimindedir. Burada bu aileden olan ve çok sık kullanılan normal dağılım ele alınacaktır.

Olasılık yoğunluk fonksiyonlarından biri olan ve istatistikte çok yaygın kullanılan normal dağılım, istatistik bilimi için büyük önem taşır. X sürekli raslantı değişkeninin olasılık yoğunluk fonksiyonu normal dağılım gösteriyorsa, aşağıda belirtilen özelliklere sahiptir:

1- Tek tepeli bir dağılımdır.

2- Ortalamasına ($X=\mu$) göre simetriktir.

3- Ortalama, ortanca, tepe değeri birbirine eşittir.

4- X raslantı değişkeni $-\infty$ ile $+\infty$ arasında değerler alır.

5- Daha önce de belirtildiği gibi gözlemlerin yaklaşık olarak

%68'i, $\mu \pm 1\sigma$

%95'i, $\mu \pm 2\sigma$

%99,7'si, $\mu \pm 3\sigma$

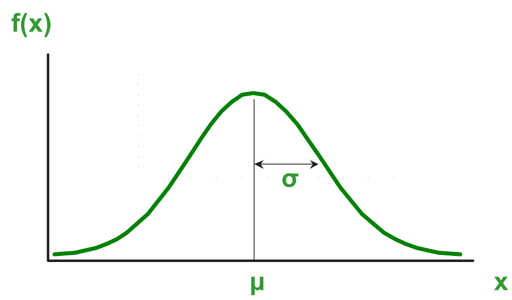
%100'ü $\mu \pm 4\sigma$

arasındadır.

Normal dağılımın olasılık yoğunluk fonksiyonu Eşitlik (5.14)'de verilmiştir:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad , \qquad -\infty < x < \infty \tag{5.14}$$

Burada; raslantı değişkeninin herhangi bir değeri, σ=Kitlenin standart sapması, μ=Kitlenin ortalaması ve e≅2,76183'dir. Şekil 5.1'de normal dağılımın grafiği verilmiştir.



Şekil 5.1 Normal Dağılım Grafiği

Eşitlik (5.14) ile verilen ifade bir olasılık yoğunluk fonksiyonu olduğu için integralinin değeri 1'e eşittir. Yani,

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 1$$

dir. Dağılımı tanımlayabilmek için μ ve σ'yı bilmek gerekir. İntegral ile bulunan belli sınırlar arasında eğrinin alanı, belli bölgelerde bulunma olasılıklarını verir. $P(a \leq X \leq b)$ olasılığı,

$$\int_{a}^{b} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

integralinin alınması ile bulunur. İstatistikte sık kullanılan bu dağılım için, bu integralin alınması pratik ve kolay bir iş değildir. Bunun için dağılımın standartlaştırılması işlemi uygulanır. Bu uygulama, olasılık yoğunluk fonksiyonunun belli sınırlar arasındaki alan hesaplamasında kolaylık sağlar.

Tanım 5.2: Bir dağılımın standartlaştırılması, alanı değişmemek şartı ile ortalamanın sıfıra kaydırılması, varyansın 1'e eşitlenmesidir. Tüm normal dağılımlar standart normal dağılıma dönüştürülebilir.

Standart Normal Dağılım (SND)

Standart normal dağılımda ortalama 0, varyans 1'dir. Ortalamaya göre tam simetrik bir dağılımdır. SND'ın olasılık yoğunluk fonksiyonu $Z = \frac{X-\mu}{\sigma}$ dönüşümü ile,

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \quad , \qquad -\infty < z < \infty \tag{5.15}$$

dur. Bu dağılım, X raslantı değişkenlerinin belli bölgelerde bulunma olasılığını bulmak için kullanılır. Örneğin bir standart normal dağılımda, Z=0,00 ile Z=1,45 arasında bulunma olasılığı,

$$P(0 \le Z \le 1{,}45) = \int_0^{1{,}45} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$$

integrali alınarak elde edilir. Ancak uygulamada, bu alanlar için daha önce oluşturulmuş tablolar kullanılır. Ek 3'de verilmiş tablo, bir SND tablosudur ve Z'nin sıfırdan, 3,09'a kadar olan değerleri için çeşitli alanları vermektedir. Simetrik bir dağılım olduğu için, sol tarafa ait bilgiler verilmemiştir. Çünkü dağılım simetrik olup sağ taraf ile aynı değerlere sahiptir. Örneğin,

$$P(0<Z<0{,}52)=P(-0{,}52<Z<0)$$

dir. Bazı kaynaklar Ek 3'de verilen tablo yerine,

$$F(Z) = \int_{-\infty}^{z} f(z)dz$$

ile tanımlanan standart normal dağılımın birikimli olasılıklarını veren tabloları kullanırlar (Ek 4).

Ortalaması, μ ve Varyansı, σ^2, Bilinen Normal Dağılımlarda Olasılık Hesabı

Normal dağılım ND($\mu;\sigma^2$) ya da N($\mu;\sigma^2$) ile gösterilir. Gösterimin anlamı, ortalaması μ ve varyansı σ^2 olan dağılım demektir. Her ND($\mu;\sigma^2$) dağılım, standart dağılım olan SND(0;1)'e dönüştürülebilir. Bu dönüştürme işlemi kitle için;

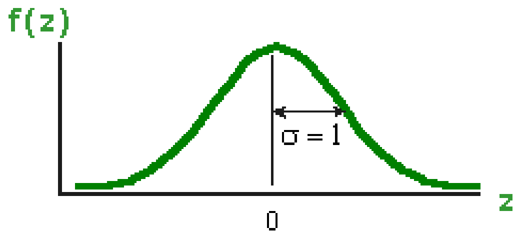
$$Z = \frac{X-\mu}{\sigma}$$

örneklem için;

$$Z = \frac{X-\overline{X}}{S}$$

ile gerçekleştirilir.

Şekil 5.2'de, standart normal dağılım (SND) grafiği verilmiştir.



Şekil 5.2 Standart Normal Dağılım

Ek 3'de verilen SND tablosunu kullanmak için; dağılımında, Z değerlerinin tam sayı kısmı ve virgülden sonra gelen ilk hane satırdan okunur. Virgülden sonra gelen ikinci hane ise, sütundan okunur, ikisinin kesiştikleri yerde olan rakam, sıfır ile o, Z değeri arasında bulunma olasılığıdır.

Örnek 5.7: Günlük tüketim ortalaması μ=3000 lt ve standart sapması σ=1000 lt olan bir petrol istasyonunun, gerekli işlemlerde iş hacmini belirleyebilmek için aşağıdaki olasılıklara ihtiyacı vardır. Aşağıdaki olasılıkları hesaplayınız.

a- Bu şirkette günlük tüketimin 2500 lt ve daha fazla olma olasılığı nedir?

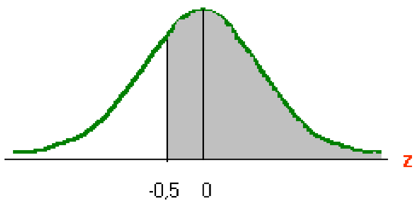
1. Adım: İstenen olasılık bölgesi yazılır.

$$P(X>2500)=P(X\geq 2500)$$

2. Adım: ND(3000;1000^2) da verilen X değerinin SND(0;1)da karşıtı bulunur.

$$Z = \frac{2500-3000}{1000} = -0{,}5$$

3. Adım: Bu olasılığın SND'da karşıtı yazılır ve tablodan bulunabilecek şekle getirilir. Yani P(Z>-0,5) olarak yazılır. Grafik çizilerek istenen bölge belirlenir.



4. Adım: 3. adımda taranarak gösterilen olasılıklar tablodan bakılıp yazılır.

P(-0,5<Z<0)=0,1915'dir. P(0<Z<∞)=0,5'dir.

Taralı alanın olasılığı 0,5+0,1915=0,6915 olacaktır.

Bu problem, Ek 4'de verilen, $F(Z) = \int_{-\infty}^{z} f(z)dz$ ile tanımlanan birikimli standart normal dağılım tablosundan yararlanılarak da hesaplanabilir. Tablonun simetrikliğinden dolayı

$$\int_{-\infty}^{0,5} f(z)dz = \int_{-0,5}^{\infty} f(z)dz$$

dir. Bu nedenle tablodan 0,5 için bulunan 0,6915 değeri,

$$\int_{-0,5}^{\infty} f(z)dz$$

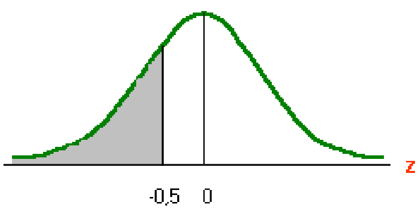
için istenen olasılıktır.

P(-0,5<Z<∞)=0,6915'dir.

b- Tüketimin 2500 lt'den az olma olasılığı nedir?

$$P(X<2500)=P\left(\frac{X-\mu}{\sigma} < \frac{2500-3000}{1000}\right) = P(Z < -0,5)$$

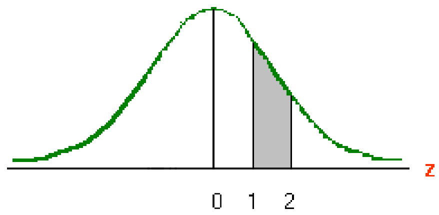
Şekil çizerek istenen bölge bulunabilir.



İstenen bölge; P(Z<-0,5)=0,5-0,1915=0,3085'dir.

c- Günlük tüketimin 4000 lt ile 5000 lt arasında olma olasılığı nedir?

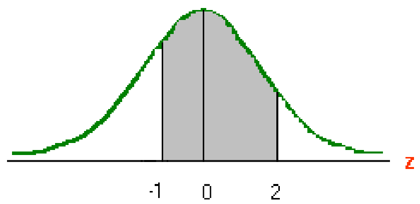
$$P(4000<X<5000)=P\left(\frac{4000-3000}{1000} < \frac{X-\mu}{\sigma} < \frac{5000-3000}{1000}\right) = P(1 < Z < 2)$$



$P(1<Z<2)=P(0<Z<2)-P(0<Z<1)=0,4772-0,3413=0,1359$

d- Günlük tüketimin 2000 lt ile 5000 lt arasında olma olasılığı nedir?

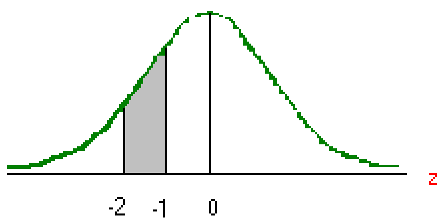
$$P(2000<X<5000)=P\left(\frac{2000-3000}{1000}<\frac{X-\mu}{\sigma}<\frac{5000-3000}{1000}\right)=P(-1<Z<2)$$



$P(-1<Z<2)=P(-1<Z<0)+P(0<Z<2)=0,3413+0,4772=0,8185$

e- Günlük tüketimin 1000 lt ile 2000 lt arasında olma olasılığı nedir?

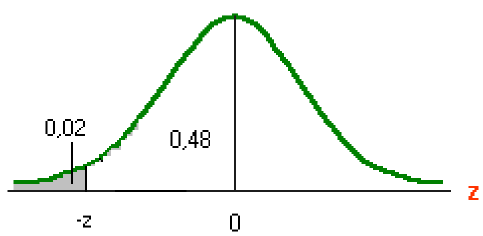
$$P(1000<X<2000)=P\left(\frac{1000-3000}{1000}<\frac{X-\mu}{\sigma}<\frac{2000-3000}{1000}\right)=P(-2<Z<-1)$$



$P(-2<Z<-1)=P(-2<Z<0)-P(-1<Z<0)=0,4772-0,3413=0,1359$ olacaktır.

f- Bu petrol şirketi bir sonraki yıl günlük tüketimin 5000 lt'den az olması olasılığının 0,02 olması halinde başarılı olacağını düşünmektedir. Bu durumda ortalama günlük tüketim kaç lt olmalıdır?

1 Adım: ND(μ; 1000^2) da istenen bölge işaretlenir.



2. Adım: P(X<5000)=0,02, SND tablosundan, bu olasılığa karşıt gelen Z değeri bulunur.

P(X<5000)=P(Z<?)=0,5-P(0<Z<?)=0,5-0,02=0,48

Tabloda P(0<Z<?)=0,48'dir. Tablonun içine bakılarak bu olasılığı veren Z değeri bulunur. Bu değer 2,05'dir.

3. Adım: $-Z=\frac{X-\mu}{\sigma}$ dönüşümünde, $-2{,}05=\frac{5000-\mu}{1000} \Rightarrow \mu = 7050$ lt olacaktır.

Örnek 5.8: Başkent radyosunda dinlenen bir müzik programının haftalık dinlenme süresi μ=4 ve σ=1 parametrelerine sahip olan bir normal dağılım gösterdiği bilinmektedir.

a- Haftalık müzik dinleme süresinin üç saatten az olma olasılığı nedir?

b- Bu sürenin 3,5 saat ile 4,5 saat arasında gerçekleşme olasılığı nedir?

a- $P(X<3)=P\left(Z<\frac{3-4}{1}\right)=P(Z<-1)$

=0,5-0,3413=0,1587

b- $P(3{,}5<X<4{,}5)=P\left(\frac{3{,}5-4}{1}<Z<\frac{4{,}5-4}{1}\right)=P(-0{,}5<Z<0{,}5)$

=0,1915+0,1915=0,3830

PROBLEMLER

1- İstatistik sınavında başarılı olma olasılığı % 50'dir. Sınava giren 3 kişiden,

a- 2'sinin başarılı olma olasılığını formül kullanarak bulunuz.

b- En çok 1 kişinin başarılı olma olasılığını formül kullanarak bulunuz.

c- En az 1 kişinin başarılı olma olasılığını tablo kullanarak bulunuz.

2- Belli bir üniversitenin I. sınıfına giren öğrencilerinin zamanında mezun olma olasılığı 0,40'dır. I. sınıfa yeni giren öğrencilerden 4'ü rasgele seçilmiştir. Bunlardan,

a- 3'ünün zamanında mezun olma olasılığını, olasılık fonksiyonu yardımıyla bulunuz.

b- En az 2'sinin zamanında mezun olma olasılığını tablo kullanarak bulunuz.

3- Bir bölgede 0-4 yaş arası çocuklarda kabakulak görülme olasılığı 0,25'dir. Bu bölgeden rasgele 4 çocuk seçildiğinde,

a- En az 3 çocuğun kabakulak olma olasılığını tablo kullanarak bulunuz.

b- En çok 1 çocuğun kabakulak olma olasılığını formül kullanarak bulunuz.

4- Bir binom raslantı değişkeni n=10 ve p=0,8 parametrelerine sahipse,

a- Raslantı değişkenine ilişkin ortalama, varyans ve standart sapmayı bulunuz.

b- $P(X>8)$, $P(X<5)$ ve $P(2 \leq X<6)$ olasılıklarını hesaplayınız.

5- Bir araştırma şirketinin yaptığı bir araştırmada, insanların %37'sinin emekli olduktan sonra dışarıda yemeğe daha fazla para harcamayı planladıklarını ortaya çıkarmıştır. Rasgele seçilen 7 kişi için aşağıdaki olasılıkları bulunuz.

a- 4 kişinin emekli olduktan sonra dışarıda yemeğe daha fazla para harcamayı planlaması olasılığını bulunuz.

b- 2 kişiden azının emekli olduktan sonra dışarıda yemeğe daha fazla para harcamayı planlaması olasılığını bulunuz.

c- 3 kişiden fazlasının emekli olduktan sonra dışarıda yemeğe daha fazla para harcamayı planlaması olasılığını bulunuz.

6- Bir müdür, şirketinde çalışan 15 kişiye bilgisayar almayı düşünüyor. Bilgisayar firmasına gidip rasgele 4 makineyi denemeye karar veriyor. Bilgisayarlardan en az birinin hasarlı çıkması halinde, başka bir firmayı denemeye karar veriyor.

a- Bilgisayarlardan %10'u hasarlı ise, 1 ve daha fazla hatalı çıkması halinde başka bir firmayı deneme olasılığını hesaplayınız.

b- Sadece %5'i hasarlı iken, 1 ve daha fazla hatalı çıkması halinde başka bir firmayı deneme olasılığını hesaplayınız.

7- Bir araştırmaya göre yetişkinlerin %82'si, çocuklarının her isteğinin yapılmamasının gerekliliğini savunuyorlar. Rasgele 6 yetişkin seçildiğinde, aşağıdaki olasılıkları hesaplayınız.

a- 5 yetişkinin, çocuklarının her isteğinin yapılmamasını savunması olasılığını bulunuz.

b- Üç ya da daha az sayıda yetişkinin, çocuklarının her isteğinin yapılmamasını savunması olasılığı nedir?

c- 2'den fazla yetişkinin, çocuklarının her isteğinin yapılmamasını savunmasının olasılığı nedir?

8- Yapılan bir ankete göre, ilkokul öğretmenlerinin %30'unun internet kullandığını belirtmektedir. Rasgele 5 öğretmen seçildiğinde,

a- Bu öğretmenlerden internet kullanıyor olanlarının beklenen sayısını bulunuz.

b- İnternet kullanan bu öğretmenlerin varyansını ve standart sapmasını bulunuz.

9- Bir bölgede ayda ortalama olarak 2 kez elektrik kesilmektedir. Elektrik kesintilerinin, Poisson dağılımı gösterdiği kabul edilirse,

a- Hiç kesintinin olmaması olasılığını bulunuz.

b- En az 2 ve en çok 3 kesintinin olması olasılığını hesaplayınız.

10- Bir X raslantı değişkeni $\lambda=5$ parametresiyle Poisson dağılımı gösteriyorsa aşağıdaki olasılıkları bulunuz.

a- $P(X=4)$

b- $P(X>3)$

c- $P(X\leq 2)$

11- Bir garson 15 dakika içerisinde ortalama 5 müşteriye hizmet vermektedir. Bu garsonun 15 dakika içerisinde,

a- 3 müşteriye hizmet vermesi olasılığını bulunuz.

b- En az 4 müşteriye hizmet vermesi olasılığını bulunuz.

c- En fazla 6 müşteriye hizmet vermesi olasılığını bulunuz.

12- Bir otoyolda 1 ay içerisinde ortalama 8 kaza olmaktadır. Bu otoyolda 1 ay içerisinde,

a- 5 kaza olma olasılığını bulunuz.

b- En az 4 kaza olma olasılığını bulunuz.

c- En fazla 6 kaza olma olasılığını bulunuz.

13- Bir sınıftaki öğrencilerin istatistik dersinden aldıkları notlar ortalaması 76 ve standart sapması 8 olan normal dağılıma uyduğuna göre,

a- Öğrencilerin 90'dan yukarı not alma olasılığını bulunuz.

b- Öğrencilerin 60'dan yukarı not alma olasılığını bulunuz.

14- Bir kahve firması öğütülmüş kahveleri çok küçük paketler halinde piyasaya sürmektedir. Kahve paketlerinin standart sapması 0,3 gr olan bir normal dağılım gösterdiği varsayılıyor. Eğer kahve paketlerinin %5'inin ağırlığı 12,492 gr'dan fazla ise kahve paketlerinin ortalama ağırlığı ne kadardır?

15- Bir fabrika ampul üretmektedir. Üretilen ampullerin ömrü 800 saat ortalama ve 40 saat standart sapma ile normal dağılıma sahiptir. 100 tane ampul rasgele olarak ele alındığında, bunların kaç tanesinin ömrü, 778 saat ile 834 saat arasında olacaktır?

16- Büyük bir otomobil fabrikasından haftalık olarak talep edilen otomobillerin ortalama 5000 ve 300 standart sapma ile yaklaşık olarak normal dağıldığı varsayılmaktadır. Eğer otomobil fabrikasının hafta başında deposunda 5300 oto varsa, bunların hepsini satmamasının olasılığı nedir?

17- Bir market kendi ürettiği deterjanını naylon torbalarla piyasaya sürmektedir. Naylon torbaların ağırlıkları standart sapması 0,4 kg olarak bir normal dağılım göstermektedir. Naylon torbaların ağırlığının 5,588 kg'dan fazla olma olasılığı 0,07 ise, naylon torbaların ortalama ağırlığı nedir?

18- Bir sınavda alınan notlar 76 ortalama ve 15 standart sapma ile normal dağılımlıdır. Öğrencilerin %15'i A notunu almıştır. A notunu alabilmek için en az alınması gereken not kaçtır?

19- Yaşlılar yurdunda kalan 75 kişinin yaş ortalaması 76 ve varyansı 9 ile normal dağılım göstermektedir.

a- Yaşı 74'den az olan yaşlıların olasılığını bulunuz.

b- Yaşı 72 ile 73 arasında olan yaşlıların olasılığını bulunuz.

20- 800 öğrencinin boylarının ortalaması 1,75 cm ve 1,2 cm standart sapma ile normal dağılıma uymaktadır.

a- Boyları 1,68 cm ile 1,80 cm arasında olma olasılığını bulunuz.

b- Boyları 1,82 cm'den uzun olan öğrencilerin sayısını bulunuz.

6. ALTINCI BÖLÜM

6.1. GİRİŞ

6.2. MERKEZİ LİMİT TEOREMİ

6.3. ÖRNEKLEME VE ÖRNEKLEM DAĞILIMLARI

PROBLEMLER

6.1. GİRİŞ

İstatistikte temel amaç, kitlenin tümü hakkında bilgi edinmektir. Önceki konularda da gördüğümüz gibi incelenen toplumun tümü kitle, kitleyi temsil edebilecek alt gruba da örneklem denir. Kitleye ait bilgi (ortalama, varyans gibi) sabit bir değer olup, genellikle bilinemez. Örneklem bilgisi ise, seçilen örneklemden örnekleme değişir. Öyleyse bu bilgilerin bir dağılımı olacaktır ki bu dağılıma örneklem dağılımı denilmektedir. Bu bölümün amacı **örneklem dağılımlarının** özelliklerini incelemektir. Öncelikle örneklem dağılımlarından ortalamaların dağılımı için yararlı bir teorem olan merkezi limit teoremine değinilecektir.

6.2. MERKEZİ LİMİT TEOREMİ

Merkezi limit teoremi, kitleye ait dağılım bilinmediğinde ya da kitlenin dağılımı normal dağılım olmadığında, normal dağılımdan yararlanarak olasılık hesaplamak için kullanılan bir teoremdir. Uygulamada büyük kolaylık sağlayan bu teorem aşağıda verilmiştir:

Teorem 6.1: X raslantı değişkeni ortalaması μ, varyansı σ^2 olan herhangi bir dağılıma sahip olsun. Bu kitleden örneklem büyüklüğü n olan örneklemler çekilsin. n arttıkça, bu örneklemlerin ortalamalarının dağılımı, ortalaması μ, varyansı σ^2/n olan bir normal dağılıma yakınsar. Başka bir gösterim ile

$$\lim_{n\to\infty} f(\overline{X}) \sim N(\mu;\sigma^2/n)$$

dir (Genelde, $n\geq 30$ olması $\overline{X}$'nın dağılımının normal kabul edilmesi için yeterlidir).

Binom Dağılımının Normal Dağılıma Yaklaşımı

X raslantı değişkeni binom dağılımı gösteriyor ise, daha önce de belirtildiği gibi dağılımının ortalaması E(X)=np ve varyansı V(X)=npq'dır. Örneklem büyüklüğü $n\geq 30$ olduğunda, merkezi limit teoremine göre,

$$Z = \frac{X - E(X)}{\sqrt{V(X)}} = \frac{X - np}{\sqrt{npq}} \tag{6.1}$$

olur. Yani binom dağılımı Eşitlik (6.1) dönüşümü sonrası ortalaması 0, varyansı 1 olan standart normal dağılıma yakınsar. Bu yaklaşım, $P(a\leq X\leq b)$ olasılığını bulmak için yararlıdır. Bu olasılık için;

$$P(a \le X \le b) = P\left[\frac{a-np}{\sqrt{npq}} \le \frac{X-np}{\sqrt{npq}} \le \frac{b-np}{\sqrt{npq}}\right]$$

$$= P\left[\frac{a-np}{\sqrt{npq}} \le Z \le \frac{b-np}{\sqrt{npq}}\right] \qquad (6.2)$$

eşitliği yazılabilir. Ancak, X kesikli bir değişkendir, sürekli olan normal dağılımla olasılığın yaklaşık değerini hesaplayabilmek için; a ve b yerine sırasıyla (a-0,5) ve (b+0,5) düzeltme terimleri kullanılır ve bu durumda $P(a \le X \le b)$ olasılığı,

$$P(a \le X \le b) = P\left[\frac{a-0,5-np}{\sqrt{npq}} \le Z \le \frac{b+0,5-np}{\sqrt{npq}}\right] \qquad (6.3)$$

olur, böylece hesaplanan olasılık değeri Binom dağılımdan hesaplanacak olasılık değerine çok yakın olur, bu yol ile olasılık hesabı daha kolaydır.

Örnek 6.1: Bir tekstil fabrikasında aktif çalışan makinelerin belli bir ayda onarım gerektirme olasılığı 0,2 olarak belirlenmiştir. Fabrikada bu tür makinelerden 900 tane vardır**.** Belli bir ayda 200'den çok makinenin onarım gerektirmesi olasılığı nedir?

$E(X)=np=900\times0,2=180$

$V(X)=npq=900\times0,2\times0,8=144$

$P(X>200)=P\left(Z>\frac{200-180}{\sqrt{144}}\right)=P(Z>1,67)=0,5-0,4525=0,0475$

olur.

Düzeltme terimleri kullanılırsa,

$P(X>200)=P\left(Z>\frac{200-0,5-180}{\sqrt{144}}\right)=P(Z>1,63)=0,5-0,4484=0,0516$

olur.

Örnek 6.2: Bir hastane, hastalarının %25'inin borçlarını en az bir ay gecikme ile ödediklerini saptamıştır. Rasgele 45 ödeme seçilmiştir.

a- En az bir ay gecikmeli ödemelerin örneklemde 10'dan az olma olasılığı nedir? (Süreklilik düzeltmesi ile)

$E(X)=45\times0,25=11,25$

$V(X)=45\times0,25\times0,75=8,4375$

$$P(X<10)=P\left[Z<\frac{10+0{,}5-11{,}25}{\sqrt{8{,}4375}}\right]=P(Z<-0{,}26)=0{,}5-0{,}1025=0{,}3974$$

b- En az bir ay gecikme ile 12 ile 15 arasında olma olasılığı nedir? (Süreklilik düzeltmesi ile)

$$P(12<X<15)=P\left[\frac{12-0{,}5-11{,}25}{2{,}9047}\le Z\le\frac{15+0{,}5-11{,}25}{2{,}9047}\right]$$

$$=P(0{,}09<Z<1{,}46)=0{,}4279-0{,}0359=0{,}392$$

Poisson Dağılımının Normal Dağılıma Yaklaşımı

X raslantı değişkeni poisson dağılımı gösteriyor ise, daha önce de belirtildiği gibi dağılımın ortalaması E(X)=λ ve varyansı V(X)= λ'dır. λ'nın büyük değerleri için merkezi limit teoremine göre,

$$Z=\frac{X-E(X)}{\sqrt{V(X)}}=\frac{X-\lambda}{\sqrt{\lambda}} \tag{6.4}$$

yazılabilir. Buradan P(a≤X≤b) olasılığı;

$$P(a\le X\le b)=P\left[\frac{a-\lambda}{\sqrt{\lambda}}\le Z\le\frac{b-\lambda}{\sqrt{\lambda}}\right] \tag{6.5}$$

dır. Süreklilik düzeltmesi yapılarak aynı olasılığın;

$$P(a\le X\le b)=P\left[\frac{a-0{,}5-\lambda}{\sqrt{\lambda}}\le Z\le\frac{b+0{,}5-\lambda}{\sqrt{\lambda}}\right] \tag{6.6}$$

eşitliğine yakınsadığı kabul edilebilir.

Örnek 6.3: Bir bankanın belli bir bölgedeki bankamatiği öğleden önceki bir saat içinde ortalama 45 kez kullanılmaktadır.

a- Öğleden önceki bir saat içinde bu bankamatiğin 70'den çok kullanılma olasılığı nedir?

$$P(X>70)=P\left[Z>\frac{70-0{,}5-45}{\sqrt{45}}\right]=P(Z>3{,}65)=0{,}5-0{,}4999=0{,}001$$

b- Bu bir saat içinde bu bankamatiğin 50'den az kullanılma olasılığı nedir?

$$P(X<50)=P\left[Z<\frac{50+0{,}5-45}{\sqrt{45}}\right]=P(Z<0{,}82)=0{,}5+0{,}2939=0{,}7939$$

c- Bu bir saat içinde kullanılma sayısının 50 ile 70 arasında olma olasılığı nedir?

$$P(50<X<70)=P\left[\frac{50-0{,}5-45}{\sqrt{45}}<Z<\frac{70+0{,}5-45}{\sqrt{45}}\right]=P(0{,}67<Z<3{,}8)=0{,}4999-0{,}2486=0{,}2513$$

6.3. ÖRNEKLEME VE ÖRNEKLEM DAĞILIMLARI

Örneklem dağılımları örnek ortalaması, örnek varyansı gibi örneklemden hesaplanacak istatistikler için düşünülebilir. Örnek 6.4, örneklem ortalamasının dağılımını incelemek için verilmiştir.

Örnek 6.4: X: 1,2,3,4,5 olan bir kitle olsun (Düzgün kesikli).

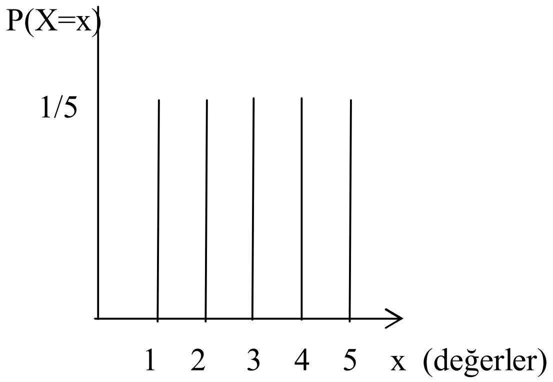
Kitleye ait ortalama ve varyans sırasıyla,

$$\mu=(1+2+3+4+5)/5=3$$

$$\sigma^2 = \frac{(-2)^2 + (-1)^2 + (0)^2 + (1)^2 + (2)^2}{5} = 2$$

olur.

X'in dağılımı düzgün, kesikli bir dağılımdır ve olasılık dağılımının grafiği aşağıda verilmiştir.



Şimdi bu kitleden, iadesiz n=3'lük örnekler alıp $\overline{X}$ değerlerini örneklerden hesaplayalım. Bu kitleden üçlü gruplar, $\binom{5}{3}=10$ farklı şekilde seçilebilir. Tablo 6.1'de bu farklı seçimler ve örneklemlere ilişkin ortalamalar verilmiştir. Bu örnekte kitle çok küçüktür ve yerine iade etmeden yapılacak örneklemede oluşturulabilecek örneklem sayısı 10'dur. Uygulamada 10 gibi az sayıda örneklem oluşturulması mümkün olmaz. Örneğin kitle büyüklüğü N=10000 olan bir kitleden n=100 olan bir örneklem çekileceği düşünülürse, oluşacak örneklemleri bu örnekteki gibi yazmak mümkün olmaz. Bu küçük örnek, konuyu daha iyi açıklayabilmek için verilmiştir.

Tablo 6.1 1,2,3,4,5 Kitlesinden Çekilen Üç Birimlik Örneklemler ve Ortalamaları

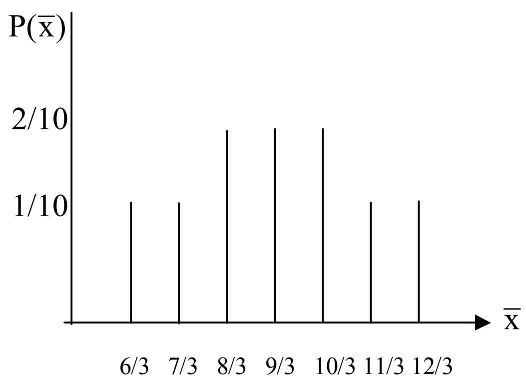
Örneklem	**Ortalama**	**Örneklem**	**Ortalama**
(1,2,3)	6/3	(2,4,5)	11/3
(1,2,4)	7/3	(1,3,4)	8/3
(1,2,5)	8/3	(1,3,5)	9/3
(2,3,4)	9/3	(1,4,5)	10/3
(2,3,5)	10/3	(3,4,5)	12/3

Tablo 6.1 incelendiğinde, kitleden çekilen örneklem ortalamalarının farklı olduğu ve bir dağılımının olacağı görülür. Ortalama 6/3 değerini alan 1, ortalama 7/3 değerini alan 1, ortalama 8/3 değerini alan 2, ortalama 9/3 değerini alan 2, ortalama 10/3 değerini alan 2, ortalama 11/3 değerini alan 1, ortalama 12/3 değerini alan 1 örneklem bulunmaktadır. Ortalamaların almış olduğu bu değerler olasılık fonksiyonu olarak ifade edilirse, aşağıdaki eşitlikler yazılabilir:

$$P(\overline{X} = 6/3) = 1/10\,;\, P(\overline{X} = 7/3) = 1/10\,;\, P(\overline{X} = 8/3) = 2/10\,;\, P(\overline{X} = 9/3) = 2/10$$

$$P(\overline{X} = 10/3) = 2/10\,;\; P(\overline{X} = 11/3) = 1/10\,;\; P(\overline{X} = 12/3) = 1/10$$

Ortalamaların olasılık dağılımına ilişkin dağılım grafiği Şekil 6.1'de verilmiştir.



Şekil 6.1 1,2,3,4,5'e İlişkin Ortalamalar Grafiği

Örneklem Ortalamasının Örneklem Dağılımı

Örnek 6.4'den görüldüğü gibi örneklem ortalaması bir dağılıma sahiptir. Ortalaması μ, varyansı σ^2 olan bir kitleden gözlemleri $X_1, X_2, ..., X_n$ olan n büyüklüğünde bir örneklemin çekildiği düşünülsün. Burada $\overline{X} = \sum_{j=1}^{n} X_j / n$ olan, $\overline{X}$'nın örneklem dağılımı incelenecektir.

Bir dağılım ortalama ve varyansı ile tanımlanır. Dağılımın ortalaması raslantı değişkenlerinin beklenen değerleri alınarak bulunur. Bu durumda $E(\overline{X})$'in bulunması gerekir. Önceki bilgilerimizden aşağıda verilen eşitlik yazılabilir:

$$E(\overline{X}) = E\left(\frac{\sum_{j=1}^{n} X_j}{n}\right) \tag{6.7}$$

Eşitlik (6.7)'de beklenen değer kuralları uygulanırsa,

$$E\left(\frac{\sum_{j=1}^{n} X_j}{n}\right) = \frac{1}{n}\left[E(X_1) + E(X_2) + ... + E(X_n)\right]$$

yazılabilir. Tüm X_j'ler ortalaması μ, varyansı σ^2 olan bir kitleden çekilmiş bir raslantı değişkeni olduğu için, her birinin beklenen değeri μ olup, aşağıdaki ifade elde edilir:

$$E\left(\frac{\sum_{j=1}^{n} X_j}{n}\right) = \frac{1}{n} n\mu$$

Sonuç olarak;

$$E(\overline{X}) = \mu \tag{6.8}$$

elde edilir. Bu sonuçtan, ortalaması μ olan bir kitleden ard arda örneklemler çekildiğinde, bu örneklemlerin ortalamalarının ortalamasının, kitle ortalamasına yakınsadığı söylenebilir.

Örnek 6.5: Örnek 6.4'de verilen problemde elde edilen sonucun doğruluğu gösterilsin.

Tablo 6.2 Örnek 6.4'de Seçilen Örneklemler için Ortalama ve Varyans Hesaplaması

$\overline{X}=\overline{x}$	$P(\overline{X}=\overline{x})$	$\overline{x}P(\overline{X}=\overline{x})$	$\overline{x}^2P(\overline{X}=\overline{x})$
6/3	1/10	6/30	36/90
7/3	1/10	7/30	49/90
8/3	2/10	16/30	128/90
9/3	2/10	18/30	162/90
10/3	2/10	20/30	200/90
11/3	1/10	11/30	121/90
12/3	1/10	12/30	144/90

$$\sum \overline{x}P(\overline{X}=\overline{x})=\frac{90}{30} \qquad \sum \overline{x}^2P(\overline{X}=\overline{x})=\frac{840}{90}$$

Örneklem ortalamalarının ortalaması,

$$\mu_{\overline{X}}=\frac{90}{3}=3=\mu$$

dır. Burada $\mu_{\overline{X}}$'nın μ olduğu görülmektedir.

Kitleyi tanımlayan ikinci parametre varyanstır. Önceki konularda varyans ile ilgili gördüğümüz tanımlar ve özelliklerden yararlanarak örneklem ortalamasının varyansı;

$$\begin{aligned}
\mathrm{Var}\,(\overline{X}) = \sigma^2_{\overline{X}} &= \mathrm{Var}\left(\frac{1}{n}\sum_{j=1}^{n} X_j\right) \\
&= \frac{1}{n^2}\mathrm{Var}\left(\sum_{j=1}^{n} X_j\right) \\
&= \frac{n\sigma^2}{n^2} \\
&= \frac{\sigma^2}{n} \qquad (6.9)
\end{aligned}$$

olarak bulunur.

Örneklem ortalamasının varyansı σ^2/n olduğundan, n örneklem büyüklüğü arttıkça, $\mathrm{Var}(\overline{X})=\sigma^2_{\overline{X}}$ küçülecek, yani örneklem ortalamalarının kitle ortalaması etrafındaki değişimi daralacak, kitleye ait daha doğru bilgi elde etme olanağı sağlanacaktır. $\overline{X}$'nın örneklem dağılımının standart sapması,

$$\sigma_{\overline{X}}=\sigma/\sqrt{n} \qquad (6.10)$$

olur. Kitleden örneklem seçimi, yerine iade edilerek ya da edilmeden gerçekleştirilir. Eğer yerine iade edilerek örneklem gerçekleştirilmiş ise sonsuz bir kitle söz konusudur. Bu durumda bireylerin seçimi birbirlerinden tamamen bağımsızdır. $\overline{X}$'nın örneklem dağılımının standart sapması da, $\sigma/\sqrt{n}$ olur. Kitleden örneklem seçimi, yerine iade etmeden yapılıyor ise ve n örneklem büyüklüğü, N kitle büyüklüğüne göre büyük ise, sonlu bir kitle söz konusu olacaktır. Burada kitleden bireylerin seçimi birbirlerinden bağımsız olamaz. Bu durumda $\overline{X}$'nın örneklem dağılımının varyansı,

$$\sigma^2_{\overline{X}} = \frac{N-n}{N-1}\left(\frac{\sigma^2}{n}\right) \tag{6.11}$$

dır. Burada $\sqrt{\dfrac{N-n}{N-1}}$ varyansa ait sonlu kitle düzeltme katsayısıdır. Eğer n örneklem büyüklüğü, N kitle büyüklüğüne göre çok küçükse, sonlu kitle düzeltme katsayısı bire yaklaşır ve etkisiz olur. Olasılık dağılımları ile ilgili hesaplamaları yapabilmek için raslantı değişkeninin dağılımının bilinmesi gerekir. Kitlenin dağılımı normal ise, bu kitleden çekilen örneklemlerin ortalamalarının dağılımı da normaldir. Ayrıca, Bölüm 6.2'de verilmiş olan merkezi limit teoreminden yararlanılarak, kitlenin dağılımı bilinmese bile n örneklem büyüklüğü arttıkça, (yaklaşık 30 ya da daha çok) $\overline{X}$'lerin dağılımının normal dağılıma yaklaştığı söylenebilir.

Normal dağılımla ilgili olasılık hesaplamaları için standart normal dağılım kullanılmaktadır. $\overline{X}$'leri standartlaştırmak için,

$$Z = \frac{\overline{X}-\mu}{\sigma_{\overline{X}}} = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \tag{6.12}$$

eşitliği kullanılır.

Örnek 6.6: Bilgisayar gereçleri üreten bir şirkette üretim oranları iki yıl takip edilmiştir. Bu üretimde artış yüzdesinin ortalaması %12,2 ve standart sapması %3,6 olan normal dağılıma uyduğu görülmektedir. Bu kitleden yerine iade edilerek 9 gözlemlik örneklem seçilmiştir. Örneklem ortalamasının %10'dan düşük olma olasılığı kaçtır?

$\mu=12{,}2$ $\sigma_{\overline{X}}=3{,}6$ $n=9$

$$P(\overline{X}\langle 10) = P\left(\frac{\overline{X}-\mu}{\sigma_{\overline{x}}}\langle \frac{10-12{,}2}{3{,}6/\sqrt{9}}\right) = P(Z\langle -1{,}83) = P(Z<0)-P(-1{,}83<Z<0)$$

$$=0{,}50-0{,}4664=0{,}0336$$

elde edilir. Ek 3 ile verilen standart normal dağılım tablosu kullanılarak yukarıda belirtildiği gibi örneklem ortalamasının %10'dan küçük olma olasılığı, 0,0336 olarak bulunur.

Örnek 6.7: Örnek 6.4'de kullanılan kitlenin varyansı $\sigma^2 = 2$ olarak bulundu. Kitleden çekilen örneklemlerin ortalamalarına ait varyans Tablo 6.2'nin dördüncü sütunundan,

$$\sigma_{\overline{x}}^2 = \frac{840}{90} - (3)^2 = \frac{1}{3}$$

dır. Aynı değeri kitle varyansını kullanarak da bulabiliriz. Bu hesaplama için Eşitlik (6.11)'den yararlanarak;

$$\sigma_{\overline{x}}^2 = \frac{\sigma^2}{n}\frac{N-n}{N-1} = \frac{2}{3}\left(\frac{5-3}{5-1}\right) = \frac{1}{3}$$

olarak elde edilir.

Örnek 6.8: Elektrik gereçleri üreten bir firmada üretilen aygıtların ohm cinsinden direnç ortalaması 92 ve standart sapması 3,6 olan normal dağılım özelliği göstermektedir. Üretilen aygıtlar her biri 4 gereçten oluşan kutulara konularak satılmaktadır. Her bir kutu bir örneklem olarak kabul edilirse,

a- Örneklemlerin ortalama direncinin örneklem dağılımının ortalamasını bulunuz?

$$E(\overline{X}) = \mu = 92$$

b- Örneklem ortalamalarının dağılımının varyansını bulunuz? (Yerine iade edilerek)

$$Var(\overline{X}) = \frac{\sigma^2}{n} = \frac{(3{,}6)^2}{4} = 3{,}24$$

c- Örneklem ortalamasının 93 ohm'u aşma olasılığı nedir?

$$P\left(Z\rangle \frac{93-92}{3{,}6/\sqrt{4}}\right) = P\left(Z\rangle \frac{1}{1{,}8}\right) = P(Z > 0{,}56) = 0{,}5 - 0{,}2123 = 0{,}2877$$

Örneklem Oranının Dağılımı

Bazı denemelerde X raslantı değişkeni belirlenen bir olayın gerçekleşme sayısı olarak karşımıza çıkar. Örneğin, bir bankanın yeni bir kredi kartı uygulamasında iki ay içinde bu uygulamayı kullanan kart sahiplerinin sayısı ya da bir partiye oy verenlerin sayısı vb gibi. Bu iki örnekteki kitle hakkında bilgi elde etmek isteyebiliriz. Ancak bu örneklerde görüldüğü gibi, kitleye ulaşmak oldukça güçtür. Kitle hakkında bilgi elde etmenin yolu daha öncede belirtildiği gibi kitleyi temsil edebilecek örneklem kullanmaktır. Örneklemden elde edilecek olan orana örneklem oranı denir ve $\hat{p}$ ile gösterilir. X örneklemde istenen olayın gerçekleşme sayısı, n örneklem büyüklüğü ise, örneklem oranı;

$$\hat{p} = \frac{X}{n} \tag{6.13}$$

olur. Bu tür olaylarda X raslantı değişkeninin dağılımı, binom dağılımına uymaktadır. Binom dağılımına ilişkin ortalama ve varyans kullanılarak, örneklem oranının ortalaması ve varyansı sırasıyla,

$$E(\hat{p}) = E\left(\frac{X}{n}\right) = \frac{1}{n}E(X) = \frac{1}{n}np = p \tag{6.14}$$

ve

$$Var(\hat{p}) = Var\left(\frac{X}{n}\right) = \frac{1}{n^2}Var(X) = \frac{1}{n^2}np(1-p) = \frac{p(1-p)}{n} \tag{6.15}$$

olur.

Ortalamaların örneklem dağılımlarında görüldüğü gibi basit rasgele örneklemede yerine iade ederek örnekleme gerçekleştirilmiş ise birimlerin örnekleme çıkışı birbirlerinden bağımsız ve eşit olasılığa sahip olacaktır. Ya da yerine iade etmeden örnekleme gerçekleştirilmiş olmasına rağmen, büyük bir kitleden küçük bir örneklem çekilmiş ise birimlerin örnekleme çıkışı birbirlerinden bağımsız ve eşit olasılığa sahip olduğu kabul edilebilir. Bu durumda, $Var(\hat{p}) = \sigma_{\hat{p}}^2 = \frac{p(1-p)}{n}$ olur. Basit rasgele örneklemede yerine iade etmeden örnekleme gerçekleştirilmiş ise ve örneklemin yapıldığı kitleye göre örneklem çok küçük değil ise, kitleden çekilen örneklemin birimleri birbirlerinden bağımsız olamazlar. Bu durumda $\hat{p}$ 'nın varyansı;

$$\sigma_{\hat{p}}^2 = \frac{p(1-p)}{n}\frac{N-n}{N-1} \tag{6.16}$$

olur. $\hat{p}$ 'nın örneklem dağılımının standart sapması,

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} \tag{6.17}$$

dir. Eşitlik (6.17) aynı zamanda $\hat{p}$'nın standart hatasıdır. Yerine iade etmeden yapılan örneklemede bu eşitlik,

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} . \sqrt{\frac{N-n}{N-1}} \tag{6.18}$$

biçiminde olur. N kitle büyüklüğüne göre, n örneklem büyüklüğü fazla ise uygulamada, n≥(0,05N) ise, $\sqrt{\frac{N-n}{N-1}}$ düzeltme katsayısı olan eşitliğin kullanılması önerilir. Eğer n<(0,05N) ise, $\sqrt{\frac{N-n}{N-1}}$ katsayısı 1'e yaklaşır, bu nedenle yerine iade etmeden yapılan örneklemede kullanılmaması sakınca çıkarmaz.

Örneklem oranına ait dağılımda belli değerler arasında bulunma, belli bir değerden küçük ya da büyük olma olasılıklarını bulabilmek için, n örneklem büyüklüğü 30 ya da daha fazla olan çalışmalar için kullanılacak olasılık dağılımı normal dağılım olur. Standartlaştırma işlemi ile standart normal dağılıma yakınsama yapılabilir. Standartlaştırma;

$$Z = \frac{\hat{p} - p}{\sigma_{\hat{p}}} \tag{6.19}$$

eşitliği ile yapılır.

Örnek 6.9: Bir şirket TV dizisine sponsor olmak için bir ön araştırma yapmıştır. Müşteri kitlesinden yerine iade edilerek 200 kişilik bir örneklem çekilmiştir. Önceki bilgilere göre bu TV dizisinin seyredilme olasılığı 0,40'dır.

a- Bu TV dizisini izleyenlerin örneklem oranının örneklem dağılımının ortalamasını hesaplayınız.

$$E(\hat{p}) = p = 0{,}40$$

b- $\hat{p}$'nın örneklem dağılımına ait varyansı hesaplayınız.

$$Var(\hat{p}) = \frac{p(1-p)}{n} = \frac{0{,}40.0{,}60}{200} = 0{,}0012$$

c- Örneklem oranının 0,60'dan fazla olma olasılığı nedir?

$$P(\hat{p} > 0{,}60) = P\left(Z > \frac{0{,}60 - 0{,}40}{\sqrt{0{,}0012}}\right) = P(Z > 5{,}71) = 0{,}5 - 0{,}499 = 0{,}001$$

d- TV dizisini izleyenlerin örneklem oranının 0,45'den küçük olma olasılığı nedir?

$$P(\hat{p} < 0{,}45) = P\left(Z < \frac{0{,}45 - 0{,}40}{\sqrt{0{,}0012}}\right) = P(Z < 1{,}44) = 0{,}5 + 0{,}4251 = 0{,}9251$$

e- TV dizisini izleyenlerin örneklem oranının 0,35 ile 0,45 arasında olma olasılığı nedir?

$$P(0{,}35 < \hat{p} < 0{,}45) = P\left(\frac{0{,}35 - 0{,}40}{\sqrt{0{,}0012}} < Z < \frac{0{,}45 - 0{,}40}{\sqrt{0{,}0012}}\right) = P(-1{,}44 < Z < 1{,}44) = 0{,}8502$$

Örnek 6.10: Bir şirket belli bir bölgedeki temsilcilerinin sattıkları ürün için müşterilerinin %30 ile bakım sözleşmesi imzalandığını saptamıştır. İki ay içinde ürünü satın alan 280 kişiyi rasgele bir örneklem olarak kabul edersek,

a- Bakım sözleşmesi imzalayanların örneklem oranının standart hatası kaçtır? (yerine iade edilerek)

$$\sigma_{\hat{p}} = \sqrt{\frac{(0{,}3)(0{,}7)}{280}} = 0{,}02739$$

b- Örneklem oranının 0,25'den büyük olma olasılığı kaçtır?

$$P(\hat{p} > 0{,}25) = P(Z > \frac{0{,}25 - 0{,}3}{0{,}02739}) = P(Z > -1{,}83) = 0{,}9664$$

c- Örneklem oranının 0,32'den küçük olma olasılığı kaçtır?

$$P(\hat{p} < 0{,}32) = P(Z < \frac{0{,}32 - 0{,}3}{0{,}02739}) = P(Z < 0{,}73) = 0{,}7673$$

Student t Dağılımı

Merkezi limit teoremine göre, μ ortalama ve σ^2 varyans ile dağılan her hangi bir kitleden çekilen örneklemlerde,

$$Z = \frac{\overline{X} - \mu}{\sigma_{\overline{x}}} = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}}$$

dönüşümünün standart normal dağılıma yakınsadığı belirtildi. Ancak her zaman kitle varyansı hakkında bilgi sahibi olmak mümkün değildir. Bu durumda standartlaştırmada kullanılacak varyans için, örneklemden elde edilen bilgiler kullanılır. n örneklem büyüklüğünün 30 ve 30'dan büyük olması durumunda;

$$Z = \frac{\overline{X} - \mu}{S / \sqrt{n}} \tag{6.20}$$

eşitliğinin standart normal dağılıma yakınsadığı kabul edilebilir. Burada S; örneklemden elde edilen standart sapmadır. Ancak n örneklem büyüklüğü 30'dan küçük olduğunda, bu yakınsama gerçekleşmez. Bu durumu ilk fark eden Gosett'tir. Bir İngiliz olan Gosset, biralar için kullanılan şerbetçi otu ile ilgili tarla deneme sonuçlarının standart normal dağılım göstermesini umduğu dönüşümlerin, küçük örneklemlerde bu özelliği göstermediğinin farkına varmış ve zamanının önemli istatistikçilerinden Pearson ile bu konuda mektuplaşmaya başlamıştır. Bugün kullandığımız Student t dağılımının temelleri de bu şekilde atılmıştır. Student t dağılımı n örneklem büyüklüğüne bağlı bir dağılım gösterir. Örneklem büyüklüğü arttıkça dağılım dikleşir ve standart normal dağılıma yaklaşır.

Tanım 6.1: $X_1, X_2, ... X_n$ raslantı değişkenleri ortalaması μ, varyansı σ^2 normal dağılımlı bir kitleden çekilmiş olsunlar. n<30 olduğunda;

$$t = \frac{\overline{X} - \mu}{S/\sqrt{n}} \qquad (6.21)$$

dönüşümü, (n-1) serbestlik dereceli t dağılımı gösterir. Burada S örneklemden hesaplanan standart sapmadır.

Student t dağılımının olasılık yoğunluk fonksiyonu;

$$f(sd, t) = \frac{\Gamma\left(\frac{sd+1}{2}\right)}{\Gamma\left(\frac{sd}{2}\right)\sqrt{\pi\, sd}} \left(1 + \frac{t^2}{sd}\right)^{-\frac{(sd+1)}{2}} \quad , \quad -\infty < t < +\infty \quad sd = 1,2,...,\infty \qquad (6.22)$$

olarak tanımlanmıştır. Bu eşitlikte, sd serbestlik derecesidir ve sd=n-1'dir. Serbestlik derecesi, bağımsız örneklem biriminin sayısıdır.

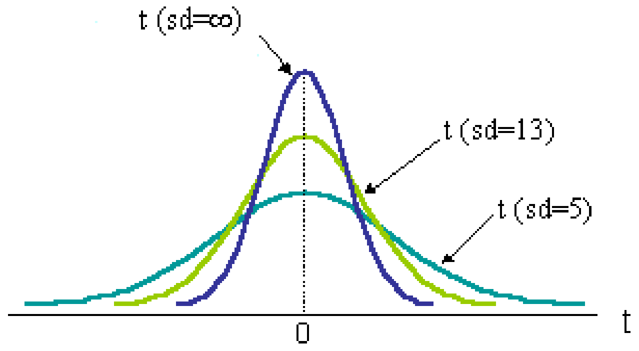
Örnek 6.11: Bir kitleden rasgele seçilmiş üç gözlemin ortalamasının 8 olduğunu varsayalım. Bu gözlemlerden ikisini istediğimiz gibi seçebiliriz. Diyelim ki; $X_1=7$; $X_2=8$ olsun. X_3'ü artık bağımsız seçmemiz olanaksızdır. Bu gözlem belirlenen ortalamayı verecek şekilde olmalıdır. Burada $X_3=9$'dur. Yani bağımsız gözlem sayısı 3-1=2 olacaktır.

Eşitlik (6.22)'de t, yatay eksende yer alacak t değeri, $\Gamma\left(\frac{sd+1}{2}\right)$ ve $\Gamma\left(\frac{sd}{2}\right)$ gama fonksiyonudur. Gosset tarafından 1908'de ilk kez oluşturulan Student t dağılımı

1926'da R.A.Fisher tarafından geliştirilmiştir. Bu dağılım normal dağılıma benzer bir dağılımdır. Dağılımın ortalaması sıfırdır, ortalamaya göre simetriktir, sd>2 olmak üzere dağılımın varyansı,

$$\sigma^2 = \frac{sd}{sd - 2} \tag{6.23}$$

dir. $sd \geq 30$ olduğunda, normal dağılıma çok yakınsar. Şekil 6.2'de farklı serbestlik derecelerine göre dağılım grafikleri verilmiştir:



Şekil 6.2 Farklı Serbestlik Derecelerine Göre t Dağılımı

t dağılımı da, normal dağılım gibi bir sürekli olasılık dağılımı olup, t'nin belli değerleri alması olasılığı f(sd, t) fonksiyonunun o değerleri için integral alınarak bulunabilir. Bu matematiksel işlem uygulamacıya bırakılmamış, t değerlerinin belli bölgelerde bulunma olasılıkları için tablolar hazırlanmıştır. Bilgisayar ortamında da bu olasılıklar hesaplanabilir.

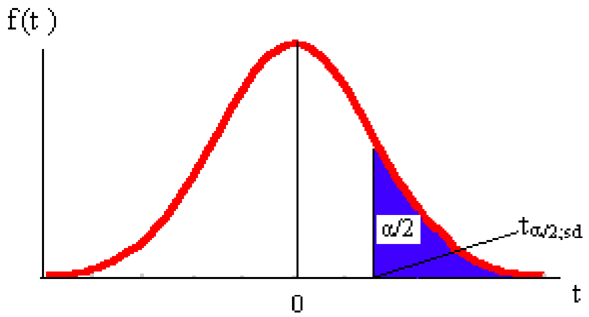
t Tablosu

Ek 5 ile verilen t tablosu, f(sd,t) olasılık yoğunluk fonksiyonundan oluşturulmuştur. Her bir serbestlik derecesi için Şekil 6.3'de gösterilen alanlar hesaplanmıştır. Şekil 6.3'de gösterilen alan tek yanlı olasılık içindir. Ek 5'de verilen t tablosunda olasılıklar, güven aralığı, tek yanlı ve iki yanlı olasılıklar olarak tablonun üst kısmında yer almışlardır. Tablo değerlerini göstermek için;

$t_{\alpha/2;sd}$ tek yanlı olasılıkta tablo değeri olarak,

$t_{\alpha;sd}$ iki yanlı olasılıkta tablo değeri olarak,

$t_{1-\alpha;sd}$ güven aralığı tablo değeri olarak kullanılacaktır.



Şekil 6.3 $P(t_{sd}>t_{\alpha/2;sd})=\alpha/2$

Bu tablodaki ikinci satır, tek yanlı olasılık olan, $P(t_{sd}>t_{\alpha/2;sd})=\alpha/2$ olasılıklarını hesaplar. Bu; serbestlik derecesine göre örneklemlerin, tabloda verilen $t_{\alpha/2;sd}$ değerden büyük olma olasılığının α/2 olduğunu gösterir. Tabloda ilk sütun sd serbestlik dereceleridir. İlk satır $P(-t_{\alpha/2;sd} < t_{sd} < t_{\alpha/2;sd}) = 1-\alpha$ olan güven aralığı olasılıklarını verir. Bu; sd serbestlik derecesindeki örneklemlerin $-t_{\alpha/2;sd}$ ile $t_{\alpha/2;sd}$ arasında bulunma olasılığının 1-α olduğunu gösterir. Üçüncü satır iki yanlı olasılık $[P(t_{sd} < -t_{\alpha/2;sd}) + P(t_{sd} > t_{\alpha/2;sd})]$'dir. Bu; sd serbestlik dereceli örneklemlerin $-t_{\alpha/2;sd}$ değerden küçük, $t_{\alpha/2;sd}$ değerden büyük olma olasılıkları toplamının α olduğunu gösterir.

Örnek 6.12: $P(t_{10} > t_{0,10;10}) = 0{,}10$ olan t değeri nedir? $t_{0,10;10}$ = 1,372'dir.

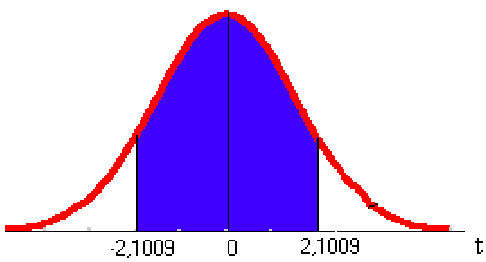
Örnek 6.13: $P(t_{13} > t_{0,1/2;13}) = 0{,}05$

α/2 = 0,05 dir. $t_{13,\,0,05}$ = 1,771'dir.

Örnek 6.14: a- sd = 18 ise P(-2,1009<t<2,1009)=?

b- $P(t_{15} > t_{0,025;15}) = 0{,}025 \Rightarrow t = ?$

a- sd=18 iken aşağıdaki alan toplamı isteniyor.



arada kalan alan toplamı için sd=18 için 2,1009'a karşılık gelen güven aralığı değerine bakılır. Bu değer 0,95'dir. P(-2,1009<t<2,1009)=0,95 olacaktır.

b- $P(t_{15} > t_{0,025;15}) = 0,025 \Rightarrow t_{0,025;15} = 2,1315$

Örnek 6.15: $P(?<t_{10}<?)=0,80$

Güven aralığında 0,80 ve sd=10 olduğu yerdeki değer 1,372 olduğu için $P(-1,372<t_{10,0.80}<1,372)=0,80$ olacaktır.

Örnek 6.16: $P\left(-\frac{\overline{X}-\mu}{S/\sqrt{25}} < -t_{24}\right) + P\left(\frac{\overline{X}-\mu}{S/\sqrt{25}} > t_{24}\right) = 0,20$ ise t nedir?

Bu değeri bulmak için sd=n-1=25-1=24 için, iki yanlı olasılık satırında 0,20 olan değere bakılır ki bu değer, 1,3178 olacaktır.

Örneklem Varyansının Dağılımı

Ortalaması μ olan, σ^2'nin bilinmediği bir kitle ele alınsın ve bu kitleden $X_1,X_2,...,X_n$ ile gösterilen n gözlemli bir örneklem çekilsin. Seçilen örneklemin kitle varyansı;

$$\sigma^2 = E(X-\mu)^2$$

dir. Kitle varyansı için, n büyüklüğündeki örneklemlerden hesaplanan $(X-\mu)^2$'nin ortalamasının alınması gerekmektedir. Ancak örneklemlerden yapılacak bu hesaplamada, kitleye ait μ bilinmeyecek, bunun yerine örnekleme ait $\overline{X}$ kullanılacaktır. Bu durumda $(X-\overline{X})^2$'nin ortalaması düşünülmelidir.

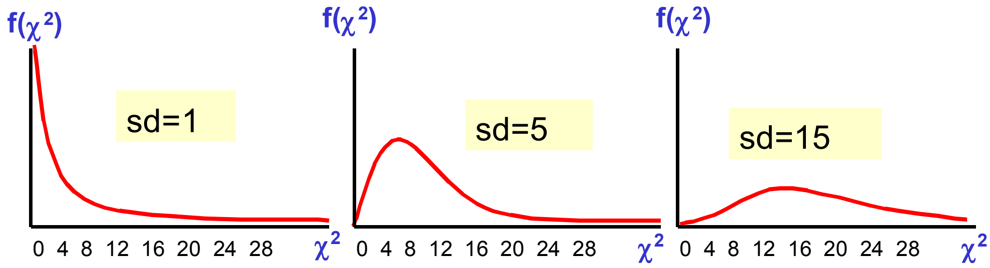
Bir kitleden $X_1,X_2,...,X_n$ rasgele çekilmiş örneklem ise bu örnekleme ilişkin varyans,

$$S^2 = \frac{1}{n-1}\sum_{j=1}^{n}(X_j-\overline{X})^2 \qquad (6.24)$$

ile bulunur. Burada $\sum_{j=1}^{n}(X_j-\overline{X})^2$'nın ortalaması alınmaktadır. Ancak burada n yerine, (n-1) kullanılmıştır. Bunun amacı Bölüm 7.2'de verilen $E(S^2)=\sigma^2$ olan, yansızlık şartını sağlamaktır. S^2 ya da $\sum_{j=1}^{n}(X_j-\overline{X})^2$ nasıl bir dağılım göstermektedir? Öncelikle bu dağılımda $X_1,X_2,...,X_n$'lerin normal dağılmış bir kitleden çekildikleri varsayımı sağlanmalıdır. Eğer bu varsayım sağlanıyor ise,

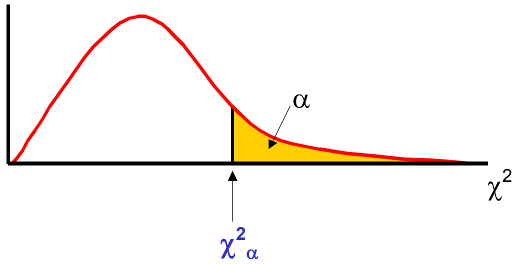
$$\frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{j=1}^{n}\left(X_j - \overline{X}\right)^2}{\sigma^2} \tag{6.25}$$

raslantı değişkeni (n-1) serbestlik dereceli (sd) Khi-kare (χ^2) dağılımı gösterir. χ^2 dağılımı, 0 ile $+\infty$ arasında tanımlı olup, simetrik bir dağılım değildir. χ^2 dağılımı serbestlik derecesine bağlı bir grafik gösterir. Şekil 6,4'de çeşitli serbestlik dereceleri için χ^2 çizimleri verilmiştir.



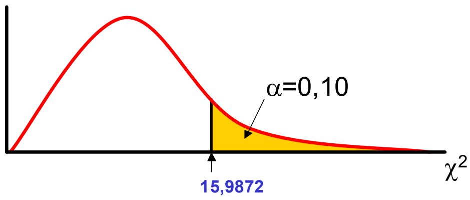
Şekil 6.4 Farklı Serbestlik Derecelerine Sahip χ^2 Dağılımı

Şekil 6.4'den görüldüğü gibi serbestlik derecesi arttıkça, χ^2 dağılımının simetrikliği de artar. χ^2 dağılımıyla ilgili değerler, kolaylık sağlaması nedeniyle tablo halinde verilmiştir. Bilgisayar ortamında bu fonksiyonlara ilişkin olasılıkları doğrudan veren paket programlar da vardır. Ek 6'da verilen χ^2 tablosunun ilk kolonunda sd, ilk satırında olasılık değerleri yer alır. Tablonun içinde yatay eksende yer alan χ^2 değerleri bulunur. Tablo, büyüklüğüne göre serbestlik derecesi, 1 ile 30 ya da 1 ile 200 arasında değişir. Olasılık değerleri, χ^2 ile $+\infty$ arasındaki alanı ya da 0 ile χ^2 arasındaki alanı verir. Bu kitap boyunca χ^2 ile $+\infty$ arasındaki olasılıkları veren tablo kullanılacaktır. İlgili grafik Şekil 6.5'de verilmiştir.



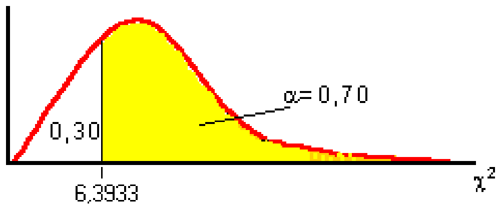
Şekil 6.5 χ^2 Dağılımının χ^2_α Yüzde Noktası

Örnek 6.17: P(χ^2_{10} > ?) = 0,10



Serbestlik derecesi 10 olan ve 0,10 olasılığını veren nokta için tabloya bakılır. Bu değer 15,9872'dir.

Örnek 6.18: P(χ^2_9< ?) = 0,30



Ek 6 ile verilen tablo, istenen değerden ∞'a kadar olan toplamı verdiğinden, 1-0,30=0,70'daki değere bakılır, bu değer 6,3933'dür.

Örnek 6.19: Bir müşteri kitlesinin yaş dağılımı standart sapması 3,6 olan normal dağılıma uymaktadır. 10 kişilik bir örneklem çekildiğinde, örneklem varyansının 30'dan büyük olma olasılığı kaçtır?

n=10, σ=3,6 ve σ^2=12,96

Aranan olasılık $P(S^2>30)$ olacaktır.

$$P\left(S^2>30\right)=P\left(\frac{(n-1)S^2}{\sigma^2}>\frac{(n-1)30}{\sigma^2}\right)$$

$$=P\left(\chi_9^2>\frac{9\times 30}{12{,}96}\right)=P\left(\chi_9^2>20{,}83\right)=?$$

Tablo kullanılarak olasılık bulunacak ise, Ek 6 ile verilen tablo kullanılır. Bu tabloda, 9 serbestlik derecesinde 20,83 değerine karşılık tam bir olasılık bulamayız. Ancak bu değeri içeren iki değer, aşağıda verildiği gibidir:

$$P(\chi_9^2>19{,}228)=0{,}025$$
$$P(\chi_9^2>20{,}83)=?$$
$$P(\chi_9^2>21{,}6660)=0{,}01$$

Aranan olasılık 0,025 ile 0,01 arasında olacaktır.

Bilgisayar ortamında paket programlarda χ^2 dağılımı var ise, χ^2 değerine karşı gelen olasılık tam olarak bulunabilir.

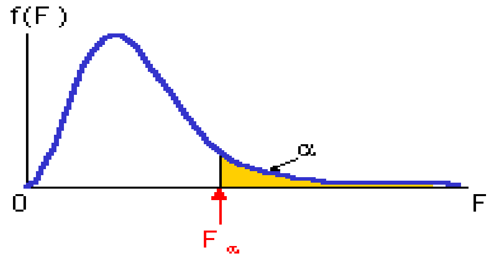
F Dağılımı

Örneklem dağılımlarından biri olan ve sık kullanılacak F dağılımı rasgele seçilmiş iki örneklemin, varyansları açısından aynı varyansa sahip olup olmadığının kontrolü ve sonraki bölümlerde verilecek olan varyans analizi alanında çok kullanılan bir dağılımdır.

sd_1 ve sd_2 serbestlik dereceli bağımsız iki χ^2 raslantı değişkeninin oranı F dağılımı gösterir.

$$F=\frac{\chi_1^2/sd_1}{\chi_2^2/sd_2} \tag{6.26}$$

Şekil 6.6’da F dağılımının grafiği verilmiştir.



Şekil 6.6 F Dağılım Grafiği

Şekil 6.6'dan görüldüğü gibi F dağılımı da, χ^2 dağılımı gibi 0 ile ∞ aralığında tanımlı değerler alır. Bu dağılım için de 0 ile F değerleri arasındaki olasılıkları veren tablolar oluşturulmuştur (Ek 7). Tablonun ilk satırı ve ilk sütunu serbestlik derecesini (sd) gösterir. Tablonun içindeki değerler F değerleridir. F oranlarında, payında büyük varyans, paydasında da küçük varyans kullanılır. Tabloda F değerine bakılırken, büyük varyans'ın sd'si ilk satırdan, küçük varyans'ın sd'si ilk sütundan bakılıp, kesiştikleri yerde bulunan F değeri tablo değeri olarak kullanılır. F dağılımı iki serbestlik derecesine bağlı olduğu için, bu hem sd hem de α olasılıklarını aynı tabloda gösterilmesini zorlaştırır, bu nedenle her olasılık için (α=0,05, α=0,01 vb.) ayrı tablolar bulunmaktadır. Genellikle α=0,05 ve α=0,01 kullanıldığı için, bu kitapta bu olasılıkları veren F tabloları verilmiştir. Bu dağılım için de bilgisayar ortamında istatistik paket programlarında F dağılımı tanımlanmıştır. İstenen olasılık ya da verilen olasılıklara karşı gelen F değerleri bulunabilir.

PROBLEMLER

1- Elektrik gereçleri üreten bir firmada üretilen aygıtların ohm cinsinden direnç ortalaması 105, standart sapması 5 olan normal dağılım özelliği göstermektedir. 50 adet aygıttan örneklem büyüklüğü 25 olan yerine iade edilmeden örneklemler çekildiğinde;

a- Örneklemin ortalama direncinin örneklem dağılımının ortalamasını bulunuz?

b- Örneklem ortalamasının varyansını bulunuz?

c- Örneklem ortalamasının 103 ohm ile 107 ohm arasında olma olasılığı nedir?

2- Belli bir grup tarafından izlenen bir müzik programının haftalık dinlenme süresinin ortalama ve standart sapması sırasıyla μ=4 saat ve σ=1 saat olan normal dağılım göstermektedir. Yerine iade edilerek rasgele seçilen 10 kişilik bir örneklem için ortalama dinlenme zamanının, a- Üç saatten az olma olasılığı nedir?

b- Bu programın ortalama dinlenme zamanının 3,5 saat ile 4,5 saat arasında gerçekleşme olasılığı nedir?

3- Kamu personelinin aylıklarındaki yıllık artış yüzdesi ortalaması 0,14 ve standart sapması 0,065 olan bir normal dağılıma uymaktadır. Bu kitleden yerine iade edilerek 25 kişilik bir örneklem seçilmiştir. Örneklem ortalamasının 0,10'dan düşük olma olasılığı kaçtır?

4- Bir denetçi 3000 dosya kontrol edecektir. Hatalı dosya sayısının ortalaması 68 ve standart sapması 15 olan bir normal dağılım gösterdiği varsayılsın. 180 dosya yerine iade etmeme koşulu altında seçilirse örneklem ortalamasının,

a- 66 ile 68 arasında olma olasılığı nedir?

b- 69'dan daha fazla dosyanın hatalı bulunması olasılığı ne olur?

5- Bir kuruluşta çalışan 900 kişinin gündelikleri 60 YTL ortalama ve 5 YTL standart sapma ile normal dağılıma sahiptir. Yerine iade edilmeden rasgele 100 kişi seçilmiştir.

a- Örneklem dağılımının ortalamasını ve varyansını bulunuz.

b- Örneklem ortalamasının 60 YTL ile 61,4 YTL arasında olma olasılığını bulunuz.

6- Bir üniversitedeki 3000 erkek öğrencinin ağırlıklarının ortalaması 68 kg ve standart sapması 3 kg olup normal dağıldığı varsayılmaktadır. Her biri 25 öğrenciden oluşan 80 örneklem yerine iade edilmeden alındığında, örneklemden kaç tanesinin ortalaması

a- 66,8 kg ile 68,3 kg arasında,

b- 66,4 kg'dan az bulunması beklenir.

7- Bir firma bisküvilerini paketler halinde piyasaya sürmektedir. Paketlerin ağırlığının standart sapması 0,3 gr olan normal dağılım gösterdiği varsayılıyor. Rasgele seçilen 9 paketin %5'inin ağırlığı 30 gr'dan fazla ise, bisküvi paketlerinin ortalama ağırlığı ne kadardır?

8- Bir şirket, pay senedine çevrilebilir tahvil çıkarmayı düşünmektedir. Yönetim, şimdiki ortaklardan %20'sinin öneri koşullarını çekici bulacağı kanısındadır. Bu kanının doğru olduğu düşünülsün. Şimdiki ortaklardan rasgele 130 örneklem yerine iade edilerek seçildiğinde,

a- Örneklem oranının ortalamasını ve varyansını bulunuz.

b- Örneklem oranının 0,18 ile 0,22 arasında olma olasılığını bulunuz.

9- Bir şirket TV dizisine sponsor olmak için bir ön araştırma yapmıştır. Müşteri kitlesinden 100 kişilik bir örneklem yerine iade edilerek çekilmiştir. Önceki bilgilere göre, bu TV dizisinin seyredilme olasılığı 0,65'dır.

a- Bu TV dizisini izleyenlerin örneklem dağılımının ortalaması kaçtır?
b- Örneklem oranının varyansı kaçtır?
c- Örneklem oranının 0,75'dan fazla olma olasılığı nedir?

10- Bir beyaz eşya firması, belli bir bölgedeki temsilcilerinin sattıkları ürün için müşterilerinin %45'i ile bakım sözleşmesi imzalandığını saptamıştır. İki ay içinde ürünü satın alan 144 kişi yerine iade edilmeden seçilen rasgele bir örneklem olarak kabul edilirse (N=200 alınız.),

a- Bakım sözleşmesi imzalayanların örneklem oranının standart sapması kaçtır?
b- Bakım sözleşmesi yapmak isteyenlerin örneklem oranının 0,50'den büyük olma olasılığı kaçtır?
c- Bakım sözleşmesi yapmak isteyenlerin örneklem oranının 0,36'den küçük olma olasılığı kaçtır?

11- Bir yatırım şirketinin elinde bulunan 100 hisse senedinin bir yıl önceki getiri oranlarının ortalaması 0,13 olan normal dağılım gösterdiğini bilmektedir. Yerine iade edilmeden 16 hisse senedi seçilmiş ve standart sapma 0,03 olarak bulunmuştur.

a- Bu şirketin hisse senetlerinin getiri oranlarının örneklem ortalamasının 0,148'den çok olma olasılığı nedir?
b- Bu şirketin hisse senetlerinin zarara (eksi değer alma) uğrama olasılığı nedir?
c- Bu şirketin hisse senetlerinin getiri oranlarının örneklem ortalamasının 0,1347 ile 0,148 arasında bulunma olasılığı nedir?

12- Belli bir üretim sürecinde üretilen malların %12'sinin kusurlu olduğu bilinmektedir. Bir günde üretilen 750 ürün arasından yerine iade edilmeden 450 tanesi rasgele seçilmiştir.

a- Örneklem oranının ortalama ve varyansını bulunuz.
b- Örneklem oranının 0,10 ile 0,15 arasında olma olasılığı nedir?

13- İşletmecilik mezunlarının %43'ü işletme ahlakı dersinin öğrencilere ahlak değerleri aktarmada yararlı olduğu görüşünü belirtmişlerdir. Yerine iade edilerek rasgele seçilen 80 kişinin bu görüşte olanlarının,

a- Ortalamasını ve varyansını bulunuz.

b- Bu görüşte olanlarının 0,50’den fazla olma olasılığını bulunuz.

14- Otomatik çalışan bir makine ile doldurulan şişelerin ağırlıklarının dağılımı ortalaması 500 gr ve standart sapması 10 gr olan normal dağılım göstermektedir. Yerine iade edilerek rasgele 4 şişe seçilmiştir.

a- Örneklem varyansının 154,72’den küçük olma olasılığını bulunuz.

b- Örneklem varyansının 78,87 ile 208,38 arasında olma olasılığını bulunuz.

15- Uygun tablolardan yararlanarak aşağıda verilen soruları, çizimlerini de yaparak cevaplayınız.

a- $P(Z> ?)=0,0375$

b- $P(t_6 >1,9432)=?$

c- $P(\chi^2_3 < 3,6649) = ?$

16- Günler itibarıyla elektrik kesintisi yapılan semtlerin sayılarına ilişkin ortalama 40’dır. Normal dağılım gösteren bu kitleden yerine iade edilerek rasgele 9 semt seçildiğinde örneklem varyansı 25 bulunmuştur. Verilere göre örneklem ortalamasının 41,84683’den küçük olma olasılığını bulunuz.

17- Bir araştırması kuruluşu, yaptığı anketlerde deneklerin %20’sinin cevaplarında samimi olmadığı görüşündedir. 500 kişiye yapılan anketlerden yerine iade edilmeden rasgele seçilen 100 kişilik bir örneklemde 27’den fazla kişinin samimi cevap vermeme olasılığı nedir?

18- Kitle varyansı 22 olan bir kitleden n=20 olan örneklemler yerine iade edilerek çekilmiştir. Örneklem varyansının hangi değerden büyük olması olasılığı 0,1’dir?

19- Aşağıda verilen soruda soru işaretli olan yerleri doldurunuz.

a- $P(\chi^2_8<5,5274)=?$

b- $P(\chi^2_5>11,0705)=?$

c- $P(?<\chi^2_4< 4,8784)=0,20$

20- Bir firma ürettiği çelik boruların çaplarının ortalamasının $\mu=3$ mm ve varyansının $\sigma^2=0,5$ mm^2 olduğunu savunmaktadır. Varyansın büyümesi kullanılamaz ürün oranını artırır. Müşterisi olan firma varyansının büyüklüğünün örneklemlerin 0,05’inde istenilmeyen varyansa sahip olmasını hoş görmektedir. Yerine iade edilerek rasgele seçilen 25 örneklem için varyans 0,7587 olarak bulunmuştur. Sizce müşteri isteği sağlanabilmiş midir?

7. YEDİNCİ BÖLÜM

7.1. GİRİŞ

7.2. NOKTA TAHMİNİ

7.3. ARALIK TAHMİNİ (GÜVEN ARALIĞI)

7.4. İSTENEN ÖRNEKLEM BÜYÜKLÜĞÜNÜN BELİRLENMESİ

PROBLEMLER

7.1. GİRİŞ

Örneklem seçilmesinin başlıca amacı, kitleye ait çıkarsamalar yapmak için gerekli olan bilgileri elde etmektir. Örneklem bilgisine dayanarak, kitle parametresinin alacağı olası değerin belirlenmesine çalışılır. Tahmin yapmak örneklem istatistiklerine ve bunların dağılımlarına bağlıdır. Bu bölümde parametrelerin tahmini için kullanılan nokta ve aralık tahminleri (güven aralığı) verilecektir.

7.2. NOKTA TAHMİNİ

Parametre için bir tek değer ile tahmin yapılmış ise bu, nokta tahmini olarak adlandırılır. Örneğin, DİE tarafından hane halkı gelir anketleri ile, Türkiye genelinde, ortalama gelirin nokta tahmin edicisi, örneklem sonucunda elde edilen ortalama değer nokta tahminidir. Nokta tahmini gösterimleri Tablo 7.1'de verilmiştir.

Tablo 7.1 Nokta Tahmin Gösterimleri

	Kitle Parametresi	**Tahmin Edici**
Ortalama	μ	$\overline{X}$
Varyans	σ^2	S^2
Standart Sapma	σ	S
Oran	p	$\hat{p}$

Örnek 7.1: 2003 yılına ait rasgele seçilen 10 hafta için Ereğli Demir Çelik hisselerinin getirileri Tablo 7.2'de verilmiştir. Bu bilgilerden yararlanarak seçilen 10 hafta için getirilere ait ortalama, varyans ve standart sapmayı bulunuz. Getirisi 0,04 üzerindeki getiriler için örneklem oranını bulunuz.

Tablo 7.2 Seçilen 10 Hafta İçin Getiri Değerleri

Seçilen Haftalar	Getirileri
5. Hafta	0,0458
18. Hafta	0,0014
20. Hafta	0,0542
31. Hafta	0,0119
35. Hafta	0,0002
40. Hafta	0,0081
43. Hafta	-0,102
48. Hafta	0,0756
50. Hafta	-0,0553
51. Hafta	-0,007

Çözüm: $n = 10 \qquad \sum_{j=1}^{10} X_j = 0{,}0329 \qquad \overline{X} = \frac{\sum_{j=1}^{10} X_j}{10} = \frac{0{,}0329}{10} = 0{,}00329$

Rasgele seçilen 10 hafta için getirilere ait ortalama 0,00329'dur. Verilere ilişkin varyans ve standart sapma da sırasıyla,

$$S^2 = \frac{1}{9}\left[(0{,}0458 - 0{,}00329)^2 + (0{,}0014 - 0{,}00329)^2 + \ldots + (-0{,}007 - 0{,}00329)^2\right] = 0{,}0027$$

ve

$$S = 0{,}052029$$

olarak elde edilir.

Getirisi 0,04'den fazla olan hafta sayısı X =3'dür. Böylece orana ait nokta tahmini,

$$\hat{p} = \frac{X}{n} = \frac{3}{10} = 0{,}30$$

olacaktır. Bu örnekte bulunan $\overline{X} = 0{,}00329$, $S^2 = 0{,}0027$, $S = 0{,}052029$ ve $\hat{p} = 0{,}30$ sırasıyla μ, σ^2, σ, p için birer nokta tahminleridir.

Nokta Tahmin Edicilerin Özellikleri

İyi bir tahmin edicinin bazı özellikleri taşıması gerekir. Bu özellikler aşağıda verilmiştir:

Tutarlılık

Herhangi bir parametre θ ve bu parametre için tahmin edici $\hat{\theta}$ ise bu özellik, örneklem büyüklüğü arttıkça, tahmin edicinin değerinin parametreye eşit olması olasılığının yüksek olmasıdır. $\hat{\theta}$, n büyüklüğündeki bir örneklemden elde edilen bir tahmin edici, θ'da tahmin edilecek parametre olarak ifade edilirse, aşağıdaki eşitlik yukarıda verilen tanımı ifade etmektedir:

$$\lim_{n \to \infty} P\left(\left|\hat{\theta} - \theta\right| = 0\right) = 1 \tag{7.1}$$

Yansızlık (Sapmazlık)

$\hat{\theta}$'nın dağılımının ortalamasının, θ'ya eşit olmasına yansızlık denir ve aşağıdaki ifade ile gösterilir:

$$E\left(\hat{\theta}\right) = \theta \tag{7.2}$$

Yansızlıkla ilgili iki önemli nokta aşağıda verilmiştir:

1- Örneklem ortalaması, örneklem varyansı ve örneklem oranı kendilerine karşılık gelen kitle parametrelerinin yansız tahmin edicileridir.
2- Örneklem standart sapması, kitle standart sapmasının yansız bir tahmin edicisi değildir. Eğer bir tahmin edici yansız değilse, sahip olduğu sapmanın bulunması için,

$$\text{Yan}(\hat{\theta}) = E(\hat{\theta}) - \theta \tag{7.3}$$

eşitliği kullanılır.

Etkinlik

Aynı parametrenin yansız iki tahmin edicisinden standart sapması küçük olan daha etkindir. θ parametresinin birden fazla yansız kestiricisi bulunabilir. Bunlar arasında etkinliği daha yüksek olanının tercih edilmesi gerekir. Bir tahmin edicinin diğerine göre göreli etkinliği (GE),

$$GE = \frac{\text{Var}(\hat{\theta}_2)}{\text{Var}(\hat{\theta}_1)} \tag{7.4}$$

ile verilir.

Örnek 7.2: $X_1, X_2, ..., X_n$ ortalaması μ, varyansı σ^2 olan normal dağılımdan rasgele çekilmiş bir örneklem olsun. $\overline{X}$ örneklem ortalaması, kitle ortalamasının yansız bir tahmin edicisidir ve varyansı,

$$\text{Var}(\overline{X}) = \frac{\sigma^2}{n}$$

ile verilir. Buna karşın aynı örneklemin ortancasına ait yansız tahmin edicisinin varyansı,

$$\text{Var}(\overline{X}') = \frac{\pi}{2}\frac{\sigma^2}{n} = 1{,}57\frac{\sigma^2}{n}$$

dir. Bu durumda konum ölçüsü olarak örneklem ortancasının ortalamaya göre göreli etkinliği,

$$GE = \frac{\text{Var}(\overline{X}')}{\text{Var}(\overline{X})} = 1{,}57$$

olur. Bu sonuç ortalamanın ortancaya göre daha etkin bir tahmin edici olduğunu göstermektedir.

7.3. ARALIK TAHMİNİ (GÜVEN ARALIĞI)

Bir önceki konuda kitle için ortalama, oran ve varyans gibi bilgiler tek bir değer olarak tahmin edildi. Bazı durumlarda, parametre için tek bir tahmin vermek yeterli olabilir. Ancak uygulamada, kitle için alınan bilgilerin belli aralıklarda olma olasılıkları da sık istenilen bilgilerdir. Örneğin; bir yatırımcı gelecek yılın planlarını yaparken elinde bulunan hisse senetlerinin %90 olasılıkla hangi değerler arasında bulunacağını öğrenmek isteyebilir. Ya da bir sigorta yöneticisi, gelecek yıl hasar tazminat oranının belli bir olasılıkla hangi değerler arasında olacağını bilmek ister. Bu soruları nokta tahmini ile çözümlemek mümkün olmaz. Tahmini yapılacak bilginin aralarında kalacağı sınırlar elde edilmek istendiğinde, **aralık tahmini (güven aralığı)** elde etmek gerekir. Aralık tahminin kullanılmasının bir başka gerekçesi, n örneklem büyüklüğünde olan artışın, nokta tahminlerinde doğrudan yansıyamamasıdır. Örneğin; tüm koşullar aynı kalmak şartı ile bir şirkette hatalı hesapların oranını belirlemek için n örneklem büyüklüğü 10 alındığında 1 hatalı hesap, 100 alındığında 10 hatalı hesap bulunursa, her iki durumda da hatalı hesap oranı 0,1 olur. Nokta tahmini yapıldığında her ikisi arasında bir fark gözlenemez. Oysaki, n örneklem büyüklüğünün 100 olduğu durumda, daha güvenli tahmin yapıldığı belli olmalıdır. Aralık tahmininde örneklem büyüklüğünün artmasının aynı olasılıkla tahmini yapılacak parametrenin alabileceği değerler aralığını daraltacağı aşikardır.

Aralık tahmininde kullanılacak olan bazı genel tanımlar aşağıda verilmiştir:

Güven aralığı tahmin edicisi: Bilinmeyen parametreyi (ortalama, oran, varyans vb.) içine alacak iki değerin bulunabileceği belirtildi. Bu anlatım olasılık kullanılarak ifade edilebilir. θ herhangi bir parametre ve α, 0 ile 1 arasında bir olasılık değeri olmak üzere,

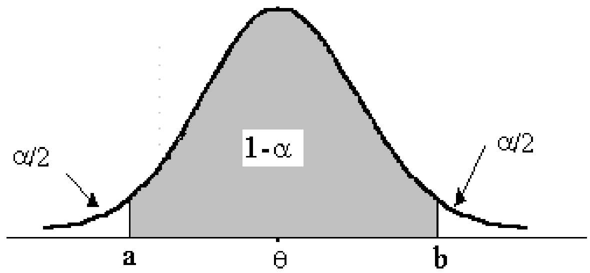
$$P(A \leq \theta \leq B) = 1 - \alpha \tag{7.5}$$

olasılığı gerçekleştiğinde, θ'yı (1- α) olasılıkla içine alacak A ile B sınırları ile verilen aralığa, θ'nın (1- α) **güven aralığı tahmin edicisi** denir.

Güven aralığı: Örneklem bilgilerine dayanarak, θ parametresi için,

$$P(A \leq \theta \leq B) = 1 - \alpha$$

eşitliğini sağlayacak A için a, B için b değerleri bulunmuş ise, a ve b aralığına θ'nın %(1-α) **güven aralığı** adı verilir. (1-α), aralığın **güven düzeyidir**. Güven düzeyinin anlamı şudur; Bir kitleden yerine iade ederek örneklemler çekildiğinde, bu örneklemlerden bulunacak güven aralıkların %(1-α)'sı θ parametresini içerir demektir. Yukarıda verilen tanımlardan, θ parametresini (1-α) olasılıkla içeren (a,b) aralığının belirlenmesi işlemine **aralık tahmini** denir. Burada α aralığın θ'yı içermeme olasılığıdır. θ'nın (a,b) aralığında olması olasılığı (1-α) ile ifade edilir. a'ya güven aralığının alt sınırı, b'ye güven aralığının üst sınırı denir. Normal dağılım için güven aralığı gösterimi Şekil 7.1'de verilmiştir.



Şekil 7.1 θ'nın Güven Aralığı Çizimi

Belli bir parametre için ne kadar dar bir güven aralığı sağlanırsa, parametre tahminin o ölçüde iyi olduğu söylenebilir.

Nokta Tahmini ile Aralık Tahmini Arasındaki Temel Farklılıklar:

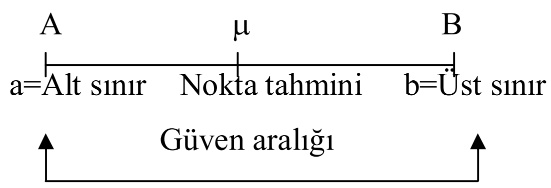
1- Nokta tahmininden tek bir nokta veya değer elde edilmesine karşın, aralık tahmini bir aralık vermektedir.

2- X raslantı değişkeninin normal ve t dağılımına sahip olması durumunda nokta tahmin değeri, aralık tahmininin orta noktasına denk gelmektedir.

3- Aralık tahmininde örneklem büyüklüğü etkili olurken, nokta tahmininde etkili değildir.

Kitle Ortalamasının Aralık Tahmini

Kitle ortalaması μ'nün içinde bulunduğu sınırlara "kitle ortalamasının güven aralığı" ya da "kitle ortalamasının güven sınırları" denir. Şekil 7.2'de gösterildiği gibi kitle ortalaması μ'nün, alt ve üst sınırların ortasında yer aldığı kabul edilir.



Şekil 7.2 Kitle Ortalamasının Aralık Tahmini

Kitle ortalamasının aralık tahmini, kitle varyansının bilinip bilinmemesi durumuna göre iki farklı yolla elde edilir.

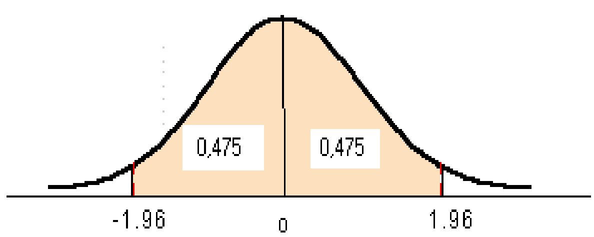
Kitle Varyansı σ^2 Bilindiğinde Kitle Ortalamasının Güven Aralığı

X, ortalaması μ ve varyansı σ^2 olan normal dağılıma sahip olduğunda, örneklem ortalaması $\overline{X}$'nin dağılımının da $\overline{X} \sim N(\mu, \sigma^2 / n)$ olduğu Bölüm 6'da belirtilmiştir. Buna göre güven aralığında kullanılacak olan istatistiğin dağılımı,

$$Z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}}$$

olan standart normal dağılımdır.

Standart normal dağılımdan çekilmiş bir örneklemde Z'nin $(1-\alpha)$=0,95 olasılıkla arasında bulunacak sınırlarını, P(?< Z <?) = 0,95 olarak yazabiliriz. Burada 0,95 değerini veren Z değeri nedir sorusunu bulalım. Ek 3'de verilen standart normal dağılım tablosu 0 dan, o Z değerine kadar olan toplam alanı verir. O halde tablonun içinden 0,4750 bulunursa buna karşılık gelen Z değeri 1,96'dır. Dağılımın simetrikliğinden dolayı sol taraf da –1,96 olacaktır. Yani; $P(-1{,}96 < Z < 1{,}96) = 0{,}95$'dır. Z dağılımına ilişkin grafik Şekil 7.3'de verilmiştir.



Şekil 7.3 0,95 Güven Düzeyi İçin Z'nin Dağılımı

$\overline{X}$'nın çekildiği kitle ortalaması μ için bir çıkarsama yapmaya çalışalım. σ bilindiğine göre, $P(-1{,}96 < Z < 1{,}96) = 0{,}95$ eşitliğinde Z yerine, $\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ eşitliğini yerleştirelim.

Bu durumda, $P(-1{,}96 < \dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} < 1{,}96) = 0{,}95$ olur. Eşitsizliği $\dfrac{\sigma}{\sqrt{n}}$ ile çarparsak,

$$P(-1{,}96\frac{\sigma}{\sqrt{n}} < \overline{X}-\mu < 1{,}96\frac{\sigma}{\sqrt{n}}) = 0{,}95$$

olur. Yukarıda verilen eşitsizlikten $\overline{X}$ çıkarılırsa;

$$P(-\overline{X}-1{,}96\frac{\sigma}{\sqrt{n}} < -\mu < -\overline{X}+1{,}96\frac{\sigma}{\sqrt{n}}) = 0{,}95$$

olur. Bu eşitsizliğin tüm tarafları (-1) ile çarpılırsa, eşitsizlik yön değiştirir ve buradan

$$P(\overline{X}-1{,}96\frac{\sigma}{\sqrt{n}} < \mu < \overline{X}+1{,}96\frac{\sigma}{\sqrt{n}}) = 0{,}95$$

elde edilir. Kitle ortalamasının güven sınırlarını verecek formülü, daha önce verilen tanımlara göre genellemek istersek;

$$P\left(\overline{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha \qquad (7.6)$$

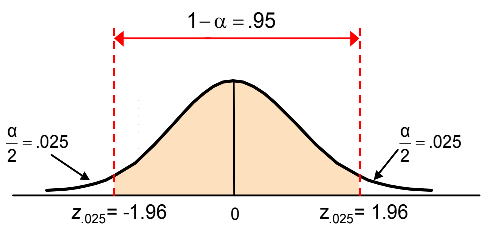
elde edilir. Eşitlik (7.6), kitle varyansı (σ^2) bilindiğinde, kitle ortalamasının güven aralığı için genel formüldür. Aralık genişliği de iki sınır arasındaki farktır. Güven aralığının genişliği aşağıdaki eşitlikten yararlanılarak da bulunabilir.

$$w = \overline{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} - \left(\overline{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 2Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \qquad (7.7)$$

w küçük ise tahmin de anlamlı ve yararlıdır.

Örnek 7.3: Kitle standart sapması 9 olan normal dağılımdan alınmış 25 büyüklüğünde bir rasgele örneklemin ortalaması 36 bulunmuştur. Kitle ortalamasının 0,95 güven düzeyinde güven aralığını ve aralık genişliğini bulunuz.

$\overline{X} = 36 \quad n = 25 \quad \sigma = 9$



$$P\left(\overline{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

$$P\left(36 - 1{,}96\frac{9}{\sqrt{25}} < \mu < 36 + 1{,}96\frac{9}{\sqrt{25}}\right) = 0{,}95$$

$P(32{,}472 < \mu < 39{,}528) = 0{,}95$ aralık genişliği =39,528-32,472= 7,056

Bu örnek üzerinde, standart sapmanın küçülmesi, örneklem büyüklüğünün artması ile güven aralığı ve aralık genişliğinin nasıl değişeceğini gösterelim.

Eğer kitle standart sapması $\sigma = 6$ olsaydı, güven aralığı ve aralık genişlikleri sırasıyla,

$P(33{,}648 < \mu < 38{,}352) = 0{,}95$ ve aralık genişliği = 4,704

olarak bulunurdu. Eğer örneklem büyüklüğü n = 50, kitle standart sapması $\sigma = 9$ olsaydı güven aralığı ve aralık genişliği sırasıyla;

$P(33{,}505 < \mu < 38{,}495) = 0{,}95$ aralık genişliği = 4,99

olarak elde edilirdi. Yukarıda verilen son iki uygulamada görüldüğü gibi, standart sapmanın küçülmesi veya örneklem büyüklüğünün artması aralık genişliğinin küçülmesine neden olmuştur.

Örnek 7.4: Dış ticaret sektöründe çalışan uzmanların gelirlerinin ortalaması kestirilmek isteniyor. Ankara'daki uzmanlar arasından rasgele olarak 50 uzman seçiliyor. 50 uzmanın ortalama aylıklarının 1100 YTL olduğu görülmektedir. Daha önce yapılan çalışmalardan sektördeki uzmanların aylık gelirlerinin varyansının 1600 YTL^2 olduğu bilindiğine göre, dış ticaret uzmanlarının ortalama gelirlerinin 0,95 güven düzeyinde güven aralığını bulunuz.

$n = 50,\quad \overline{X} = 1100,\quad \sigma^2 = 1600,\quad 1-\alpha = 0{,}95,\quad \alpha/2 = 0{,}025,\quad Z_{\alpha/2} = 1{,}96$

$$P\left(\overline{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

$$P\left(1100 - 1{,}96\frac{40}{\sqrt{50}} < \mu < 1100 + 1{,}96\frac{40}{\sqrt{50}}\right) = 0{,}95$$

$P(1088{,}91 < \mu < 1111{,}08) = 0{,}95$ aralık genişliği = 22,17

Yorum: Buna göre dış ticaret sektöründeki uzmanların aylık ortalama gelirlerinin 0,95 güven düzeyinde 1088,91 YTL ile 1111,08 YTL arasında olduğu söylenebilir.

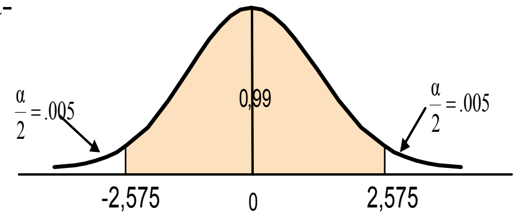
Örnek 7.5: Bir fabrikada üretilen ürünlerin uzunlukları kontrol edilmiş ve kitle standart sapması 0,45 mm olarak bulunmuştur. Bir kalite kontrol uzmanı rasgele seçilen 40 ürün üzerinde kontrol yapmış ve ortalama uzunluğu 35,62 mm olarak bulmuştur. Buna göre,

a- 0,99, 0,95 ve 0,90 güven düzeylerinde kitle ortalamasının güven aralıkları nedir?

b- 0,90 güven düzeyinde kitle ortalaması hangi değerden küçüktür?

$$\overline{X} = 35{,}62 \quad \sigma = 0{,}45 \quad n = 40$$

a-



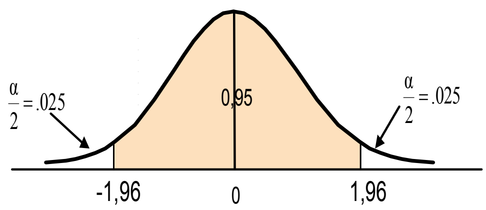
1-α=0,99 ve $Z_{\alpha/2}$=2,575

$$P\left(35{,}62-2{,}575\frac{0{,}45}{\sqrt{40}}<\mu<35{,}62+2{,}575\frac{0{,}45}{\sqrt{40}}\right)=0{,}99$$

$$P(35{,}62-0{,}18<\mu<35{,}62+0{,}18)=0{,}99$$

$$P(35{,}44<\mu<35{,}80)=0{,}99$$

aralık genişliği = 0,36

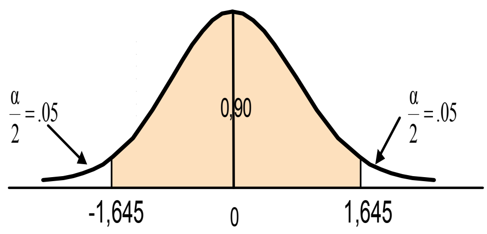


1-α=0,95 ve $Z_{\alpha/2}$=1,96

$$P(35{,}62-0{,}372<\mu<35{,}62+0{,}1372)=0{,}95$$

$$P(35{,}48<\mu<35{,}76)=0{,}95$$

aralık genişliği = 0,28



1-α=0,90 ve $Z_{\alpha/2}$=1,645

$$P\left(35{,}62-1{,}645\frac{0{,}45}{\sqrt{40}}<\mu<35{,}62+1{,}645\frac{0.45}{\sqrt{40}}\right)=0{,}90$$

$$P(35{,}50<\mu<35{,}74)=0{,}90$$

aralık genişliği = 0,24

b- $P\left(\mu<\overline{X}+Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)=0{,}90 \Rightarrow P\left(\mu<35{,}62+1{,}645\frac{0{,}45}{\sqrt{40}}\right)=0{,}90 \Rightarrow P(\mu<35{,}74)=0{,}90$

Yorum: 0,90 güven düzeyinde kitle ortalamasının üst sınırı 35,74'dür.

Yukarıda verilen örneklerden aşağıda verilen sonuçlar çıkarılabilir:

1- Kitle standart sapması büyüdükçe, kitle ortalamasına ait güven aralığı da genişler.

2- Örneklem büyüklüğü arttıkça, kitle ortalamasının güven aralığı daralır.

3- Güven düzeyi yükseldikçe, güven aralığının genişliği artar.

Kitle Varyansı σ^2 Bilinmediğinde Büyük Örneklemlerde Kitle Ortalaması Güven Aralığı

Bilindiği gibi örneklem büyüklüğü arttıkça örneklem standart sapması, kitle standart sapmasının iyi bir tahmin edicisidir. Böylece kitle varyansı bilinmediğinde, eğer örneklem büyüklüğü yeterince büyükse (n>30) örneklem standart sapması kullanılarak yapılan standartlaştırmada, $\frac{\overline{X}-\mu}{S/\sqrt{n}}$ olan değerler standart normal dağılıma yakınsar. Bu yaklaşım sonucunda, kitle ortalamasının güven aralığı,

$$P\left(\overline{X} - Z_{\alpha/2}\frac{S}{\sqrt{n}} < \mu < \overline{X} + Z_{\alpha/2}\frac{S}{\sqrt{n}}\right) = 1-\alpha \qquad (7.8)$$

olarak elde edilir.

Örnek 7.6: Bir sigorta şirketindeki hayat sigortası poliçeleri arasından 200 tanesi rasgele seçilerek, aylık ödenen prim miktarı incelenmiştir. Seçilen bu aylık prim ödemelerinin ortalaması 82,4 YTL, standart sapması 4,2 YTL olduğuna göre şirketteki hayat sigortası poliçeleri için ortalama aylık ödenen primlerin 0,95 güven düzeyinde güven aralığını bulunuz.

$$\overline{X} = 82{,}4 \quad S = 4{,}2 \quad n = 200 \quad 1 - \alpha = 0{,}95$$

$$P\left(82{,}4 - 1{,}96\frac{4{,}2}{\sqrt{200}} < \mu < 82{,}4 + 1{,}96\frac{4{,}2}{\sqrt{200}}\right) = 0{,}95 \Rightarrow P(81{,}82 < \mu < 82{,}98) = 0{,}95$$

Yorum: Ödenen primlerin ortalaması 0,95 güven düzeyi ile 81,82 YTL ile 82,98 YTL arasındadır.

Kitle Varyansı σ^2 Bilinmediğinde Küçük Örneklemlerde Kitle Ortalaması Güven Aralığı

Kitle varyansının bilinmediği durumda, örneklem büyüklüğü küçük (n<30) ve kitlenin dağılımı normal ise (μ ve σ^2 bilinmiyor), $\dfrac{\overline{X}-\mu}{S/\sqrt{n}}$ dönüşümün t dağılımı gösterdiği bilinmektedir. Bu durumda kitle ortalaması için verilen güven aralığı formülü;

$$P\left(\overline{X}-t_{\alpha/2;\,n-1}\frac{S}{\sqrt{n}}<\mu<\overline{X}+t_{\alpha/2;\,n-1}\frac{S}{\sqrt{n}}\right)=1-\alpha \qquad (7.9)$$

olur. Eşitlik (7.9)'da bulunan $t_{\alpha/2;\,n-1}$ tablo değeri Ek 5'de verilen t tablosu kullanılarak bulunur. t dağılım tablosunda (1-α) güven düzeyi için olasılık değerleri verilmiştir. Uygun serbestlik derecesi olan (n-1) kullanılarak (1-α) güven düzeyindeki değere bakılarak güven aralığı formülünde kullanılacak $t_{\alpha/2;\,n-1}$ değeri okunur.

Örnek 7.7: Araba lastiği ihracatı yapan bir firma, lastiklerin ortalama ömrünü tahmin etmek istiyor. Bu amaçla rasgele 21 lastik seçiyor. Seçilen 21 lastiğin parçalanmadan kat ettiği yolun ortalaması 65000 km ve standart sapması 9500 km olarak bulunuyor. Lastik ömrünün kitle dağılımının normal dağılıma sahip olduğu bilindiğine göre, bu firmanın ürettiği lastiklerin ortalama ömrü için 0,99 güven düzeyinde, güven aralığı bulunuz.

$$P\left(\overline{X}-t_{\alpha/2;\,n-1}\frac{S}{\sqrt{n}}<\mu<\overline{X}+t_{\alpha/2;\,n-1}\frac{S}{\sqrt{n}}\right)=1-\alpha$$

$t_{\alpha/2;\,n-1}$ için, Ek 5 ile verilen t tablosundan, 20 sd'li 0,99 güven düzeyi değerine bakılır ki bu değer 2,845'dir.

$$P\left(65000-2{,}845\frac{9500}{\sqrt{21}}<\mu<65000+2{,}845\frac{9500}{\sqrt{21}}\right)=0{,}99$$

$$P(59102{,}12<\mu<70897{,}8)=0{,}99$$

Yorum: 0,99 güven düzeyinde, araba lastiklerinin ortalama ömrü için güven aralığı 59102,12 km ile 70897,8 km arasındadır.

Kitle Oranının Güven Aralığı

Kitleyi belirleyen parametrelerden biri de özel bir durum için bu durumun kitlede olma oranıdır. Örneğin bir seçim bölgesinde belli bir partiye oy verenlerin oranı, belli bir hastalığın kitlede rastlanma oranı gibi bilgilerin kitledeki değeri, kitleye ait oran olarak alınabilir. Böylece bu kitleye ait oranın (binom parametresi) tahmininde $\hat{p} = \frac{X}{n}$ kullanılır. Burada, X ilgilenilen olayın örneklemde bulunma sayısı, n örneklemin büyüklüğüdür. Merkezi limit teoremi ile,

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \sim N(0,1)$$

dönüşümü yazılabilir. Uygulamada genellikle p bilinmez. Ancak n büyük olduğu için ;

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} \cong \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

yaklaşımı kullanılır.

Böylece kitle oranına ait güven aralığı,

$$P\left(\hat{p} - Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < p < \hat{p} + Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right) = 1 - \alpha \qquad (7.10)$$

ile bulunur.

Örnek 7.8: Bir bölgede yaşayan 18 yaşın üzerinde rasgele seçilen 100 kişiden 50'si A partisini desteklemektedir. Söz konusu bölgede A partisini destekleyen kişilerin oranı için, 0,90 ve 0,95 güven düzeylerinde güven aralığını bulunuz.

$$\hat{p} = \frac{50}{100} = 0,5$$

$1 - \alpha = 0,90 \quad Z_{\alpha/2}=1,645$

$$P\left(0,5 - 1,645\sqrt{\frac{0,5(1-0,5)}{100}} < p < 0,5 + 1,645\sqrt{\frac{0,5(1-0,5)}{100}} \right) = 0,90$$

$$P(0,418 < p < 0,582) = 0,90$$

Yorum: A partisini destekleyenlerin oranı 0,90 güven düzeyinde 0,418 ile 0,582 arasında bir değer alabilir.

$1 - \alpha = 0,95 \quad Z_{\alpha/2}=1,96$

$$P\left(0,5 - 1,96\sqrt{\frac{0,5(1-0,5)}{100}} < p < 0,5 + 1,96\sqrt{\frac{0,5(1-0,5)}{100}} \right) = 0,95$$

$$P(0,402 < p < 0,598) = 0,95$$

Yorum: A partisini destekleyenlerin oranı 0,95 güven düzeyinde 0,402 ile 0,598 arasında bir değer alabilir.

Örnek 7.9: Bir bölgede yapılan sağlık taraması sonuçlarına göre 500 kişide gözlenen bazı hastalık oranları aşağıda verilmiştir:

Hastalıklar	**Oranlar** $\hat{p}$
Sarılık	0,20
Kronik Bronşit	0,35

Bu bölge için sarılık hastalığı oranının 0,95 güven düzeyinde güven aralığını bulunuz.

$$P\left(0,20 - 1,96\sqrt{\frac{0,20(1-0,2)}{500}} < p < 0,20 + 1,96\sqrt{\frac{0,20(1-0,2)}{500}}\right) = 0,95$$

$$P(0,165 < p < 0,235) = 0,95$$

Yorum: Bu bölgede sarılık hastalığı oranının 0,95 güven düzeyinde 0,165 ile 0,235 arasında olabileceği söylenebilir.

Kitle Varyansının Güven Aralığı

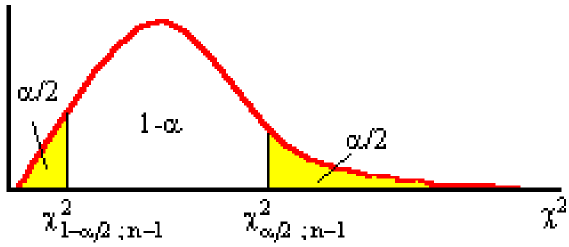
Kitle varyansının içinde bulunduğu sınırların tahmini için örneklem bilgilerinden yararlanılarak güven aralığı bulunur. χ^2 dağılımı, σ^2'nin güven aralığının elde edilmesinde kullanılır. Daha önce belirtildiği gibi, varyansı σ^2 olan normal dağılımlı bir kitleden n gözlemli rasgele bir örneklem çekildiğinde, örneklem varyansı S^2 ise, $(n-1)S^2/\sigma^2$ dönüşümü n-1 sd'li χ^2 dağılımına yakınsar. Yani,

$$\chi^2_{n-1} = \frac{(n-1)S^2}{\sigma^2}$$

dir. Buradan σ^2,

$$\sigma^2 = \frac{(n-1)S^2}{\chi^2_{n-1}}$$

yazılabilir. χ^2 dağılımında σ^2'yi (1-α) güven düzeyinde içine alacak iki χ^2 değeri Şekil 7.4'de verilmiştir.



Şekil 7.4 χ^2 Dağılımı

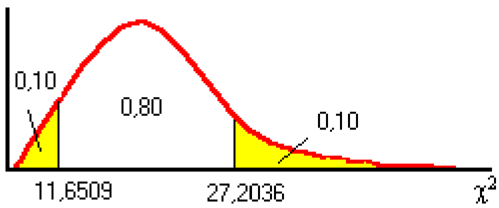
P(?<σ^2<?)=1-α biçiminde verilen χ^2 değerleri ve yazılan olasılık birlikte ele alınarak;

$$P\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2;n-1}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\alpha/2;n-1}}\right) = 1-\alpha \qquad (7.11)$$

elde edilebilir. Eşitlik (7.11) kitle varyansının güven aralığı formülüdür.

Örnek 7.10: Bir sosyolog bisiklet hırsızları üzerinde bir araştırma yapmıştır. 20 çocuk üzerinde yapılan araştırmada çocukların yaşları ortalaması 16,9 ve standart sapması 3,2 olarak bulunmuştur. 0,80 güven düzeyinde kitle varyansının güven aralığını bulunuz (Çocukların yaşlarının dağılımının normal olduğu varsayılmıştır).

S=3,2 n=20 1-α=0,80 $\chi^2_{0,10;19} = 27,2036$ $\chi^2_{0,90;19} = 11,6509$



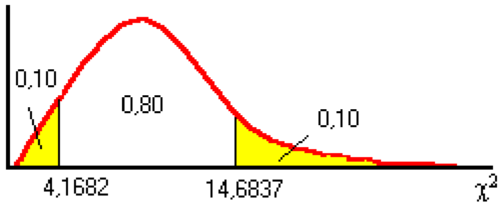
$$P\left(\frac{(20-1)3,2^2}{27,2036} < \sigma^2 < \frac{(20-1)3,2^2}{11,6509}\right) = 0,80$$

$$P\left(7,15 < \sigma^2 < 16,69\right) = 0,80$$

Yorum: 0,80 güven düzeyinde, bisiklet hırsızı çocukların yaşlarının kitle varyansının 7,15 ile 16,69 arasında olduğu söylenebilir.

Örnek 7.11: Bir makinede üretilen elektrik tellerinin gerilme dayanıklılığı test ediliyor. İncelenen 10 telin geriliş dayanıklılığı ölçülmüş ve varyansı 0,058 bulunmuştur. Kitle varyansının 0,80 güven düzeyi için güven aralığını bulunuz (Gerilme dayanıklılık ölçüsü X'in normal olduğu varsayılmıştır).

S^2=0,058 ; n=10 ; 1-α=0.80 ; $\chi^2_{0,10;9} = 14,6837$; $\chi^2_{0,90;9} = 4,1682$



$$P\left(\frac{(10-1)0,058}{14,6837} < \sigma^2 < \frac{(10-1)0,058}{4,1682}\right) = 0,80$$

$$P\left(0,035 < \sigma^2 < 0,125\right) = 0,80$$

Yorum: Üretilen elektrik tellerinin gerilme dayanıklılık ölçüsünün varyansı 0,80 güven düzeyinde 0,035 ile 0,125 arasında değişen değerler alabilir.

7.4. İSTENEN ÖRNEKLEM BÜYÜKLÜĞÜNÜN BELİRLENMESİ

Güven aralığı genişliği; güven düzeyine, kitle varyansına ve örneklem büyüklüğüne bağlıdır. Bu genişlik güven düzeyi ve varyans büyüdükçe artar, örneklem büyüklüğü arttıkça daralır. Amaç, tutarlı tahmin yapmaktır, bu nedenle aralık genişliğinin belli bir değer üzerine çıkması istenmez.

Örneklem büyüklüğü, aralık genişliğine bağlı olarak $Z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$ eşitliğinden bulunabilir. Bu eşitlik ile bulunacak bilgiye, kabul edilebilir hata denir. Kabul edilebilir hata ,e, araştırıcı tarafından seçilir ve

$$e = Z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$$

olarak verilir. Buradan

$$e^2 = Z^2_{\alpha/2}\frac{p(1-p)}{n}$$

olup örneklem büyüklüğü,

$$n \geq Z^2_{\alpha/2}\frac{p(1-p)}{e^2} \tag{7.12}$$

olarak elde edilir. p değeri ya önceki bilgilerden ya da pilot çalışması ile belirlenir. Bu konuda hiç bilgi yok ise, p=0,50 almak, daha büyük örneklem büyüklüğü vereceği için uygundur.

Örnek 7.12: Sağlık taramasının yapıldığı bir çalışmada güven düzeyi 0,95, kabul edilebilir hata 0,01 olarak alınırsa, örneklem büyüklüğü ne olmalıdır?

p ile ilgili hiçbir bilgi olmadığını kabul edip, p=0,50 alalım.

Güven düzeyi, 0,95 iken $Z_{\alpha/2}$=1,96 olur.

$$n = Z_{\alpha/2}^2 \frac{p(1-p)}{e^2}$$

$$n = 1{,}96^2 \frac{0{,}50(1-0{,}50)}{0{,}01^2} = \frac{0{,}9604}{0{,}0001} = 9604$$

Eğer önceki bilgilerden p=0,35 ve e= ±0.04 ise,

$$n \geq 1{,}96^2 \frac{0{,}35(1-0{,}35)}{0{,}04^2} = 546$$

olacaktır.

Not: σ bilindiğinde, kitle ortalaması tahmini için gerekli örneklem büyüklüğü aynı yolla aşağıda verildiği biçimde elde edilir:

$$n \geq \frac{Z_{\alpha/2}^2 \sigma^2}{e^2} \qquad (7.13)$$

Örnek 7.13: Bir sigorta şirketi kasko yaptıran müşterilerinin araçları ile ilgili bir araştırma yapmak istiyor. Önceki bilgilerden araçların yaşlarına ilişkin varyansı 5'dir. Kabul edilebilir hata 0,28 olarak kabul edildiğinde, 0,05 anlamlılık düzeyinde örneklem büyüklüğü ne olmalıdır?

$$n \geq \frac{Z_{\alpha/2}^2 \sigma^2}{e^2} = \frac{1{,}96^2 \times 5}{0{,}28^2} = \frac{19{,}208}{0{,}0784} = 245$$

PROBLEMLER

1- Bir şirketteki çalışanlardan 12'si rasgele seçilmiştir. Bu 12 çalışanın bir ay içinde yaptıkları fazla mesai saatleri:

22; 16; 28; 12; 18; 36; 23; 11; 41; 29; 26; 31

olarak bulunmuştur.

a- Kitle ortalamasını tahmin ediniz.
b- Kitle varyansını tahmin ediniz.
c- Bu işletmede incelemenin yapıldığı ayda 30 saatin üstünde mesai yapanların oranı nedir?

2- Ortalaması μ, varyansı σ^2 olan bir kitleden seçilen iki gözlemli rasgele örneklem $\{X_1, X_2\}$ olsun. μ'nün aşağıda verilen üç tahmin edicisi ele alınsın;

$$\overline{X}_a = \frac{1}{2}X_1 + \frac{1}{2}X_2 \qquad \overline{X}_b = \frac{1}{4}X_1 + \frac{3}{4}X_2 \qquad \overline{X}_c = \frac{1}{3}X_1 + \frac{2}{3}X_2$$

a- Bu üç tahmin edicinin yansız olduğunu gösteriniz.

b- Bu tahmin edicilerden hangisi daha etkindir?

c- $\overline{X}_a$'nın diğer iki tahmin ediciye göre göreli etkinliğini (GE) bulunuz.

3- Ortalaması μ, varyansı σ^2 olan bir kitleden seçilen örneklem, X_1, X_2, X_3, X_4 olsun. μ'nün aşağıda verilen tahmin edicileri ele alınsın.

$$\overline{X}_a = \frac{X_1 + 2X_2 + 3X_3 + 4X_4}{10} \qquad \overline{X}_b = \frac{X_1 + 4X_2 + 4X_3 + X_4}{10}$$

a- Her iki tahmin edicinin de yansız olduğunu gösteriniz.

b- Hangi tahmin edici daha etkindir?

c- Göreli etkinliği bulunuz.

4- θ_1'in yansız bir tahmin edicisi $\hat{\theta}_1$ ve θ_2'nin yansız bir tahmin edicisi $\hat{\theta}_2$ olsun.

a- $(\hat{\theta}_1 + \hat{\theta}_2)$'nin $(\theta_1+\theta_2)$'nin yansız bir tahmin edicisi olduğunu gösteriniz.

b- $(\hat{\theta}_1 - \hat{\theta}_2)$'nin $(\theta_1-\theta_2)$'nin yansız bir tahmin edicisi olduğunu gösteriniz.

5- 100 denekli bir örneklemden elde edilen ortalama 3 ve standart sapma 5 olduğuna göre kitle ortalamasının 0,95 güven düzeyinde güven aralığını elde ediniz.

6- Aşağıda verilen sorulara ilişkin güven aralıklarını bulunuz.

a- $\overline{X} = 3$; $S = 2$; $n = 9$ ise, kitle ortalaması için 0,99 güven düzeyinde güven aralığını bulunuz.

b- S=1,2 ; n=5 ise, kitle varyansı için 0,80 güven düzeyinde güven aralığını bulunuz.

7- Bir şirketin insan kaynakları bölümünde çalışan yöneticisi, şirketin orta düzey elemanlarına kalite eğitimi uygulamıştır. Geçmişteki bilgilerden bu eğitim sonucundan alınan memnuniyet puanlarının varyansı 2,1 olan bir normal dağılım göstermektedir. Son yıl uygulamasından 10 gözlemlik bir örneklem çekilmiştir. Örneklemde ortalama puan 6,3'dür.

a- Son yıl eğitim verilen adayların kitle ortalamasının 0,80, 0,90 ve 0,95 güven düzeylerinde güven aralıklarını bulunuz ve elde edilen sonuçları tartışınız.

b- Bir istatistikçi çekilen 10 gözlemlik örneklem bilgilerinden güven aralığı sınırlarını 5,3 ile 7,3 olarak bulmuştur. Kitle ortalamasının bu sınırlar arasında bulunma olasılığı nedir?

8- Yeni çıkan bir vergi yasasının yapılan işlemlerde ek bir zaman harcaması getirdiğini gören bir muhasebe bürosu, hesaplarını tuttuğu 800 defter içinden 30 tanesini rasgele seçip harcadığı ek zamanı belirlemiştir. Dakika olarak harcanılan zamana ait aşağıda verilen bilgiler elde edilmiştir:

$$\sum_{j=1}^{30} X_j = 1320 \quad \sum_{j=1}^{30} \left(X_j - \overline{X}\right)^2 = 3210$$

a- Gerekli gördüğünüz bir varsayım var ise onu belirterek yeni yasanın getirdiği ek zamanın ortalamasının 0,95 güven düzeyinde güven aralığını bulunuz.

b- Örneklem büyüklüğünü 15 olarak aldığınız bir örneklemin kitle ortalaması için bulacağınız 0,95 güven düzeyinde güven aralığının daha dar, daha geniş ya da eşit olacağını hiç hesaplama yapmadan tartışınız.

9- Yıllık üretimi 25 ton olan bir sucuk fabrikası ortalama 300 gr'lık kangallar ile satış yapmaktadır. Çeşitli nedenlerle bu ağırlık tam 300 gr olamamaktadır. Rasgele seçilen 50 kangal sucuk ayrı ayrı tartılmış ağırlıklarının ortalaması 285 gr ve standart sapması 0,25 gr olarak bulunmuştur. Bu üretilen sucukların kangal ortalama ağırlığı için 0,95 güven düzeyinde güven aralığını bulunuz.

10- Toplam kalite yönetiminde müşterinin fikri çok önemsenir. Bir şirket yöneticilerinin bu konudaki fikirlerini almak için rasgele seçtiği 350 yöneticisine bir anket vermiştir. Bu ankette sorulardan biri "Müşteriye sunulan kusurlu malların azaltılması maliyeti yükseltir" şeklindedir. Seçenekler 1'den 5'e kadar puanlandırılmıştır. Bu fikre kesinlikle katılıyorum 1..., hiç katılmıyorum 5 puan olarak değerlendirilmiştir. Cevapların örneklem ortalaması 2,75 ve örneklem standart sapması 1,25 olan normal dağılıma sahip olduğu bilinmektedir.

a- Bu bilgilere dayanarak bu soru için yöneticilerin aldığı ortalama puanların 0,90 güven düzeyinde güven aralığını bulunuz.

b- Kitle ortalaması alt sınırı 2,55 olan güven aralığına ait güven düzeyi ne kadardır?

11- Bir kırtasiye mağazası, üniversite öğrencilerinin öğretim yılının ilk ayı içindeki kırtasiye harcamaları ile ilgilenmektedir. Bu amaçla 20 kişilik rasgele bir örneklem çekip ortalama harcamanın 40,23 YTL ve standart sapmanın 1,2 YTL olduğunu bulmuştur. Ortalama harcamanın 0,95 güven düzeyinde güven aralığını bulunuz. Yorumlayınız.

12- Bir benzin dağıtım şirketi, yeni dağıtım tankerleri almak amacı ile bir araştırma yapmıştır. Bir bölgede bulunan 30 istasyondan rasgele 10 tanesini seçip, bir yıl içinde yine rasgele seçilen 150 güne ait satış miktarlarını ton olarak kaydetmiştir. Satış istasyonlarının günlük tüketiminin birbirinden bağımsız normal dağıldığı varsayılmaktadır. Satış miktarlarına ilişkin bilgiler aşağıdadır:

	Ortalama	**Standart Sapma**	**Gün**
1-	2,3 ton	0,9 ton	150
2-	5,1 ton	1,7 ton	150
3-	8 ton	2 ton	150
4-	1,2 ton	1 ton	150
5-	3,8 ton	1,2 ton	150
6-	4,1 ton	2 ton	150
7-	6,5 ton	1,2 ton	150
8-	6 ton	2 ton	150
9-	9 ton	3,2 ton	150
10-	5 ton	1,5 ton	150

a- Her benzin istasyonu için kitle ortalamasının 0,95 ve 0,90 güven düzeylerinde güven aralığını bulunuz.

b- Yaptığınız aralık tahminlerini kullanarak istasyonları karşılaştırınız.

13- Rasgele seçilen 7 hastanın iyileşme sürelerine ait ortalama 24 gün ve standart sapma 4 gün olarak bulunmuştur. 0,95 güven düzeyinde ortalama iyileşme sürelerine ait güven aralığını bulunuz.

14- Bir banka şubesinin bir aylık işlem sayısı 850'dir. Yapılan bir aylık bu işlemlerden 60 işlem rasgele seçilmiştir. Bu işlemlerin dağılımı ;

Kambiyo	;	10
Havale	;	18
Para çekme	;	25
Para yatırma	;	7

dır. Bu işlemlerin oranları için birer nokta tahmini veriniz.

15- Bir ay içinde hastane polikliniklerine başvuran hastalardan rasgele seçilen 900 hastanın 360'ı kadın, 225'i çocuk ve geri kalanı erkektir. Erkek hastaların kitle oranının 0,90 güven düzeyinde güven aralığını bulunuz.

16- Bir kütüphaneden alınan 400 kitabın konulara göre dağılımı şöyledir. Fen 100, Kimya 50, Sosyal Bilim 120, Mühendislik 130. Sosyal bilimlere ait kitapların kitle oranı için 0,95 güven düzeyinde güven aralığını bulunuz.

17- Otomatik çalışan makine ile doldurulan şişelerin ağırlıklarının dağılımı normal dağılım göstermektedir. 8 makine üzerine yapılan çalışma sonucunda örneklem standart sapması 22 gr olarak bulunuyor. 0,80 güven düzeyinde kitle varyansı için güven aralığını bulunuz.

18- Otomatik çalışan bir makine ile doldurulan şişelerin ağırlıklarının dağılımı ortalaması 500, varyansı 100 olan normal dağılım göstermektedir. Rasgele 4 şişe seçilmiştir. Örneklem varyansının 78,87’den küçük olma olasılığını bulunuz.

19- Bir borsa analisti müşterileri arasında yapacağı araştırma için örneklem büyüklüğü belirlemek istiyor. Kabul edilebilir hatanın 0,03 olarak alınması uygun bulunuyor. Anlamlılık düzeyi 0,05 olarak alınır ve p’nin bilinmediği varsayılırsa, örneklem büyüklüğü ne olmalıdır?

20- Kabul edilebilir hata 0,50 , anlamlılık düzeyi 0,05 ve kitle varyansı 200 olan bir çalışma için örneklem büyüklüğü ne olmalıdır?

8. SEKİZİNCİ BÖLÜM

8.1. GİRİŞ

8.2. ORTALAMAYA İLİŞKİN HİPOTEZ TESTLERİ

8.3. KİTLE ORANINA AİT HİPOTEZ TESTLERİ

8.4. KİTLE VARYANSINA AİT HİPOTEZ TESTLERİ

PROBLEMLER

8.1. GİRİŞ

İstatistik biliminin amacı, kitle bilgilerine ulaşmak, kitleyi tanımaktır. Kitleyi tanıma ile ilgili bilgiler şu ana kadar nokta tahmini ve aralık tahmini yapmak şeklinde verilmiştir. Bazı durumlarda, kitle hakkında bir ön bilgimizin, doğru olup olmadığını da sınamak isteyebiliriz. Örneğin, yeni vergi yasasının orta ve küçük işletmeler ile büyük işletmelerde kabul edilme oranı ve bu iki grup arasında farkın olup olmayacağını; yeni bir pazarlama stratejisinin önce uygulanandan daha başarılı olup olmadığını; bir parça üretimde hatalı oranın beklenen bir değeri geçip geçmediğini; paket halinde sunulan deterjanların belirlenen 3 kg'lık ortalama ağırlığı sağlayıp sağlamadığını bilmek isteyebiliriz. Bu örneklerde ortak bilgilerin örneklemden elde edilmiş olmasıdır. Ancak bu sonuçlar kitle için de geçerli olacak mıdır? Bunu sınamak istemekteyiz. Konuyu daha genel incelemek için, kitleye ait bir θ parametresi ile ilgilenildiği düşünülsün. Eğer kitle için bilgimiz θ ise ve bunun tersi olan bilginin varlığı kanıtlanmamış ise, θ'nın geçerli olduğu varsayılır. Bu görüş, ortaya atılan hipotezdir. Ortaya atılan hipotez sıfır hipotezidir. Yukarıda verilen örneklerde, yeni vergi yasasının küçük ve orta ölçekli sanayiciler ile büyük sanayiciler tarafından aynı oranda kabul edildiği gibi, paket halinde sunulan deterjanların ortalama ağırlığının 3 kg olduğu ile ilgili görüş ortaya atılan hipotezlerdir. Bu hipotezler ilgili kitlelerden çekilen örneklemlerle sınanır. Hipotezimiz ya reddedilir ya da reddedilemez. İstatistikte önerilen ilk görüş sıfır hipotezi, karşı görüş de karşıt hipotez olarak adlandırılır. Yukarıda verilen paket ağırlıklarına ilişkin örnekte karşıt hipotez, ağırlığın 3 kg'dan daha hafif, daha ağır ya da 3 kg'dan farklıdır şeklinde olacaktır. Hipotez test edilirken, bu karşıt hipotezlerden biri kullanılır.

İstatistikte sıfır hipotezi H_0, karşıt hipotez H_1 ya da H_A ile gösterilir. Bu kitap boyunca karşıt hipotez H_A ile gösterilecektir. Kitle için kabul edilen hipotez kitlenin bir parametresi için bir değeri kullanıyor ise basit hipotez, bir aralık değeri veriyorsa bileşik hipotez olarak adlandırılır. Örneğin, deterjan paketleri 3 kg'dır hipotezi basit sıfır hipotezidir. Karşıt hipotez de basit ya da bileşik olabilir.

Kitle parametresi θ, θ_0 gibi bir değere sahip olsun, yukarıda değinilen basit ve bileşik hipotezler için örnekler;

H_0: $\theta=\theta_0$ ise H_0 hipotezi basittir. Bu hipotez için karşıt hipotez H_A: $\theta>\theta_0$ ya da H_A: $\theta<\theta_0$ şeklinde ise bu karşıt hipotezler bileşik hipotezlerdir. Eğer H_0: $\theta\leq\theta_0$; H_A: $\theta>\theta_0$ ise her iki hipotez de bileşiktir.

Deterjan paketleri ile ilgili örnekte, karşıt hipotezi, paketlerin 3 kg'dan hafif olduğunu iddia eden bir hipotez:

H_0:$\mu \geq 3$kg

H_A:$\mu<3$ kg

şeklinde olacaktır. Bu hipotezlerde her iki hipotez de bileşik birer hipotezdirler. Aynı zamanda da bu hipotez tek yanlı bir hipotezdir.

H_0:$\mu=3$ kg

H_A:$\mu\neq3$ kg

şeklindeki hipotezde ise H_0 basit, H_A bileşik hipotez olup, bu hipotez iki yanlı bir hipotezdir.

Örnek 8.1: Bir parça üretimde, hata oranının 0,01'den fazla olduğu iddia ediliyor. Üretimde hata oranının bu oranı geçtiği ile ilgili bir kanıt olmadıkça, sistem iyi çalışıyor denilmektedir. Hata oranı p ile gösterilir ise hipotez

H_0:$p\leq0,01$

H_A:$p>0,01$

olan H_0'ında H_A'nında bileşik hipotez olduğu tek yanlı hipotezdir. Bunun anlamı, hata oranı $p\leq0,01$'e karşı örneklem kullanılarak, $p>0,01$ karşıt hipotezi test edilecek demektir.

Yukarıda verilen bilgiler kullanılarak bir karar vermek, hipotez testlerinin ikinci aşamasıdır. Karar verme aşamasında sıfır hipotezi ya reddedilemez ya da karşıt hipotezin lehine sıfır hipotezi reddedilir. Hipotez testinde sıfır hipotezinin reddi ya da reddedilemeyeceği örneklem bilgisine dayanır. Kitle için kabul edilen değer ile örneklemden elde edilen değer arasında ne kadar çok fark var ise, sıfır hipotezinin reddi olasılığı da o kadar yüksektir. Ancak, örneklem kitleden rasgele çekilmiş olduğu için örneklemden elde edilen değerin kitle değeri ile aynı olması beklenemez. Güven aralıkları konusunda görüldüğü gibi aynı kitleden çekilen örneklemlerin kitle bilgisini içermeme olasılığı α kadar olacaktır. Bu durumda, sıfır hipotezi için karar sürecinde aşağıda verilen tabloda belirtilen sonuçlarla karşılaşılır.

Tablo 8.1 Karar Süreç Tablosu ve Olasılıklar

Karar	**H_0 Doğru**	**H_0Yanlış**
H_0 Red	Hatalı Karar I. Tür Hata α	Doğru Karar 1-β
H_0 Reddedilemez	Doğru Karar 1-α	Hatalı Karar II. Tür Hata β

Tablo 8.1'den görüldüğü gibi, sıfır hipotezi doğru ve bu hipotez reddedilmiş ise, yanlış bir karar verilmiştir. Bu hatalı karar I. tür hata adını alır. I. tür hata yapma olasılığı α olacaktır. α'ya hipotezin anlamlılık düzeyi denir. Karar ya red ya da reddedilemez olabileceğine göre, doğru bir hipotezin reddedilmeme olasılığı da 1-α olacaktır. Tablo 8.1'e göre, yapılacak ikinci hata, II. tür hata olarak adlandırılır. Belli bir karar kuralı için, sıfır hipotezi yanlış iken reddedilememe olasılığı β ile gösterilir. Yanlış bir hipotezi reddetme olasılığı da 1-β'dır. 1-β'ya testin gücü denir.

Hipotez testinde kullanacak terimler aşağıda özetlenmiştir.

Sıfır Hipotezi (H_0): Tersine yeterli kanıt bulununcaya kadar doğru kabul edilen savdır.

Karşıt Hipotez (H_A): Sıfır hipotezi karşısında test edilen, sıfır hipotezi reddedildiğinde kabul edilen hipotezdir.

Basit Hipotez: İlgilenilen kitlede testi yapılan parametre için bir tek değerin belirtilmesidir.

Bileşik Hipotez: İlgilenilen kitlede testi yapılan parametre için bir değerler aralığı belirleyen hipotezdir.

Tek Yanlı Karşıt Hipotez: Kitlede ilgilenilen parametre için sıfır hipotezince, belirlenen bir değerden küçük ya da büyük olanaklı tüm değerleri içeren karşıt hipotezdir.

İki Yanlı Karşıt Hipotez: (Biz çalışmalarımızda bu hipotezi basit sıfır (H_0) hipotez ile kullanacağız) Kitlede ilgilendiğimiz parametre için sıfır hipotezince belirlenen değer dışında olanaklı bütün değerleri içeren karşıt hipotezdir.

Hipotez Testi Kararı: Araştırmacıyı örneklem kanıtına dayanarak, sıfır hipotezini red ya da reddetmemeye götürecek şekilde geliştirilmiş karar kuralıdır.

I. Tür Hata: Doğru hipotezin reddedilmesi.

II. Tür Hata: Yanlış hipotezin reddedilmemesi.

Anlamlılık Düzeyi: Doğru olan sıfır hipotezini reddetme olasılığı, α.

Testin Gücü: Yanlış bir sıfır hipotezinin reddedilme olasılığı, 1-β.

8.2. ORTALAMAYA İLİŞKİN HİPOTEZ TESTLERİ

Hipotez testlerinde de güven aralığı konusunda verildiği gibi, kitle varyansının bilinip bilinmemesi, örneklem büyüklüğünün 30'dan büyük ya da küçük olmasına göre farklı dağılımlar söz konusu olacaktır. Buna ilişkin durumlar sırasıyla aşağıda verilmiştir.

Kitle varyansı σ^2'nin bilindiği durumda hipotez testi:

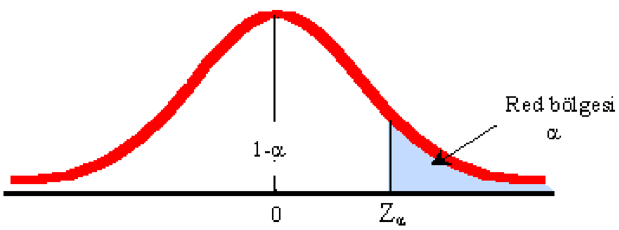
Bu durumda kullanılacak test istatistiği,

$$Z_H = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \tag{8.1}$$

ile verilen Z test istatistiğidir. Eğer hipotez,

$$\begin{array}{l} H_0 : \mu \le \mu_0 \\ H_A : \mu > \mu_0 \end{array} \quad \text{ya da} \quad \begin{array}{l} H_0 : \mu = \mu_0 \\ H_A : \mu > \mu_0 \end{array}$$

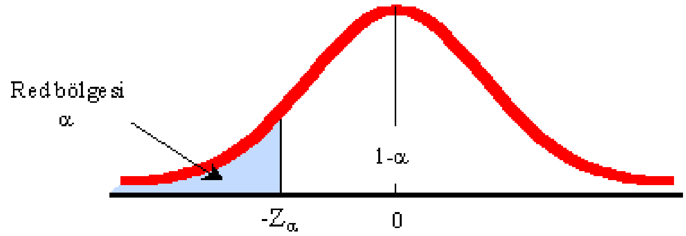
biçiminde ise, $Z_H > Z_\alpha$ olduğunda H_0 hipotezi reddedilir, $Z_H \le Z_\alpha$ olduğunda H_0 hipotezi reddedilemez.



Eğer hipotez,

$$\begin{array}{l} H_0 : \mu \ge \mu_0 \\ H_A : \mu < \mu_0 \end{array} \quad \text{ya da} \quad \begin{array}{l} H_0 : \mu = \mu_0 \\ H_A : \mu < \mu_0 \end{array}$$

ise, $Z_H < -Z_\alpha$ olduğunda H_0 hipotezi reddedilir, $Z_H \ge -Z_\alpha$ olduğunda H_0 hipotezi reddedilemez.

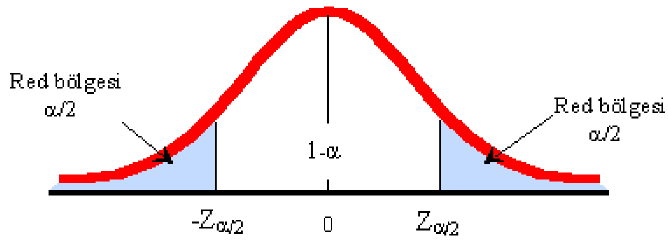


Eğer hipotez,

H_0: $\mu = \mu_0$

H_A: $\mu \neq \mu_0$

ise, $Z_H > Z_{\alpha/2}$ ya da $Z_H < -Z_{\alpha/2}$ olduğunda H_0 hipotezi reddedilir, diğer durumda reddedilemez.



H_0: $\mu = \mu_0$; H_A: $\mu \neq \mu_0$ hipotezi güven aralığından yararlanılarak da test edilebilir. Eğer H_0 hipotezinde verilen μ_0 ortalama değeri,

$$P\left(\overline{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

ile hesaplanan güven aralıklarının içine düşüyorsa, H_0 hipotezi reddedilemez.

p değeri: Bilgisayarlarda kullanılan istatistik paket programlarındaki hipotez testlerinin sonuçlarında genellikle bir olasılık değeri de verilir. Bu olasılık, H_0 hipotezi için hesaplanan test istatistiğinin H_A hipotezine göre sağ uçta veya sol uç ta veya her iki uçta birden bıraktığı olasılığı verir. Bu olasılığa p denir. Hesaplanan p değeri anlamlılık düzeyi α'dan küçük ya da eşit ise ($p \leq \alpha$), H_0 hipotezi reddedilir, $p > \alpha$ olduğunda H_0 hipotezi reddedilemez.

Bir testte p değerinin hesaplanması kolay olmayabilir. Fakat normal dağılım testlerinde oldukça kolaydır. Bölüm 5.3'de verilen F(Z) birikimli standart normal dağılımdan yararlanarak p değeri aşağıdaki eşitlik olarak tanımlanır:

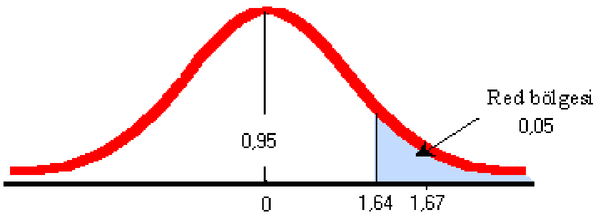
$$p=\begin{cases}2\left[1-F\left(\left|Z_H\right|\right)\right] & ; \quad H_A:\mu\neq\mu_0 \text{ ise}\\ 1-F(Z_H) & ; \quad H_A:\mu>\mu_0 \text{ ise}\\ F(Z_H) & ; \quad H_A:\mu<\mu_0 \text{ ise}\end{cases} \tag{8.2}$$

Örnek 8.2: Belli bir ilacın ortalama etki süresinin 90 dakikadan fazla olduğu iddia ediliyor. Rasgele seçilen hastalardan 16'sına ilgili ilaç veriliyor ve ortalama etki süresi 92,5 dakika olarak bulunuyor. Kitle varyansı önceden 36 dak^2 olarak bilindiğine göre, α=0,05 anlamlılık düzeyinde iddianın doğruluğunu test ediniz.

H_0: $\mu\leq 90$ dakika

H_A: $\mu>90$ dakika

$$Z_H=\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}=\frac{92{,}5-90}{6/\sqrt{16}}=1{,}67$$



$Z_{0,05}$=1,64'dür. 1,67 değeri red bölgesine düşmektedir. $Z_H>Z_\alpha$ olduğu için H_0 reddedilir.

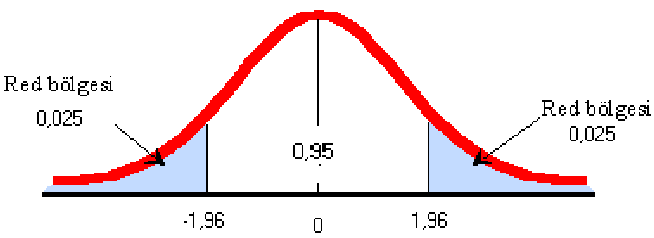
Yorum: 0,05 anlamlılık düzeyinde, ilacın ortalama etki süresinin 90 dakikadan fazla olduğu söylenebilir ($p<0{,}05$).

Örnek 8.2 de iddia; etki süresinin 90 dakikadan farklı olduğu söyleniyor ise hipotez,

H_0: $\mu=90$

H_A: $\mu\neq 90$

dir.



Z_H=1,67 pozitif değerdir ve pozitif bölgedeki $Z_{\alpha/2}$=1,96 değeri ile karşılaştırılır. $Z_H<Z_{\alpha/2}$ olduğunda H_0 hipotezi reddedilemez. Bu hipotez güven aralığından yararlanılarak da aşağıda verildiği biçimde çözülebilir:

$$P\left(\overline{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

$$P\left(92{,}5 - 1{,}96\frac{6}{\sqrt{16}} < \mu < 92{,}5 + 1{,}96\frac{6}{\sqrt{16}}\right) = 0{,}95$$

$$P(89{,}56 < \mu < 95{,}44) = 0{,}95$$

Hipotezin testinde verilen 90 değeri, güven aralıklarının içinde kaldığı için H_0 hipotezi reddedilemez.

Yorum: 0,05 anlamlılık düzeyinde kitle ortalamasının 90 olduğu söylenebilir.

Kitle varyansı σ^2 bilinmiyor ve $n \geq 30$ ise hipotez testi

Kullanılacak test istatistiği,

$$Z_H = \frac{\overline{X} - \mu_0}{S/\sqrt{n}} \tag{8.3}$$

ile verilen Z test istatistiğidir. Eğer hipotez,

$$\begin{array}{lcl} H_0: \mu \leq \mu_0 & \text{ya da} & H_0: \mu = \mu_0 \\ H_A: \mu > \mu_0 & & H_A: \mu > \mu_0 \end{array}$$

biçiminde ise, $Z_H > Z_\alpha$ olduğunda H_0 hipotezi reddedilir, $Z_H \leq Z_\alpha$ olduğunda H_0 hipotezi reddedilemez.

Problem için uygun hipotez,

$$\begin{array}{lcl} H_0: \mu \geq \mu_0 & \text{ya da} & H_0: \mu = \mu_0 \\ H_A: \mu < \mu_0 & & H_A: \mu < \mu_0 \end{array}$$

biçiminde ise, $Z_H < -Z_\alpha$ olduğunda H_0 hipotezi reddedilir, $Z_H \geq -Z_\alpha$ olduğunda H_0 hipotezi reddedilemez.

Eğer hipotez,

H_0: $\mu = \mu_0$

H_A: $\mu \neq \mu_0$

biçiminde ise, $Z_H > Z_{\alpha/2}$ ya da $Z_H < -Z_{\alpha/2}$ olduğunda H_0 hipotezi reddedilir, diğer durumda reddedilemez.

H_0: $\mu = \mu_0$, H_A: $\mu \neq \mu_0$ hipotezi güven aralığından yararlanılarak da, test edilebilir. Eğer H_0 hipotezinde verilen μ_0 ortalama değeri,

$$P\left(\overline{X} - Z_{\alpha/2}\frac{S}{\sqrt{n}} < \mu < \overline{X} + Z_{\alpha/2}\frac{S}{\sqrt{n}}\right) = 1-\alpha$$

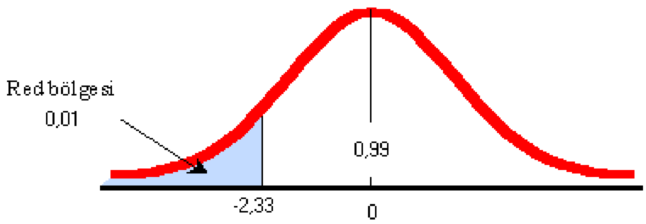
ile hesaplanan güven aralıklarının içine düşüyorsa, H_0 hipotezi reddedilemez.

Örnek 8.3: Bir yazılım şirketi yeni bir muhasebe paketi oluşturduğunu ve eski paketle 16 saniyede gerçekleşen işlemlerde daha kısa sürede sonuç alınacağı iddia ediyor. Bu iddianın doğrulanması için 36 işlem uygulanıyor ve ortalama işlem süresi 14 saniye, standart sapma 2 saniye olarak bulunuyor. Şirketin iddiası 0,01 anlamlılık düzeyinde doğru mudur?

H_0: $\mu \geq 16$

H_A: $\mu < 16$

$$Z_H = \frac{\overline{X} - \mu_0}{S/\sqrt{n}} = \frac{14-16}{2/\sqrt{36}} = -6$$



–6<-2,33 olduğundan H_0 hipotezi reddedilir.

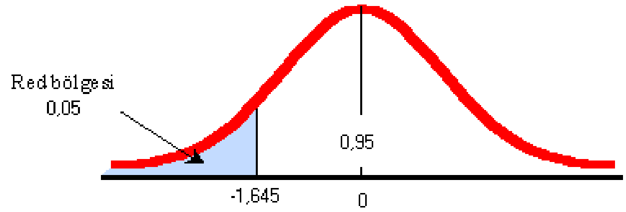
Yorum: 0,01 anlamlılık düzeyinde yazılım daha kısa sürede sonuç vermektedir.

Örnek 8.4: Kardeşler süpermarket saat 5'den sonra alışveriş eden müşterilerinin ortalama 50 YTL'den az harcama yaptığını iddia ediyor. Bu iddiayı doğrulamak için, saat 5'den sonra alış-veriş yapan 60 müşterisini rasgele seçiyor. Yapılan harcamaların ortalaması 45 YTL ve standart sapması 15 YTL olarak bulunduğuna göre, 0,05 anlamlılık düzeyinde hipotezi test ediniz.

H_0: $\mu \geq 50$

H_A: $\mu < 50$

$$Z_H = \frac{\overline{X} - \mu_0}{S/\sqrt{n}} = \frac{45-50}{15/\sqrt{60}} = -2{,}58$$



$-2,58<-1,645$ olduğundan H_0 hipotezi reddedilir.

Yorum: 0,05 anlamlılık düzeyinde, market müşterileri saat 5'den sonra 50 YTL'den daha az tutarda alışveriş yapmaktadırlar.

Kitle varyansı σ^2'nin bilinmediği ve n<30 durumunda hipotez testi:

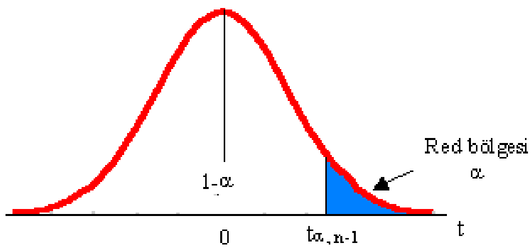
Kitle dağılımı normal varsayımı altında kullanılacak test istatistiği, Bölüm 7'de güven aralığında verildiği gibi t test istatistiği olup, Eşitlik (8.4)'de verilmiştir.

$$t_H = \frac{\overline{X}-\mu_0}{S/\sqrt{n}} \tag{8.4}$$

Eğer hipotez,

$$\begin{matrix} H_0 : \mu \le \mu_0 \\ H_A : \mu > \mu_0 \end{matrix} \quad \text{ya da} \quad \begin{matrix} H_0 : \mu = \mu_0 \\ H_A : \mu > \mu_0 \end{matrix}$$

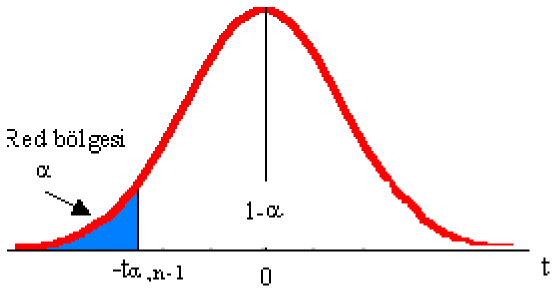
biçiminde ise, $t_H > t_{\alpha;n-1}$ olduğunda H_0 hipotezi reddedilir, $t_H \le t_{\alpha;n-1}$ olduğunda H_0 hipotezi reddedilemez. $t_{\alpha;n-1}$, Ek 5'de verilen n-1 serbestlik derecesinde α olasılığından alınan değerdir.



Problem için uygun hipotez,

$$\begin{matrix} H_0 : \mu \ge \mu_0 \\ H_A : \mu < \mu_0 \end{matrix} \quad \text{ya da} \quad \begin{matrix} H_0 : \mu = \mu_0 \\ H_A : \mu < \mu_0 \end{matrix}$$

biçiminde ise, $t_H < -t_{\alpha;n-1}$ olduğunda H_0 hipotezi reddedilir, $t_H \ge -t_{\alpha;n-1}$ olduğunda H_0 hipotezi reddedilemez.

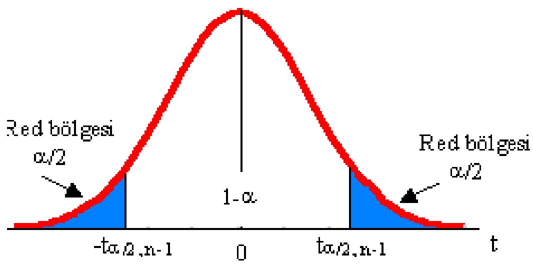


Eğer hipotez,

H_0: $\mu=\mu_0$

H_A: $\mu\neq\mu_0$

biçiminde ise, $t_H>t_{\alpha/2;n-1}$ ya da $t_H<-t_{\alpha/2;n-1}$ olduğunda H_0 hipotezi reddedilir, diğer durumda reddedilemez.



H_0: $\mu=\mu_0$, H_A: $\mu\neq\mu_0$ hipotezi güven aralığından yararlanılarak da, test edilebilir. Eğer H_0 hipotezinde verilen μ_0 ortalama değeri,

$$P\left(\overline{X}-t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}<\mu<\overline{X}+t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}\right)=1-\alpha$$

ile hesaplanan güven aralıklarının içine düşüyorsa, H_0 hipotezi reddedilemez.

Örnek 8.5: Belli bir şehirdeki 24 aylık çocukların ortalama ağırlığının 12,5 kg'dan az olduğu iddia ediliyor. Rasgele seçilen 5 çocuğun ortalaması 11 kg; standart sapması 1,17 kg olarak bulunuyor. α=0,10 anlamlılık düzeyinde aşağıdaki hipotezi test ediniz.

H_0:$\mu\geq 12,5$

H_A:$\mu<12,5$

$$t_H=\frac{11-12,5}{1,17/\sqrt{5}}=-2,88$$

t_H=-2,88 ve $-t_{0,10;\,4}$=-1,533'dır. $t_H<-t_{\alpha;\,n-1}$ olduğundan H_0 hipotezi reddedilir.

Yorum: 0,10 anlamlılık düzeyinde 24 aylık çocukların ortalama ağırlığının 12,5'dan az olduğu söylenebilir.

Örnek 8.5'de verilen soru, belli bir şehirdeki 24 aylık çocukların ortalama ağırlığının 12,5 kg'dan farklı olduğu iddia ediliyor biçiminde sorulmuş olsaydı, hipotez;

H_0: μ=12,5

H_A: $\mu \neq 12,5$

olurdu. t_H=-2,88 olmak üzere eksi bir değer olup, negatif bölge değeri olan $-t_{\alpha/2;n-1}$= -2,13 ile karşılaştırılır. –2,88<-2,13 olduğundan H_0 hipotezi reddedilir.

Yorum: 0,10 anlamlılık düzeyinde 24 aylık çocukların ortalama ağırlığının 12,5'den farklı olduğu söylenebilir.

8.3. KİTLE ORANINA AİT HİPOTEZ TESTLERİ

Oranlar ile ilgili testlerde, p_0 kitle oranı ve buna bağlı bulunan varyans kitle varyansı olacağı için $\dfrac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)/n}}$ oranının dağılımı, büyük n için standart normal dağılıma yakınsar. Burada kullanılacak test istatistiği aşağıda verilmiştir:

$$Z_H = \frac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)/n}} \qquad (8.5)$$

Eğer hipotez,

$$\begin{array}{l} H_0 : p \le p_0 \\ H_A : p > p_0 \end{array} \quad \text{ya da} \quad \begin{array}{l} H_0 : p = p_0 \\ H_A : p > p_0 \end{array}$$

biçiminde ise, $Z_H > Z_\alpha$ olduğunda H_0 hipotezi reddedilir, $Z_H \le Z_\alpha$ olduğunda H_0 hipotezi reddedilemez.

Eğer hipotez,

$$\begin{array}{l} H_0 : p \ge p_0 \\ H_A : p < p_0 \end{array} \quad \text{ya da} \quad \begin{array}{l} H_0 : p = p_0 \\ H_A : p < p_0 \end{array}$$

biçiminde ise $Z_H < -Z_\alpha$ olduğunda H_0 hipotezi reddedilir, $Z_H \ge -Z_\alpha$ olduğunda H_0 hipotezi reddedilemez.

Eğer hipotez,

H_0: p = p_0

H_A: p $\neq$ p_0

biçimde ise $Z_H < -Z_{\alpha/2}$ ya da $Z_H > Z_{\alpha/2}$ olduğunda H_0 hipotezi reddedilir, diğer durumda reddedilemez. Bu hipotez, güven aralığından yararlanılarak da, test edilebilir. Eğer H_0 hipotezinde verilen p_0 oranı,

$$P\left(\hat{p} - Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < p < \hat{p} + Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right) = 1-\alpha$$

ile hesaplanan güven aralıklarının içine düşüyorsa, H_0 hipotezi reddedilemez. Burada $\hat{p}$ örneklemden hesaplanan değerdir. $n \geq 30$ olduğu için standart normal dağılımdan (Z) değerler alınır.

Örnek 8.6: Bir sigorta şirketi eski müşterilerinin poliçelerini her yıl yenileme oranının 0,90 olmadığını iddia etmektedir. Rasgele seçilen 150 eski müşteriden 110'unun poliçelerini yenilediği görülmüştür. 0,10 anlamlılık düzeyinde iddianın doğruluğunu test ediniz.

H_0: p=0,90

H_A: $p \neq 0,90$

$$\hat{p} = \frac{110}{150} = 0,73$$

$$Z_H = \frac{0,73 - 0,90}{\sqrt{0,90(1-0,90)/150}} = -7,08 \qquad -Z_{0,025} = -1,65$$

-7,08<-1,65 olduğundan H_0 hipotezi reddedilir.

Yorum: 0,10 anlamlılık düzeyinde, poliçeleri yenileme oranının 0,90 olmadığı söylenebilir.

Örnek 8.7: Türkiye'de bir şirket 2004 yılı satışlarının geçen yılın mart ayı değerine göre 0,02 oranından fazla artırdığını iddia etmektedir. Rasgele seçilen 36 bayiinin satışları incelendiğinde, geçen yıla göre satışların 0,025 oranında arttığı saptanmıştır. Şirketin görüşünün doğruluğunu 0,05 anlamlılık düzeyinde test ediniz.

H_0: $p \leq 0,02$

H_A: $p > 0,02$

$$Z_H = \frac{0,025 - 0,02}{\sqrt{0,02(1-0,02)/36}} = 0,208$$

α=0,05 için $Z_{0,05=}$1,65'dir. $Z_H < Z_\alpha$ olduğundan H_0 hipotezi reddedilemez.

Yorum: 0,05 anlamlılık düzeyinde şirketin iddiası doğru değildir.

8.4. KİTLE VARYANSINA AİT HİPOTEZ TESTLERİ

Eğer normal dağılımlı kitleden bir örneklem çekip bu örneklemin çekildiği kitlenin varyansına ilişkin bir savı test etmek istersek;

$$\chi^2_{n-1} = \frac{(n-1)S^2}{\sigma^2}$$

dağılımı kullanılır. Bu dağılım Bölüm 6'da açıklanmıştır. Burada kullanılacak test istatistiği,

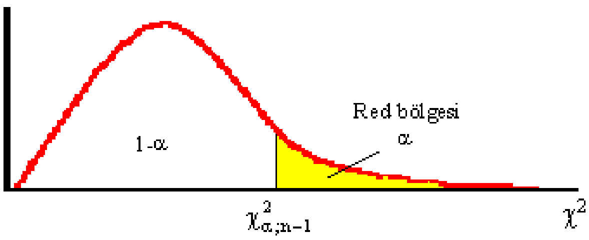
$$\chi^2_H = \frac{(n-1)S^2}{\sigma_0^2} \tag{8.6}$$

olur. Burada S^2: örneklem varyansı; $\sigma_0{}^2$: kitleye ait varyans için H_0'da verilen değer, χ^2_H: hesaplanan khi-kare değeridir. Konuya ilişkin hipotez ve sınamaları aşağıda verilmiştir:

Eğer hipotez,

$$H_0: \sigma^2 \le \sigma_0^2$$
$$H_A: \sigma^2 > \sigma_0^2$$

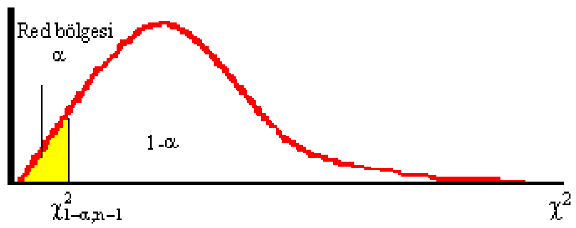
biçiminde ise, $\chi^2_H > \chi^2_{\alpha;n-1}$ olduğunda H_0 hipotezi reddedilir, $\chi^2_H \le \chi^2_{\alpha;n-1}$ olduğunda H_0 hipotezi reddedilemez. $\chi^2_{\alpha;n-1}$, α anlamlılık düzeyinde n-1 serbestlik derecesinde khi-kare tablo değeridir.



Eğer hipotez

$$H_0: \sigma^2 \ge \sigma_0^2$$
$$H_A: \sigma^2 < \sigma_0^2$$

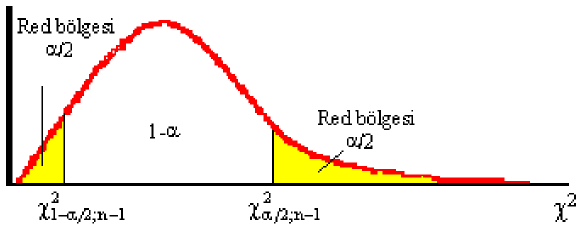
biçiminde ise, $\chi^2_H < \chi^2_{1-\alpha;n-1}$ olduğunda H_0 hipotezi reddedilir, $\chi^2_H \ge \chi^2_{1-\alpha;n-1}$ olduğunda H_0 hipotezi reddedilemez.



Eğer hipotez,

$$H_0 : \sigma^2 = \sigma_0^2$$
$$H_A : \sigma^2 \neq \sigma_0^2$$

biçiminde ise, $\chi_H^2 < \chi^2_{1-\alpha/2;n-1}$ ya da $\chi_H^2 > \chi^2_{\alpha/2;n-1}$ olduğunda H_0 hipotezi reddedilir, diğer durumda reddedilemez.



$$H_0 : \sigma^2 = \sigma_0^2$$
$$H_A : \sigma^2 \neq \sigma_0^2$$

hipotezi güven aralığı formülünden yararlanarak da bulunabilir. H_0 hipotezindeki σ_0^2 değeri,

$$P\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2;n-1}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\alpha/2;n-1}}\right) = 1-\alpha$$

güven aralıklarının içine düşüyorsa H_0 hipotezi reddedilemez.

Örnek 8.8: Sürücü belgesi almak için sınava girenler arasından 18 kişi rasgele seçilmiştir. Bu 18 kişinin aldığı puanların varyansı 4,41'dir. Kitle varyansının 6,25 olup olmadığını,

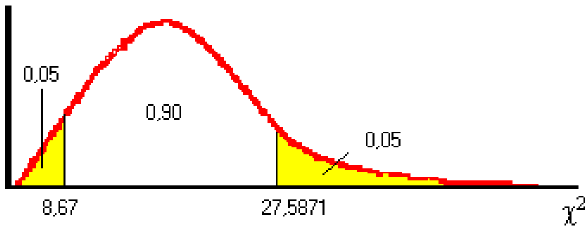
a- 0,10 anlamlılık düzeyinde test ediniz.

b- Güven aralığından yararlanarak test ediniz (0,80 güven düzeyi için).

a-
$$H_0 : \sigma^2 = 6{,}25$$
$$H_A : \sigma^2 \neq 6{,}25$$

$$\chi^2_H = \frac{(n-1)S^2}{\sigma_0^2} = \frac{(18-1)4{,}41}{6{,}25} = 11{,}99$$

α =0,10 iken $\chi^2_{0,95;17} = 8{,}67$; $\chi^2_{0,05;17} = 27{,}5871$ olur.



11,96 değeri 8,67 ile 27,5871 arasına düşüyor. Yani reddedilemeyen bölgededir. Bu nedenle H_0 hipotezi reddedilemez.

Yorum: 0,10 anlamlılık düzeyinde kitle varyansının 6,25 olduğu söylenebilir.

b- $H_0 : \sigma^2 = 6{,}25$
$H_A : \sigma^2 \neq 6{,}25$

0,80 güven düzeyinde $\chi^2_{1-\alpha/2;n-1} = \chi^2_{0,90;17} = 10{,}0852$; $\chi^2_{\alpha;n-1} = \chi^2_{0,10;17} = 24{,}7690$ olur.

$$P\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2;n-1}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\alpha/2;n-1}}\right) = 1-\alpha$$

$$P\left(\frac{(18-1)4{,}41}{24{,}7690} < \sigma^2 < \frac{(18-1)4{,}41}{10{,}0852}\right) = 0{,}80$$

$$P(3{,}026 < \sigma^2 < 7{,}433) = 0{,}80$$

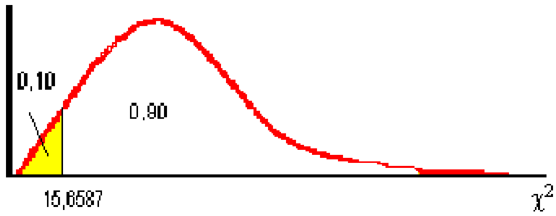
H_0 hipotezinde verilen 6,25 değeri, kitle varyansı için bulunan 3,026 ile 7,433 sınırları arasına düştüğü için H_0 hipotezi reddedilemez.

Yorum: 0,20 anlamlılık düzeyinde kitle varyansının 6,25 olduğu söylenebilir.

Örnek 8.9: Bir firma eski tip makine kullanarak 0,042 varyansla 10 cm. uzunluğunda vida üretmektedir. Firma, satın almak amacıyla 25 tane yeni makineyi denediğinde, aynı tip vidalar için 0,028 varyansını bulmaktadır. 0,10 anlamlılık düzeyinde yeni tip makineler ile aynı tip vidaların daha küçük varyansla imal edildiği ispatlanırsa, yeni makine satın alınacaktır. Bu örneklemde nasıl karar verilir?

$H_0 : \sigma^2 \geq 0{,}042$
$H_A : \sigma^2 < 0{,}042$

$$\chi^2_H = \frac{(n-1)S^2}{\sigma_0^2} = \frac{24.0,028}{0,042} = 16$$



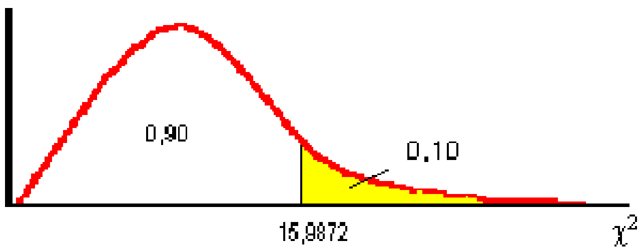
Bu hipotez için tablo değeri, $\chi^2_{0,90;24} = 15,6587$'dir. 16>15,6587 olduğundan H_0 reddedilemez.

Yorum: 0,10 anlamlılık düzeyinde kitle varyansının 0,042 eşit ya da daha büyük olduğu söylenebilir. Yeni makine almaya gerek yoktur.

Örnek 8.10: Küçük iletken parçalar üreten bir fabrikanın ürettiği ürünlerin kullanılabilir olması için, belirlenen kalınlıktan çok sapma göstermemesi gerekmektedir. Fabrikanın ürünlerini değişkenlik yönünden kontrol altında tutabilmesi için varyans 0,02 mm^2'den büyükse sistem durdurulacaktır. Günlük kontrolü yapan kalite kontrol mühendisi 11 parçada, parça kalınlıklarına ait varyansı 0,03 mm^2 olarak bulmuştur. Üretim kontrol altında mıdır? α=0,10 alınız.

$$H_0 : \sigma^2 \le 0,02$$
$$H_A : \sigma^2 > 0,02$$

$$\chi^2_H = \frac{(n-1)S^2}{\sigma_0^2} = \frac{10.0,03}{0,02} = 15 \quad ; \quad \chi^2_{0,10;10} = 15,9872$$



15<15,9872 olduğundan H_0 hipotezi reddedilemez.

Yorum: 0,10 anlamlılık düzeyinde kitle varyansının 0,02'den küçük veya eşit olduğunu söyleyebiliriz. Yani, sistem kontrol altındadır.

PROBLEMLER

1- Bir montaj hattında 1 saatte çıkan ürün miktarları, süreç içinden seçilen rasgele 20 saat için aşağıda verilmiştir:

1 saatte üretilen ürün sayıları :

3	5	3	4	6
2	3	3	5	3
1	8	4	4	4
4	2	3	3	2

a- Bu örneklem için bir saatte ortalama üretim miktarı ve varyansı ne kadardır?

b- Bu montaj hattında ortalama üretimin kitle ortalamasına ait 0,99 güven düzeyinde güven aralığını bulunuz.

c- Bu montaj hattında saatte ortalama 7 üretim yapıldığı iddia ediliyor. α=0,01 anlamlılık düzeyinde iddianın doğruluğunu test ediniz.

d- b ve c şıklarından elde edilen sonuçları karşılaştırıp tartışınız.

2- Bir boru fabrikası çapı 3,5 cm olan standart borular üretmektedir. Rasgele 36 boru seçiliyor ve her borunun çapı ölçülüyor. Aşağıda bu boruların çaplarına ilişkin bilgiler verildiğine göre. fabrikanın standarda uygun üretim yaptığı 0,01 ve 0,10 anlamlılık düzeylerinde söylenebilir mi? Bu iki sonucu karşılaştırınız.

$$\sum_{j=1}^{36} X_j = 216 \ ; \ \sum_{j=1}^{36}\left(X_j - \overline{X}\right)^2 = 1260 \ ; \ S^2 = \frac{\sum_{j=1}^{36}\left(X_j - \overline{X}\right)^2}{n-1}$$

3- Bir okuldaki başarılı öğrenci sayısı 409, başarısız öğrenci sayısı 240’dır.

a- 0,90 güven düzeyinde kitle başarı oranı için güven aralığını bulunuz.

b- 0,10 anlamlılık düzeyinde kitle başarı oranının 0,50 olup olmadığını test ediniz.

4- Bir ilaç fabrikasında üretilen ilaç paketlerinin ağırlığının 15 gram olması gerekiyor. İncelenen 20 paketin ağırlıkları için,

$$\sum_{j=1}^{20} X_j = 261 \ ; \ \sum_{j=1}^{20} X_j^2 = 3463$$

bulunmuştur. İlaç paketlerinin ağırlıklarının standarda uygun olup olmadığını 0,05 anlamlılık düzeyinde araştırınız.

5- Bir firmanın ürettiği lambalardan 100 tanesi seçilmiş ve ortalama ömürleri 1570 saat, standart sapmaları 120 saat bulunmuştur. 0,05 anlamlılık düzeyinde. firmanın ürettiği lambaların ortalama ömrünün a- 1600'den fazla

b- 1600'den az

c- 1600'den farklı

alternatif hipotezlerine göre çözünüz.

6- Bir ilaç fabrikasında üretilen ilaç paketlerinin ağırlığının 30 gr olması gerekiyor. İncelenen 49 paketin ağırlıklarının ortalaması 28 gr ve varyansı 196 gr^2 bulunmuştur. İlaç paketlerinin ağırlıklarının standarda uygun olup olmadığını 0,05 anlamlılık düzeyinde test ediniz. Sonucu yorumlayınız.

7- Bir boya üreticisi boyasının 1 litresinin ortalama olarak en az 600 cm^2 lik bir alanı boyayabileceğini iddia etmektedir. Piyasaya sürülen kutulardan 12 tanesi rasgele seçiliyor ve ortalama boyanan alan 589, standart sapma ise 26 olarak bulunuyor. Bu bilgiye göre 0,05 anlamlılık düzeyinde üreticinin iddiasının doğruluğunu test ediniz.

8- Bir borsa aracısı danışmanlık yaptığı müşterilerinin haftalık getirisinin ortalama olarak 100 YTL'den fazla olduğunu iddia etmektedir. Rasgele alınan 15 yatırımcının haftalık getiri ortalamasının 95 YTL ve standart sapmasının 3 YTL olduğu görüldüğüne göre. 0,05 anlamlılık düzeyinde aracı kurumunun iddiasını test ediniz.

9- Araba akülerinin üretildiği bir fabrikanın müdürü, akü ömrünün varyansının 0,81'den küçük olduğunu iddia etmektedir. Üretilen aküler arasından rasgele seçilen 10 akü için standart sapma 1,2 yıl bulunmuş ise 0,10 anlamlılık düzeyinde iddiayı test ediniz.

10- ATM makineleri için kullanılan bir programda, yapılan işlemlere ait harcanan süre için varyans 16 dak^2 olarak kabul edilmektedir. Yeni bir işlemin devreye girmesiyle bu varyans'ın arttığı düşünülmektedir. Varyans'a ilişkin bu iddiayı doğrulamak için 16 deneme uygulanmış ve standart sapma 6 dakika olarak bulunmuştur. 0,10 anlamlılık düzeyinde iddianın doğruluğunu test ediniz. Sonucu yorumlayınız.

11- Bir ilde 40 yaşın üzerinde olanların %20'den fazlasının diş protezi kullandığı öne sürülmektedir. Bu yöredeki 40 yaşın üzerinde rasgele 200 kişi seçilmiş ve 45 kişide protez görülmüştür. 0,10 anlamlılık düzeyinde öne sürülen fikrin doğruluğunu test ediniz.

12- Metal levha ölçümlerinin varyansının 0,18'den büyük olup olmadığı hipotezi test edilmek isteniyor. Yeni bir kontrolör 25 metal levha ölçmüş ve varyansı 0,25 olarak bulmuştur. Varyansa ilişkin hipotezi kurunuz ve 0,05 anlamlılık düzeyinde hipotezi test ediniz.

13- 5 uzman rasgele seçilip İş Bankası B tipi hisse senetlerinin 2005 yılı pay senedi başına kazancını tahmin etmeleri istenmiştir. Bu 5 kişiden alınan değerlerin standart sapması 1,3 olarak bulunmuştur. 0,10 anlamlılık düzeyinde kitle varyansının 1,69'dan büyük olup olmadığı hipotezini test ediniz.

14- Bir konserve fabrikasının yaptığı konservelerden 9 kutu seçiliyor. Seçilen 9 kutunun ağırlıklarının ortalaması 975 gr ve standart sapması 30 gr olan normal dağılım göstermektedir. Kitle ağırlıklarına ilişkin ortalamanın 1000 gr olup olmadığını 0,05 anlamlılık düzeyinde test ediniz.

15- Bir iş yerinde çalışan işçilere verilen öğle yemeğinin 1800 kalori değerinde olması gerekiyor. 16 gün çıkan yemeklerin kalori değerleri ölçülmüş ve

1500	1400	1700	1600	1700	1680	1710	1600
1700	1800	1600	1750	1600	1500	1760	1750

olarak bulunmuştur. 0,01 anlamlılık düzeyinde yemeklerin kurallara uygun verildiği söylenebilir mi?

16- Radyoaktif serpintinin olmadığı bir bölgede yeni doğan bebeklerin 0,001'i genetik bozukluk göstermektedir. Aynı bölgede radyoaktif serpinti deneyler sonucu artmış ve yeni yapılan tarama sonucu bu yörede bir yıl içinde 10000 bebekten 20'sinde genetik bozukluk gözlenmiştir. Radyoaktif serpintinin genetik bozukluk oranını artırıp artırmadığını 0,01 anlamlılık düzeyinde test ediniz.

17- Bulaşık deterjanını plastik kaplara dolduran bir makinede dolumlarının varyansının 9 kg^2 dan fazla olduğu iddia ediliyor. Makineyi test etmek için 7 rasgele örneklem seçilmiş ve varyans 1,2 kg^2 olarak bulunmuştur. Varyansa ilişkin bu iddianın doğruluğunu 0,20 anlamlılık düzeyinde test ediniz.

18- Bir firma tarafından piyasaya sürülen kutu ağırlıkları ile ilgili bir araştırma yapılmıştır. Firma kutuları arasından 5'ini rasgele ve bunların ağırlıkları ortalamasını 30 gr ve varyansını 20 gr^2 olarak buluyor. 0,01 anlamlılık düzeyinde kutuların ağırlıklarının kitle ortalamasının 30 gr'dan büyük olduğu iddiasını test ediniz

19- Bir otomobil firması ürettiği arabaların fren mesafeleri için bir araştırma yapmıştır. Bu amaçla 90km/saat hızla giden 50 araçta durma uzaklıklarını ölçmüştür. Ortalama 21 metre, standart sapma 12 metre olarak bulunmuştur. Firma durma mesafesinin 20 metreden kısa olduğunu savunmaktadır. Bu iddianın doğruluğunu 0,05 anlamlılık düzeyinde test ediniz.

20- Yapılan bir piyasa araştırmasına göre 2006 yılında internet üzerinden yapılan alışverişlerde 0,14 oranında artış kaydedilmiştir. Bir bilgisayar oyun firmasında yıllık satış 3700 parça olup bunun 0,18'i internet üzerindendir. Firmanın E-ticaret satışlarında olan artışın, piyasa genelinin üstüne çıktığı 0,05 anlamlılık düzeyinde söylenebilir mi?

9. DOKUZUNCU BÖLÜM

9.1. GİRİŞ

Çoğu zaman farklı kitlelerin parametrelerinin karşılaştırılması istenir. Burada, iki kitlenin ortalamaları, varyansları ve oranları arasında bir fark olup olmadığı, birinin diğerinden büyük ya da küçük olduğu hipotez testleri üzerinde durulacaktır. Tek kitle hipotez testleri için izlenen adımlar burada da uygulanacaktır.

İki kitle ortalamaları μ_1 ve μ_2, varyansları σ_1^2 ve σ_2^2 ile normal dağılıma sahip olsunlar. Bu iki kitlenin farklı olup olmadığını araştırmak için, varyanslarının ve ortalamalarının farklı olup olmadığının araştırılması gerekir. Öncelikle, iki kitle varyanslarının farkına ilişkin hipotez testleri verilecektir.

9.2. İKİ KİTLE VARYANSINA İLİŞKİN HİPOTEZ TESTLERİ

İki kitle varyansının karşılaştırılması ile ilgili hipotezler için uygulamada genellikle tek yanlı testler kullanılır.

Tek yanlı	**Tek yanlı**	**İki yanlı**
H_0: $\sigma_1^2-\sigma_2^2=0$	H_0: $\sigma_1^2-\sigma_2^2=0$	H_0: $\sigma_1^2-\sigma_2^2=0$
H_A: $\sigma_1^2-\sigma_2^2>0$	H_A: $\sigma_1^2-\sigma_2^2<0$	H_A: $\sigma_1^2-\sigma_2^2\neq 0$
H_0: $\sigma_1^2=\sigma_2^2$	H_0: $\sigma_1^2=\sigma_2^2$	H_0:$\sigma_1^2=\sigma_2^2$
H_A: $\sigma_1^2>\sigma_2^2$	H_A: $\sigma_1^2<\sigma_2^2$	H_A: $\sigma_1^2\neq\sigma_2^2$
H_0: $\sigma_1^2/\sigma_2^2=1$	H_0: $\sigma_1^2/\sigma_2^2=1$	H_0: $\sigma_1^2/\sigma_2^2=1$
H_A: $\sigma_1^2/\sigma_2^2>1$	H_A: $\sigma_1^2/\sigma_2^2<1$	H_A: $\sigma_1^2/\sigma_2^2\neq 1$

Yukarıda alt alta 3 farklı şekilde verilen hipotezler aynı anlamdadır. Bu hipotezlerin testinde kullanılacak olasılık dağılımı, Bölüm 6.3'de anlatılan F dağılımıdır. F istatistiği H_A daki σ_1^2 , σ_2^2 ile ilişkilidir. Tek yanlı hipotez, H_A: $\sigma_1^2<\sigma_2^2$ ise;

$$F_H = \frac{S_2^2}{S_1^2}$$

dir. Tek yanlı hipotez, H_A: $\sigma_1^2>\sigma_2^2$ ise;

$$F_H = \frac{S_1^2}{S_2^2} \qquad (9.1)$$

alınmalıdır. Eşitlik (9.1)'de, serbestlik derecesi sırasıyla, $sd_1=n_1-1$ ve $sd_2=n_2-1$'dir. Burada; n_1 ve n_2 sırasıyla 1. ve 2. örneklemin büyüklüğü; S_1^2 ve S_2^2 örneklemlerin varyanslarıdır. Örnek 9.1'de verilecek örnek ile hipotezin test edilmesindeki aşamalar açıklanacaktır.

Örnek 9.1: İki farklı yöntem ile üretilen demir çubuklarının kırılma dirençlerine ilişkin varyansların karşılaştırılması istenmiştir. Karşılaştırmayı yapabilmek için A yöntemi ile üretilen 11 ve B yöntemi ile üretilen 21 tane ürün rasgele seçilmiştir. A ve B yöntemleri ile üretilen ürünlerin varyansları sırasıyla 529 ve 361 olarak elde edildiğine göre, 0,05 anlamlılık düzeyinde A yöntemi ile üretilen ürünlerin kırılma dirençlerine ait varyansın daha büyük olup olmadığını test ediniz.

1. Hipotez kurulur.

   H_0: $\sigma_1^2=\sigma_2^2$

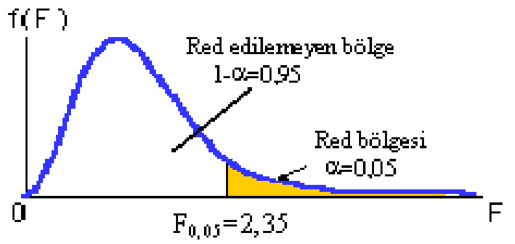
   H_A: $\sigma_1^2>\sigma_2^2$

2. α anlamlılık düzeyi belirlenir. Bu soru için α=0,05 alınacaktır.
3. F istatistiği hesaplanır.

$$F_H = \frac{529}{361} = 1{,}465$$

4. Ürünlere ilişkin serbestlik dereceleri ve bu serbestlik derecelerine ait Ek 7'de verilen F tablo değeri bulunur.

$sd_1=n_1-1=10$ ve $sd_2=n_2-1=20$ olup α=0,05 için F tablo değeri 2,35'dir.



5. Hipotez hakkında karar verilir. $F_H > F_{\alpha;n_1-1;n_2-1}$ ise H_0 hipotezi reddedilir, diğer durumda reddedilemez.

Problemde 1,465<2,35 olduğundan H_0 reddedilemez.

Yorum: 0,05 anlamlılık düzeyinde, bu iki farklı yöntem ile üretilen ürünlerin demir çubukların kırılma dirençlerine ait varyansın eşit olduğu söylenebilir.

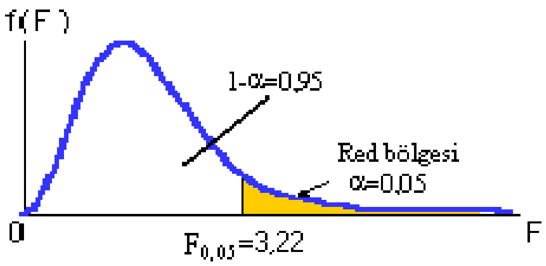
Örnek 9.2: Bir banka ATM işlemlerinde harcanan zamandaki değişkenliği (varyansı) azaltacak yeni bir algoritma geliştirmiştir. Bu yeni yazılımın harcanan zamanla ilgili varyansı küçülttüğü iddia edilmektedir. Bu iddianın doğruluğunu sınamak için eski yöntem ile 7, yeni yöntem ile 11 işlem uygulanmış ve varyanslar sırasıyla 612,68 sn^2 ve 51,49 sn^2 olarak bulunmuştur. Bankanın iddiasının doğruluğunu 0,05 anlamlılık düzeyinde test ediniz.

H_0: $\sigma_1^2=\sigma_2^2$

H_A: $\sigma_1^2>\sigma_2^2$

$$F_H=\frac{612,68}{51,49}=11,898$$

α=0,05 için 6;10 serbestlik derecesindeki F tablo değeri 3,22'dir.



11,898>3,22 olduğundan, H_0 hipotezi reddedilir.

Yorum: 0,05 anlamlılık düzeyinde eski yöntem ile harcanan zaman daha fazla değişkenlik göstermektedir.

9.3. İKİ KİTLE ORTALAMASINA İLİŞKİN HİPOTEZ TESTLERİ VE GÜVEN ARALIĞI

İki grubu karşılaştırma işlemi, bu grupların aynı ya da farklı kitlelerden çekilip çekilmediklerinin araştırılmasıdır. Bu karşılaştırma önce varyanslar sonra da ortalamaların karşılaştırılması işlemidir. Ortalamaların karşılaştırılması, grupların bağımlı ya da bağımsız olması durumuna göre iki farklı şekilde ele alınır. Bağımsız grupların karşılaştırılması kitle varyansının bilinmesi ya da bilinmemesi durumları için yine iki farklı yöntem ile test edilir. Kitle varyansları bilinmeyen iki grubun karşılaştırılması gruplarda bulunan gözlem sayılarının 30'dan büyük ya da küçük olmasına göre iki farklı yöntem kullanılır. Bu ayrıntılar dikkate alınarak ortalamaların karşılaştırılması bu bölümde verilecektir.

İki Bağımsız Grubun Kitle Ortalamalarının Karşılaştırılması ve Güven Aralığı

İki bağımsız grubun ortalamalarını karşılaştırmak için kurulabilecek hipotezler aşağıda verilmiştir:

Tek Yanlı	**Tek Yanlı**	**İki Yanlı**
$H_0:\mu_1-\mu_2\leq 0$	$H_0:\mu_1-\mu_2\geq 0$	$H_0:\mu_1-\mu_2=0$
$H_A:\mu_1-\mu_2>0$	$H_A:\mu_1-\mu_2<0$	$H_A:\mu_1-\mu_2\neq 0$
$H_0:\mu_1\leq\mu_2$	$H_0:\mu_1\geq\mu_2$	$H_0:\mu_1=\mu_2$
$H_A:\mu_1>\mu_2$	$H_A:\mu_1<\mu_2$	$H_A:\mu_1\neq\mu_2$

Not: Bu hipotezlerde genel olarak $\mu_1-\mu_2\leq D_0$ ya da $\mu_1-\mu_2\geq D_0$ ya da $\mu_1-\mu_2=D_0$ olarak alınır. Burada kullanılacak olan hipotezler $D_0=0$ için özelleştirilmiştir.

Yukarıda iki farklı şekilde verilen hipotezlerin testinde kullanılacak test istatistikleri için belirleyici olan iki kitlenin varyanslarının durumudur. Aşağıda bu durumlara ilişkin hipotez ve güven aralığı verilecektir.

σ^2_1 ve σ^2_2 Bilindiğinde Hipotez Testi ve Güven Aralığı

İki kitleye ait σ_1^2 ve σ_2^2 varyansları biliniyor ve kitlenin dağılımı normal ve bağımsız ise, hipotezleri test etmek için kullanılacak istatistik,

$$Z_H=\frac{\left(\overline{X}_1-\overline{X}_2\right)-\left(\mu_1-\mu_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \tag{9.2}$$

dir. Burada eğer hipotez;

$H_0:\mu_1-\mu_2\leq 0$ ya da $H_0:\mu_1\leq\mu_2$

$H_A:\mu_1-\mu_2>0$ $H_A:\mu_1>\mu_2$

biçiminde ise, $Z_H>Z_\alpha$ olduğunda H_0 hipotezi reddedilir, diğer durumda reddedilemez.

Eğer hipotez,

$H_0:\mu_1-\mu_2\geq 0$ ya da $H_0:\mu_1\geq\mu_2$

$H_A:\mu_1-\mu_2<0$ $H_A:\mu_1<\mu_2$

biçiminde ise, $Z_H<-Z_\alpha$ olduğunda H_0 hipotezi reddedilir, diğer durumda reddedilemez.

Eğer hipotez,

$H_0:\mu_1-\mu_2=0$ ya da $H_0:\mu_1=\mu_2$

$H_A:\mu_1-\mu_2\neq 0$ $H_A:\mu_1\neq\mu_2$

biçiminde ise, $Z_H > Z_{\alpha/2}$ ya da $Z_H < -Z_{\alpha/2}$ olduğunda H_0 hipotezi reddedilir, diğer durumda reddedilemez.

σ^2_1 ve σ^2_2 bilindiğinde, iki kitle ortalaması arasındaki farkın güven aralığı,

$$P\left[\left(\overline{X}_1-\overline{X}_2\right)-Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}<\mu_1-\mu_2<\left(\overline{X}_1-\overline{X}_2\right)+Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right]=1-\alpha \qquad (9.3)$$

eşitliği ile bulunur.

Örnek 9.3: Bir firma A ve B olmak üzere iki ayrı şirketten mal almaktadır. Firma, A ve B şirketlerinin ürünlerinin paket ağırlıklarını karşılaştırmak istemiştir. Önceki bilgilerden A şirketinin ürün paket ağırlıklarının varyansı 34,5 kg^2 ve B şirketinin ürün paket ağırlıklarının varyansı 38,4 kg^{2}'dır. A ve B şirketlerinden sırasıyla 10 ve 8'lik örneklemler seçilmiş ve ortalamalar sırasıyla 150 kg ve165 kg olarak bulunmuştur.

a- 0,05 anlamlılık düzeyinde, ürünlerin paket ağırlıkları arasında fark olup olmadığını test ediniz.

b- 0,95 güven düzeyinde ortalamalar arası fark için güven aralığı bulunuz.

a- $H_0:\mu_1-\mu_2=0$

$H_A:\mu_1-\mu_2\neq 0$

$$Z_H=\frac{\left(\overline{X}_1-\overline{X}_2\right)-\left(\mu_1-\mu_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}=\frac{(150-165)-0}{\sqrt{\frac{34,5}{10}+\frac{38,4}{8}}}=-5,2223$$

$Z_H=-5,2223$ negatif değerdir ve negatif bölgedeki $-Z_{0,025}=-1,96$ değeri ile karşılaştırılır. $-5,2223<-1,96$ olduğu için H_0 hipotezi reddedilir.

Yorum: 0,05 anlamlılık düzeyinde şirketlerin ürünlerinin paket ağırlıkları arasında fark vardır.

b- $$P\left[\left(\overline{X}_1-\overline{X}_2\right)-Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}<\mu_1-\mu_2<\left(\overline{X}_1-\overline{X}_2\right)+Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right]=1-\alpha$$

$$P\left[(150-165)-1,96\sqrt{\frac{34,5}{10}+\frac{38,4}{8}}<\mu_1-\mu_2<(150-165)+1,96\sqrt{\frac{34,5}{10}+\frac{38,4}{8}}\right]=0,95$$

$$P\left[-20,5331<\mu_1-\mu_2<-9,3703\right]=0,95$$

Yorum: 0,95 güven düzeyinde şirketlerin ürünlerinin paket ağırlıkları ortalamaları farkının güven aralığı –20,5331 ile –9,3703 arasındadır. Güven aralığında sıfırın bulunmaması iki grup arasındaki farkın önemli olduğunun başka bir göstergesidir.

σ^2_1 ve σ^2_2 Bilinmediğinde Büyük Örneklemlerde Hipotez Testi ve Güven Aralığı

İki kitleye ait σ_1^2 ve σ_2^2 varyansları bilinmiyor ve n_1, $n_2 \geq 30$ ise, hipotezleri test etmek için kullanılacak istatistik,

$$Z_H = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \qquad (9.4)$$

dir. Burada eğer hipotez;

$H_0: \mu_1 - \mu_2 \leq 0$ ya da $H_0: \mu_1 \leq \mu_2$

$H_A: \mu_1 - \mu_2 > 0$ $\quad$ $H_A: \mu_1 > \mu_2$

biçiminde ise, $Z_H > Z_\alpha$ olduğunda H_0 hipotezi reddedilir, diğer durumda reddedilemez. Eğer hipotez,

$H_0: \mu_1 - \mu_2 \geq 0$ ya da $H_0: \mu_1 \geq \mu_2$

$H_A: \mu_1 - \mu_2 < 0$ $\quad$ $H_A: \mu_1 < \mu_2$

biçiminde ise $Z_H < -Z_\alpha$ olduğunda H_0 hipotezi reddedilir, diğer durumda reddedilemez. Eğer hipotez,

$H_0: \mu_1 - \mu_2 = 0$ ya da $H_0: \mu_1 = \mu_2$

$H_A: \mu_1 - \mu_2 \neq 0$ $\quad$ $H_A: \mu_1 \neq \mu_2$

biçiminde ise $Z_H > Z_{\alpha/2}$ ya da $Z_H < -Z_{\alpha/2}$ olduğunda H_0 hipotezi reddedilir, diğer durumda reddedilemez.

σ^2_1 ve σ^2_2 bilinmediğinde, büyük örneklemlerde iki kitle ortalaması farkın güven aralığı,

$$P\left[(\overline{X}_1 - \overline{X}_2) - Z_{\alpha/2}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} < \mu_1 - \mu_2 < (\overline{X}_1 - \overline{X}_2) + Z_{\alpha/2}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}\right] = 1 - \alpha \qquad (9.5)$$

eşitliği ile bulunur.

Örnek 9.4. Bir ilaç aynı hastalığa yakalanmış iki hasta grubuna veriliyor. Hastalığın iyileşme sürelerine ilişkin veriler aşağıdadır:

I. Grup	**II. Grup**
$n_1=40$	$n_2=50$
$\overline{X}_1 = 32$ saat	$\overline{X}_2 = 30$ saat
$S_1 = 20$ saat	$S_2 = 18$ saat

a- 0,05 anlamlılık düzeyinde ilacın her iki grupta da aynı etkiyi gösterdiği söylenebilir mi?

b- 0,95 güven düzeyinde hastalıkların ortalama iyileşme süreleri farkına ilişkin güven aralığını bulunuz.

a- $H_0: \mu_1 = \mu_2$

$H_A: \mu_1 \neq \mu_2$

$$Z_H = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{(32-30)-0}{\sqrt{\frac{20^2}{40} + \frac{18^2}{50}}} = 0{,}49$$

Z_H=0,49 pozitif değerdir ve pozitif bölgedeki $Z_{0,025}$=1,96 değeri ile karşılaştırılır. 0,49<1,96 olduğu için H_0 hipotezi reddedilemez.

Yorum: 0,05 anlamlılık düzeyinde, ilacın her iki grupta da aynı etkiyi gösterdiği söylenebilir.

b- $$P\left[(\overline{X}_1 - \overline{X}_2) - Z_{\alpha/2}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} < \mu_1 - \mu_2 < (\overline{X}_1 - \overline{X}_2) + Z_{\alpha/2}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}\right] = 1 - \alpha$$

$$P\left[(32-30) - 1{,}96\sqrt{\frac{20^2}{40} + \frac{18^2}{50}} < \mu_1 - \mu_2 < (32-30) + 1{,}96\sqrt{\frac{20^2}{40} + \frac{18^2}{50}}\right] = 0{,}95$$

$$P[-5{,}9576 < \mu_1 - \mu_2 < 9{,}9576] = 0{,}95$$

Yorum: 0,95 güven düzeyinde hastalıkların ortalama iyileşme süreleri farkının güven aralığı –5,9576 ile 9,9576 arasındadır. Güven aralığında sıfırın bulunması iki grup arasındaki farkın önemsiz olduğunun başka bir göstergesidir.

σ^2_1 ve σ^2_2 Bilinmediğinde Küçük Örneklemlerde Hipotez Testi ve Güven Aralığı

Bu durumda, hipotez testi ve güven aralığı iki kitle varyansının eşit ya da farklı olmasına göre yapılır. Her iki durumda da test istatistiğinin dağılımı Bölüm 6.3'de gördüğümüz t dağılımıdır.

$\sigma^2_1 = \sigma^2_2$ durumunda hipotezi test etmek için kullanılacak test istatistiği;

$$t_H = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \qquad (9.6)$$

olur. Burada,

$$S_p = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$$

dir. t dağılımı için kullanılacak serbestlik derecesi; n_1+n_2-2'dir. Ek 5 ile verilen t tablosundan bu serbestlik derecesine karşılık gelen değer bulunur.

Eğer hipotez;

$H_0:\mu_1-\mu_2\leq 0$ ya da $H_0:\mu_1\leq\mu_2$

$H_A:\mu_1-\mu_2>0$ $H_A:\mu_1>\mu_2$

biçiminde ise, $t_H > t_{\alpha;\, n_1+n_2-2}$ olduğunda H_0 hipotezi reddedilir, diğer durumda reddedilemez.

Eğer hipotez,

$H_0:\mu_1-\mu_2\geq 0$ ya da $H_0:\mu_1\geq\mu_2$

$H_A:\mu_1-\mu_2<0$ $H_A:\mu_1<\mu_2$

biçiminde ise, $t_H < -t_{\alpha;\, n_1+n_2-2}$ olduğunda H_0 hipotezi reddedilir, diğer durumda reddedilemez.

Eğer hipotez,

$H_0:\mu_1-\mu_2=0$ ya da $H_0:\mu_1=\mu_2$

$H_A:\mu_1-\mu_2\neq 0$ $H_A:\mu_1\neq\mu_2$

biçiminde ise $t_H > t_{\alpha/2;\, n_1+n_2-2}$ ya da $t_H < -t_{\alpha/2;\, n_1+n_2-2}$ olduğunda H_0 hipotezi reddedilir, diğer durumda reddedilemez.

$\sigma^2_1=\sigma^2_2$ Durumunda Güven Aralığı

$$P\left[\left(\overline{X}_1-\overline{X}_2\right)-t_{\alpha/2;n_1+n_2-2}\,S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}<\mu_1-\mu_2<\left(\overline{X}_1-\overline{X}_2\right)+t_{\alpha/2;n_1+n_2-2}\,S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right]=1-\alpha \tag{9.7}$$

Burada S_p, Eşitlik (9.6)'da verildiği gibidir. $t_{\alpha/2;\; n1+n2-2}$ yerine; Ek 5'de verilen t tablosunda tek yanlı olasılık değerleri satırında $\alpha/2$, iki yanlı olasılık satırında α'daki, n_1+n_2-2 serbestlik derecesinin t tablo değeri yazılır.

Örnek 9.5: Siyasi bir araştırmada farklı iki partiye oy veren seçmenlerin gelir durumlarının aynı olup olmadığı test edilmek istenmiştir. Bu amaçla A partisine ve B partisine oy veren 15'er seçmen rasgele alınmış ve gelir durumları incelenmiştir.

Örneklemin çekildiği kitlede gelir durumunun normal dağılıma sahip olduğu kabul edilirse, 0,05 anlamlılık düzeyinde gelir durumlarının aynı olup olmadığını test ediniz.

A partisi	**B partisi**
$n_1=15$	$n_2=15$
$\overline{X}_1 = 2255$ YTL	$\overline{X}_2 = 2140$ YTL
$S_1=645$ YTL	$S_2=708$ YTL

Öncelikle kitle varyanslarının eşitliğinin kontrol edilmesi gerekir. Bu amaçla aşağıdaki hipotez kurulur:

H_0: $\sigma_1^2=\sigma_2^2$

H_A: $\sigma_2^2>\sigma_1^2$

$F_H = \dfrac{708^2}{645^2} = 1{,}205$; $F_{0,05;14;14} = 2{,}484$ 'dir. 1,205<2,484 olduğundan H_0 hipotezi reddedilemez. Yani iki kitle varyansları eşittir.

Bu durumda ortalamaları karşılaştırmak için eşit varyans durumuna ait formül kullanılmalıdır.

H_0: $\mu_1=\mu_2$

H_A: $\mu_1\neq\mu_2$

$$S_p = \sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}} = \sqrt{\frac{(15-1)645^2+(15-1)708^2}{15+15-2}} = 677{,}23$$

$$t_H = \frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{(2255-2140)-(0{,}0)}{677{,}23\sqrt{\frac{1}{15}+\frac{1}{15}}} = 0{,}465$$

hipotez iki yanlı kurulduğundan, anlamlılık düzeyi $\alpha/2=0{,}05/2=0{,}025$ ve serbestlik derecesi 15+15-2=28 olan t tablo değerine bakılır ki bu değer 2,0484'dir. 0,465<2,0484 olduğundan H_0 hipotezi reddedilemez.

Yorum: 0,05 anlamlılık düzeyinde gelir durumlarının aynı olduğu söylenebilir.

Örnek 9.6: Bir ihracat firması iki farklı ürün ihraç etmektedir. A ve B ürünlerinin satışlarına ilişkin bilgiler aşağıda verilmiştir. Bilgilere göre, iki ortalama farkının 0,95 güven düzeyinde güven aralığını bulunuz ($\sigma^2_1=\sigma^2_2$ varsayılıyor).

A bölüm	**B bölüm**
$n_1=15$	$n_2=7$
$\overline{X}_1 = 556$ YTL	$\overline{X}_2 = 600$ YTL
$S_1=200$ YTL	$S_2=275$ YTL

$$S_p = \sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}} = \sqrt{\frac{14.200^2+6.275^2}{15+7-2}} = 233{,}85$$

$t_{\alpha/2}$ yerine; $\alpha/2=0{,}05/2=0{,}025$ anlamlılık düzeyi ve serbestlik derecesi 15+7–2=20 olan t tablo değeri yazılır ki bu değer 2,086'dır.

$$P\left[(556-600)-2{,}086.233{,}85\sqrt{\frac{1}{15}+\frac{1}{7}} < \mu_1-\mu_2 < (556-600)+2{,}086.233{,}85\sqrt{\frac{1}{15}+\frac{1}{7}}\right] = 0{,}95$$

$$P[-267{,}25 < \mu_1-\mu_2 < 179{,}25] = 0{,}95$$

Yorum: İki kitle ortalamaları arasındaki farkın 0,95 güven düzeyinde güven aralığı, -267,25 ile 179,25 arasındadır.

$\sigma^2_1 \neq \sigma^2_2$ Durumunda μ_1-μ_2 İçin Hipotez Testi

$$t_H = \frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}} \tag{9.8}$$

bu eşitlik için kullanılacak serbestlik derecesi;

$$sd = \frac{\left(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1-1}+\frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2-1}} \tag{9.9}$$

dir. Eşitlik (9.9)'dan ondalıklı bir sayı elde edilir ise, bu sayı en yakın tam sayı olarak alınır. Ek 5'de verilen t tablosundan bu serbestlik derecesine karşılık gelen değer bulunur.

Eğer hipotez;

$H_0:\mu_1-\mu_2\leq 0$	ya da	$H_0:\mu_1\leq\mu_2$
$H_A:\mu_1-\mu_2>0$		$H_A:\mu_1>\mu_2$

biçiminde ise $t_H>t_{\alpha;\ sd}$ olduğunda H_0 hipotezi reddedilir, diğer durumda reddedilemez.

Eğer hipotez,

$H_0:\mu_1-\mu_2\geq 0$ ya da $H_0:\mu_1\geq\mu_2$

$H_A:\mu_1-\mu_2<0$ $H_A:\mu_1<\mu_2$

biçiminde ise $t_H<-t_{\alpha;\,sd}$ olduğunda H_0 hipotezi reddedilir, diğer durumda reddedilemez.

Eğer hipotez,

$H_0:\mu_1-\mu_2=0$ ya da $H_0:\mu_1=\mu_2$

$H_A:\mu_1-\mu_2\neq 0$ $H_A:\mu_1\neq\mu_2$

biçiminde ise $t_H>t_{\alpha/2;\,sd}$ ya da $t_H<-t_{\alpha/2;\,sd}$ ise H_0 hipotezi reddedilir, diğer durumda reddedilemez.

$\sigma^2_1\neq\sigma^2_2$ Durumda μ_1-μ_2 İçin Güven Aralığı

$$P\left[\left(\overline{X}_1-\overline{X}_2\right)-t_{\alpha/2;\,sd}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}<\mu_1-\mu_2<\left(\overline{X}_1-\overline{X}_2\right)+t_{\alpha/2;\,sd}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}\right]=1-\alpha \quad (9.10)$$

$t_{\alpha/2;\,sd}$ yerine, Eşitlik (9.9)'dan hesaplanan serbestlik dercesine sahip Ek 5'de verilen t tablo değeri yazılır.

Örnek 9.7: Bir firma müşterilerinin şikayet ve isteklerini internet aracılığıyla belirlemek amacıyla yeni bir yazılım geliştirmiştir. Geliştirilen programın eski yöntem ile ortalama kullanım süresi yönünden karşılaştırılması için rasgele 8 müşteri eski yöntem ile, 10 müşteri yeni algoritma ile şikayetlerini belirtmişlerdir. Alınan sonuçlar aşağıda verilmiştir. Yeni yöntem farklılık yaratmış mıdır? α=0,05 alınız.

$n_1=8$ $n_2=10$

$\overline{X}_1=10$ dak $\overline{X}_2=7$ dak

$S_1^2=25$ dak^2 $S_2^2=5$ dak^2

Öncelikle kitle varyanslarının eşit olup olmadığı test edilmelidir. Bu amaçla öncelikle aşağıdaki hipotez kurulabilir:

H_0: $\sigma_1^2=\sigma_2^2$

H_A:$\sigma_1^2>\sigma_2^2$

$F_H = \frac{25}{5} = 5$; $F_{0,05;7;9} = 3{,}20$'dir. 5>3,20 olduğundan H_0 hipotezi reddedilir. Yani iki kitle varyansları eşit değildir. O halde farklı varyans durumuna ait formül kullanılmalıdır. Probleme ilişkin hipotez aşağıdaki gibi kurulur.

$H_0: \mu_1 \leq \mu_2$

$H_A: \mu_1 > \mu_2$

t hesap değeri,

$$t_H = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{(10-7)-(0)}{\sqrt{\frac{25}{8} + \frac{5}{10}}} = 1{,}575$$

olarak bulunur, tablo için kullanılacak serbestlik derecesi Eşitlik (9.9) kullanılarak;

$$sd = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2 - 1}} = \frac{\left(\frac{25}{8} + \frac{5}{10}\right)^2}{\frac{\left(\frac{25}{8}\right)^2}{7} + \frac{\left(\frac{5}{10}\right)^2}{9}} = 9{,}25 \approx 9$$

alınabilir.

9 serbestlik derecesinde 0,05 anlamlılık düzeyinde t tablo değeri 1,8331'dir. Bu değer 1,575 hesap değerinden büyük olduğu için H_0 hipotezi reddedilemez.

Yorum: 0,05 anlamlılık düzeyinde yeni yöntemin zamanı kısalttığı söylenemez.

Örnek 9.8: Örnek 9.7'de verilen soru için $\sigma^2_1 \neq \sigma^2_2$ varsayımı altında, 0,95 güven düzeyinde güven aralığını bulunuz.

Tablo değeri için, tablodaki tek yanlı $\alpha/2 = 0{,}05/2 = 0{,}025$'de, ya da iki yanlı 0,05'deki, 9 serbestlik dercesine sahip değer kullanılır ki bu değer, 2,2622'dir.

$$P\left[(10-7) - 2{,}2622\sqrt{\frac{25}{8} + \frac{5}{10}} < \mu_1 - \mu_2 < (10-7) + 2{,}2622\sqrt{\frac{25}{8} + \frac{5}{10}}\right] = 0{,}95$$

$$P[-1{,}307 < \mu_1 - \mu_2 < 7{,}307] = 0{,}95$$

Yorum: İki kitle ortalamaları arasındaki farkın 0,95 güven düzeyinde güven aralığı –1,307 ile 7,307'dir.

Bağımlı İki Kitle Ortalamasına İlişkin Hipotez Testleri ve Güven Aralığı

Araştırma yapılacak kitleler, bir özellik yönünden aynı kabul ediliyor ise, bu iki kitleye bağımlı kitleler diyoruz. Örneğin bir şahsın iki farklı zamanda alınan bilgileri ya da cinsiyet özelliği etkin olan bir değerin ölçülmesinde anne kız, baba oğul bağımlı kitleden çekilmiş kabul edilebilir. Tıbbi deneylerde kullanılan kan akrabalığı yüksek deney hayvanlarının oluşturduğu grup da bu özelliktedir. Bir kişinin ameliyat öncesi kan basıncı ile ameliyat sonrası kan basıncı arasında fark olup olmadığı ya da aynı şirketin belli bir özellik yönünden, iki farklı zamanda karşılaştırılması bağımlı grup karşılaştırılmasıdır.

$\mu_d=\mu_1-\mu_2$ olmak üzere kullanılacak hipotezler aşağıda verilmiştir:

a- $H_0:\mu_d\leq 0$	b- $H_0:\mu_d\geq 0$	c- $H_0:\mu_d=0$
$H_A:\mu_d>0$	$H_A:\mu_d<0$	$H_A:\mu_d\neq 0$

Yukarıda verilen hipotezleri test etmek için kullanılacak test istatistiği,

$$t_H=\frac{\bar{d}-\mu_d}{S_d/\sqrt{n}} \qquad (9.11)$$

dir. Burada,

$$d_j=x_{1j}-x_{2j}\ ,\quad \bar{d}=\frac{\sum_{j=1}^{n} d_j}{n}\ ,\quad S_d=\sqrt{\frac{\sum_{j=1}^{n}\left(d_j-\bar{d}\right)^2}{n-1}}$$

olarak hesaplanır. Kurulan hipotezi test ederken hipotezlere göre kararlar aşağıda verilmiştir:

Eğer hipotez a biçiminde kurulmuş ise; $t_H>t_{\alpha;\ n-1}$ ise H_0 hipotezi reddedilir, diğer durumda reddedilemez.

Eğer hipotez b biçiminde kurulmuş ise; $t_H<-t_{\alpha;\ n-1}$ ise H_0 hipotezi reddedilir, diğer durumda reddedilemez.

Eğer hipotez c biçiminde kurulmuş ise; $t_H>t_{\alpha/2;\ n-1}$ ya da $t_H<-t_{\alpha/2;\ n-1}$ ise H_0 hipotezi reddedilir, diğer durumda reddedilemez.

Bağımlı iki grup için güven aralığı,

$$P\left[\bar{d}-t_{\alpha/2,n-1}\frac{S_d}{\sqrt{n}}<\mu_1-\mu_2<\bar{d}+t_{\alpha/2,n-1}\frac{S_d}{\sqrt{n}}\right]=1-\alpha \qquad (9.12)$$

biçimindedir**.**

Örnek 9.9: Bir sigorta şirketi çeşitli pirim ödeme gruplarına göre düzenlenen poliçe sayılarının 11 Eylül saldırısından etkilenip etkilenmediğini araştırmak amacıyla, saldırıdan önce ve sonra poliçe türüne göre satılan poliçe sayılarını incelemiş ve aşağıdaki bulgulara ulaşmıştır.

a- 0,05 anlamlılık düzeyinde poliçe sayılarının 11 Eylül saldırısından etkilenip etkilenmediğini test ediniz.

b- 0,95 güven düzeyinde gruplar arası farkın güven aralığını bulunuz.

Poliçe türü	**Saldırı Öncesi** (X_1)	**Saldırı Sonrası** (X_2)	**d**
1	40	52	-12
2	47	45	2
3	33	51	-18
4	54	60	-6
5	61	58	3
6	39	69	-30
7	42	65	-23
8	47	63	-16
9	58	59	-1
10	50	72	-22

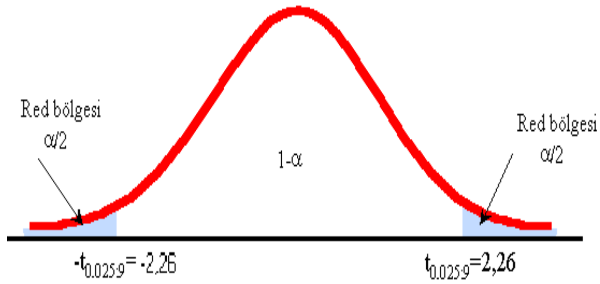
a- $H_0: \mu_d = 0$

$H_A: \mu_d \neq 0$

$$\bar{d} = \frac{\sum_{j=1}^{n} d_j}{n} = \frac{-12+2-18-6+3-30-23-16-1-22}{10} = \frac{-123}{10} = -12{,}3$$

$$S_d = \sqrt{\frac{\sum_{j=1}^{n}\left(d_j - \bar{d}\right)^2}{n-1}} = \sqrt{\frac{(-12-12{,}3)^2 + (2-12{,}3)^2 + (-18-12{,}3)^2 + \ldots + (-22-12{,}3)^2}{10-1}} = 11{,}42$$

$$t_H = \frac{\bar{d} - \mu_d}{S_d / \sqrt{n}} = \frac{-12{,}3 - 0}{11{,}42/\sqrt{10}} = -3{,}88$$



-3,88<-2,26 olduğu için H_0 hipotezi reddedilir.

Yorum: 0,05 anlamlılık düzeyinde, poliçeler saldırıdan etkilenmiştir.

b- $P\left[-12,3-2,26\frac{11,42}{\sqrt{10}}<\mu_1-\mu_2<-12,3+2,26\frac{11,42}{\sqrt{10}}\right]=0,95$

$P[-20,462<\mu_1-\mu_2<-4,138]=0,95$

Yorum: 0,95 güven düzeyinde gruplar arası farkın güven aralığı –20,462 ile –4,138'dir.

9.4. İKİ ORAN ARASINDAKİ FARKA İLİŞKİN HİPOTEZ TESTLERİ VE GÜVEN ARALIĞI

Bazı problemlerde karşılaştırılması istenen bilgi, oranlarla ifade edilen değerler olabilir. Örneğin, siyaset ile ilgilenen kişilerde, üniversite mezunu olanların oranı ile olmayanların oranının karşılaştırılması gibi.

İlgilenilen oranı p_1 olan bir kitleden çekilen n_1 örneklem büyüklüğüne sahip rasgele örneklemin oranı $\hat{p}_1$, ilgilenilen oranı p_2 olan bir kitleden çekilen n_2 örneklem büyüklüğüne sahip rasgele örneklemin oranı $\hat{p}_2$ olarak verilsin. Burada ilgilenilen kitle oranları arasındaki (p_1-p_2) farktır. Bu amaçla test edilebilecek hipotezler aşağıda verilmiştir:

a- H_0: p_1-$p_2 \leq 0$	b- H_0: p_1-$p_2 \geq 0$	c- H_0: p_1-$p_2 = 0$
H_A: p_1-$p_2 > 0$	H_A: p_1-$p_2 < 0$	H_A: p_1-$p_2 \neq 0$

Not: Bu hipotezlerde genel olarak p_1-$p_2 \leq D_0$ ya da p_1-$p_2 \geq D_0$ ya da p_1-$p_2 = D_0$ olarak alınır. Burada kullanılacak olan hipotezler D_0=0 için özelleştirilmiştir.

Yukarıda verilen hipotezlere ilişkin test istatistiği ,

$$Z_H = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\overline{p}(1-\overline{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \tag{9.13}$$

olup,

$$\overline{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} \tag{9.14}$$

dir.

Oranlarla ilgili çalışmalarda örneklem büyüklüğü genellikle 30'dan büyük olduğu için test istatistiğinin standart normal dağılım olduğu kabul edilir.

Eğer hipotez a biçiminde kurulmuş ise; $Z_H > Z_\alpha$ ise H_0 hipotezi reddedilir, diğer durumda reddedilemez.

Eğer hipotez b biçiminde kurulmuş ise; $Z_H < - Z_\alpha$ ise H_0 hipotezi reddedilir, diğer durumda reddedilemez.

Eğer hipotez c biçiminde kurulmuş ise; $Z_H > Z_{\alpha/2}$ ya da $Z_H < -Z_{\alpha/2}$ ise H_0 hipotezi reddedilir, diğer durumda reddedilemez.

İki kitle oranları arasındaki farkın için güven aralığı,

$$P\left[(\hat{p}_1 - \hat{p}_2) - Z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} < p_1 - p_2 < (\hat{p}_1 - \hat{p}_2) + Z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}\right] = 1 - \alpha \tag{9.15}$$

eşitliği kullanılarak elde edilir.

Örnek 9.10: Belirli bir hasta grubundan rasgele seçilen 250 hastaya serum veriliyor, 250 hastaya serum verilmiyor. Serum verilen 75 hasta ve serum verilmeyen 55 hasta iyileşiyor. 0,05 anlamlılık düzeyinde, serumun iyileşme üzerinde etkili olduğu aşağıda verilen hipoteze göre söylenebilir mi?

H_0: $p_1 - p_2 \geq 0$

H_A: $p_1 - p_2 < 0$

n_1=250; n_2=250; $\hat{p}_1$=55/250=0,22; $\hat{p}_2$=75/250=0,30

$$\overline{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} = \frac{250.0,22 + 250.0,30}{250 + 250} = 0,26$$

$$Z_H = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0{,}22 - 0{,}30) - 0}{\sqrt{0{,}26.(1-0{,}26).\left(\frac{1}{250} + \frac{1}{250}\right)}} = -2{,}04$$

Z_α=1,645

-2,04<-1,645 olduğu için H_0 hipotezi reddedilir.

Yorum: 0,05 anlamlılık düzeyinde, serumun etkili olduğunu söylemek mümkündür.

Örnek 9.11: n_1=200 olan bir örnekten $\hat{p}_1 = 0{,}40$ ve n_2=400 olan bir örnekten $\hat{p}_2 = 0{,}15$ olarak hesaplanmıştır.

a- 0,99 güven düzeyinde p_1-p_2 farkı için güven aralığı bulunuz.

b- 0,01 anlamlılık düzeyinde,

H_0: p_1-p_2=0

H_A: p_1-p_2≠0

hipotezini test ediniz.

a-

$$P\left[(0{,}40 - 0{,}15) - 2{,}58\sqrt{\frac{0{,}40.0{,}60}{200} + \frac{0{,}15.0{,}85}{400}} < p_1 - p_2 < (0{,}40 - 0{,}15) + 2{,}58\sqrt{\frac{0{,}40.0{,}60}{200} + \frac{0{,}15.0{,}85}{400}}\right] = 0{,}99$$

$$P[0{,}15 < p_1 - p_2 < 0{,}35] = 0{,}99$$

Yorum: 0,99 güven düzeyinde, p_1-p_2 farkı için güven aralığı 0,15 ile 0,35 arasındadır.

b- p_1-p_2 farkı için bulunan güven aralığı sınırlarına sıfır rakamı girmediği için H_0 hipotezi reddedilir. 0,01 anlamlılık düzeyinde bu iki oran arasında önemli fark vardır.

Örnek 9.12: Bir TV programı seyreden 400 yetişkin ve 600 gençten oluşan bir örnekte 100 yetişkin ve 300 gencin programı beğendiği belirlenmiştir. Programı beğenen yetişkin ve gençlerin oranları arasındaki farkın 0,95 güven düzeyinde güven aralığını bulunuz. Programı beğenme yönünden oranlar arasında fark olup olmadığını test ediniz.

$$\hat{p}_1 = \frac{100}{400} = 0{,}25 \;;\; \hat{p}_2 = \frac{300}{600} = 0{,}5 \quad ; \quad Z_{0,025}=1{,}96$$

a- $$P\left[(0{,}25 - 0{,}50) - 1{,}96\sqrt{\frac{0{,}25.0{,}75}{400} + \frac{0{,}50.0{,}50}{600}} < p_1 - p_2 < (0{,}25 - 0{,}50) + 1{,}96\sqrt{\frac{0{,}25.0{,}75}{400} + \frac{0{,}50.0{,}50}{600}}\right] = 0{,}95$$

$$P[-0{,}31 < p_1 - p_2 < -0{,}19] = 0{,}95$$

Yorum: 0,95 güven düzeyinde, p_1-p_2 farkı için güven aralığı –0,31 ile -0,19 arasındadır.

b- H_0: p_1-p_2=0

H_A: p_1-p_2≠0

p_1-p_2 farkı için bulunan güven aralığı sınırları arasına sıfır girmediği için H_0 reddedilir. İki oran arası fark önemlidir.

PROBLEMLER

1- İki çeşit ürünün dayanıklılığı karşılaştırılmak istenmiştir. Bu amaçla A tip üründen 25 ve B tip üründen 11 tane seçilmiştir. A ve B tip ürünlerin dayanıklılık ortalamaları sırasıyla 204 ve 251 ; varyansları 10 ve 10 olarak elde edildiğine göre 0,05 anlamlılık düzeyinde ortalama dayanıklılık bakımından A ve B tipleri arasında fark olup olmadığını test ediniz.

2- Beslenmemizde kullandığımız etler doğru saklanmadıkları takdirde bakteriler içerir. Etlerde bulunan bakterileri ölçmek için kullanılan çok sayıda yöntem vardır. Burada, bu ölçümlerde kullanılan iki farklı yöntemin varyansları karşılaştırılmak istenmektedir. Ölçümlerin küçük varyansa sahip olması istenen bir özelliktir. Karşılaştırmayı yapmak için birinci yöntem ile 11 ölçüm, ikinci yöntem ile 13 ölçüm yapılmış ve varyanslar sırasıyla 0,025 ve 0,017 olarak bulunmuştur. Bu iki ölçümün varyansları arasında fark olup olmadığını 0,10 anlamlılık düzeyinde test ediniz.

3- Aşağıdaki verilere göre iki grup ortalamaları arasında fark olup olmadığını 0,05 anlamlılık düzeyinde test ediniz.

$n_1 = 4 \quad \overline{X}_1 = 15 \quad S_1^2 = 8$

$n_2 = 5 \quad \overline{X}_2 = 10 \quad S_2^2 = 10$

4- Bulaşık deterjanını plastik kaplara dolduran iki makine göz önüne alınsın. Birinci makineden 5, ikinci makineden 6 plastik kap alınsın. Bu kaplar incelendiğinde, birinci makinenin ortalaması 50 birim ve varyansı 10 birim^2, ikinci makinenin ortalaması 40 birim ve varyansı 12 birim^2 olarak bulunmuştur. 0,05 anlamlılık düzeyinde ortalamalar arasında fark olup olmadığını test ediniz.

5- Bilgileri aşağıda verilen iki grubun kitle ortalamalarının farklı olup olmadığını 0,05 anlamlılık düzeyinde test ediniz.

$n_1 = 9 \qquad \overline{X}_1 = 4 \qquad S_1 = 8$

$n_2 = 15 \qquad \overline{X}_2 = 12 \qquad S_2 = 6$

6- Aşağıda verilen bilgilere göre kız ve erkek öğrencilerin sosyal çalışmaya katılma yüzdeleri arasındaki farkı 0,05 anlamlılık düzeyinde test ediniz.

	Sosyal Çalışmaya Katılan	Toplam
Kız	59	429
Erkek	131	634
Toplam	190	1063

7- Farklı yörelerde bulunan iki köyde yapılan bir araştırma sonuçları aşağıdaki tabloda verilmiştir. Kadınların okur yazar olmama oranına göre köyler arasında farklılık var mıdır? 0,05 anlamlılık düzeyinde test ediniz.

	A Köyü	B Köyü
Okuma-Yazma bilmiyor	20	8
Okuma-Yazma biliyor	111	234

8- Bir işletmede A ve B yöntemleri ile üretilen ürünlerden 100'er birimlik örneklem alındığı varsayılsın. A yöntemiyle üretilen ürünlerden 10 tanesinin, diğer yöntemle üretilen ürünlerden 3 tanesinin standartlara uymadığı gözlendiğine göre,

a- p_1-p_2 (standartlara uymama oranları) farkı için 0,99 güven düzeyinde güven aralığını bulunuz.

b- Elde edilen sonuçtan yararlanarak A ve B yöntemleri arasında (standartlara uymama oranları arasında) fark olduğu 0,01 anlamlılık düzeyinde söylenebilir mi?

9- Grip aşısının etkisini görebilmek için bir çalışma yapılmıştır. Rasgele 3000 kişiye A grip aşısı yapılmış ve 130 kişinin grip olduğu gözlenmiştir. 2500 kişiye de B grip aşısı yapılmış ve 170 kişinin grip olduğu görülmüştür.

a- 0,05 anlamlılık düzeyinde A ve B grip aşıları arasındaki farkın önemli olup olmadığını test ediniz.

b- 0,05 anlamlılık düzeyinde oranlar arası farkın güven aralığını bulunuz.

10- İki farklı ilacın etkisini görebilmek için bir çalışma yapılmıştır. Rasgele 500 kişiye A ilacı verilmiş, 120 kişide etkili olduğu görülmüştür. 250 kişiye de B ilacı verilmiş ve 125 kişide etkili olduğu görülmüştür.

a- 0,05 anlamlılık düzeyinde oranlar arası farkın önemli olup olmadığını test ediniz.

b- 0,05 anlamlılık düzeyinde oranlar arası farkın güven aralığını bulunuz.

11- Aşağıdaki verilere göre ameliyatın tansiyona etkisi var mıdır? 0,05 anlamlılık düzeyinde araştırınız.

Ameliyat öncesi tansiyon değerleri :	11	12	10	14	12	10	9	9	9	9	9	10	11	12	11
Ameliyat sonrası tansiyon değerleri :	10	11	9	12	11	10	9	8	9	8	8	9	10	12	11

12- Albümin miktarı fazla olan 12 hastaya yeni bir ilaç uygulanmış ve ilacın etkisi araştırılmıştır. İlaç öncesi ve sonrası albümin değerleri aşağıdaki gibi verildiğine göre ilacın etkisinin olduğu 0,05 anlamlılık düzeyinde söylenebilir mi?

İlaç öncesi :	139	120	118	119	133	120	122	116	114	127	140	136
İlaç sonrası :	130	115	118	117	130	119	118	114	110	124	136	132

13- Dış ticaret elemanlarından 9 tanesine satış teknikleri ve uluslararası ticarette önemli konuları içeren bir seminer verilmiştir. Seminerin başarısını ölçmek için seminerden önce ve sonra adaylar sınava alınmıştır. Sonuçlar aşağıda verilmiştir. Elde edilen yararın sıfır puandan büyük olduğunu 0,05 anlamlılık düzeyinde test ediniz.

Önce	**Sonra**
64	67
86	87
80	82
82	81
79	85
85	87
83	80
86	84
89	95

14- 170^0 C'lik ısının patateste C vitamini etkisini belirlemek için kızartmadan önce ve sonra C vitamini değerleri ölçülmüştür. 0,05 anlamlılık düzeyinde, ısının C vitamini kaybını yapıp yapmadığını test ediniz.

Önce	21	22	20	19	15	23	25	26	24
Sonra	15	16	17	18	11	20	12	19	18

15- Bir firma yeni bir yoğurt üretimi yapmıştır. Eski yoğurt ile yeni yoğurdun viskositelerini karşılaştırmak için aşağıda verilen sonuçlar alınmıştır. Yoğurtlar viskosite bakımından 0,01 anlamlılık düzeyinde farklı mıdır?

	Yeni Yoğurt	Eski Yoğurt
Ortalama	3000	2500
Standart Sapma	300	150
Gözlem Sayısı	20	10

16- 50'şer kişilik iki grup ilkokul öğrencilerinin A marka şeker için harcadıkları para 2 aylık süre için takip edilmiştir. Birinci grup bu süre içinde ortalama 20 YTL, ikinci grup ortalama 22 YTL harcamıştır. Birinci grupta standart sapma 1,1 YTL, ikinci grupta 0,9 YTL'dir. Bu iki grupta şekerin satışları arasındaki fark 0,05 anlamlılık düzeyinde önemli midir?

17- 8 deneme alanına ekilen belli bir üründen her birim alanda 5,7 standart sapma ile ortalama 94,3 kg ürün alınıyor. Öte yandan 7 deneme alanına ekilen ikinci tür üründen her birim alanda 6,2 standart sapma ile ortalama 85,7 kg ürün alınıyor. 0,05 anlamlılık düzeyinde aşağıdaki hipotezi test ediniz.

H_0: $\mu_1=\mu_2$

H_A: $\mu_1\neq\mu_2$

18- İki tür sigarada bulunan nikotin miktarının varyansları sırasıyla 1,2 mg^2 ve 1,4 mg^2'dır. İkinci tür 40 sigara miktarında bulunan nikotin miktarının ortalaması 24,1 mg ve birinci tür 40 sigarada bulunan nikotin miktarının ortalaması 23,8 mg'dir. 0,01 anlamlılık düzeyinde iki tür sigarada bulunan ortalama nikotin miktarı arasında fark olup olmadığını test ediniz.

19- İki ayrı firmadan alınan 15'er süt örneğinin elektrik iletkenlikleri ölçülmüş, aşağıdaki tablo elde edilmiştir.

I. Firma	$n_1 = 15$	$\overline{X}_1 = 110$	$S_1=5$
II. Firma	$n_2 = 15$	$\overline{X}_2 = 85$	$S_2=4$

a- 0,05 anlamlılık düzeyinde firmaların sütleri elektrik iletkenliği yönünden farklı mıdır?

b- 0,95 güven düzeyinde ortalamalar arası farkın güven aralığını bulunuz.

20- Üniversite sınavına Ankara ve İzmir bölgesinden giren 20'şer kişi rasgele seçilmiş ve aldıkları sayısal puanları sorulmuştur. Ankara bölgesinden katılan öğrencilerin aldıkları puanlara ilişkin ortalama 384 ve varyans 10,85; İzmir bölgesinden katılan öğrencilerin aldıkları puanlara ilişkin ortalama 400 ve varyans 11,23'dür.

a- Ortalamalar arası fark için 0,90 güven düzeyinde güven aralığını bulunuz.

b- 0,05 anlamlılık düzeyinde puan ortalamaları arasında fark olup olmadığını test ediniz.

10. ONUNCU BÖLÜM

10.1. GİRİŞ

Bölüm 9'da bağımlı ya da bağımsız iki grup ortalama ve varyans karşılaştırmalarına ilişkin yöntemler incelenmiştir. Bu bölümde, ikiden çok grup karşılaştırmasında ortalamaların farklılığını test eden varyans analizi yöntemi ele alınacaktır. Varyans analizi; bağımsız ve bağımlı k grup karşılaştırmalarında, tekrarlı denemelerde ve içiçe grupların etkilerini ortaya çıkarmada, birden çok etkenin bulunduğu etkenler arasındaki etkileşimlerin test edilmesinde kullanılır. Ancak, burada sadece tek etkenli k bağımsız grup ortalamalarını test eden tek yönlü varyans analizi anlatılacaktır.

10.2. TEK YÖNLÜ VARYANS ANALİZİ

Tek yönlü varyans analizi, tek etkenli örneğin yalnız sıcaklık ya da yalnız yöntem olduğunda ikiden fazla bağımsız grubun ortalamalarının karşılaştırılmasında kullanılır. Bu bölümde tamamen homojen bir kitleden rasgele oluşturulmuş gruplarda, tek bir etken için, farklı gruplar oluşturan problemler incelenecektir. Örneğin, dört farklı öğretim yönteminin başarısı karşılaştırılmak istensin. Her grupta 10'ar öğrencinin bulunması uygun görülmüş olsun. Bu durumda 40 öğrencinin eşit özelliklere sahip olduğu varsayılsın ve rasgelelik kurallarına uyarak bunlar dört gruba ayrılsın. Her gruba rasgele seçimle farklı bir eğitim yöntemi uygulansın. Oluşturulan bu yapıda, sonuçların karşılaştırılmasında varyans analizi kullanılabilir. Burada öğretim yöntemi etken, farklı öğretim yöntemleri de gruplardır.

Üç farklı araba markasının kilometre (km) başına yakıt tüketimi araştırılıyor. Araştırmayı gerçekleştirmek için yalnız araba markalarının farklılığının dışında, tüm şartların eşit olduğu bir deney hazırlanmalıdır. Bu eşitlik kullanılan yakıt, kat edilen yol, güzergah, araba yaşı gibi yakıt tüketimini etkileyecek tüm etkenler için sağlanmalıdır. Her araba markasından 7 deneme yapılması uygun bulunmuştur. Ancak deney sırasında üçüncü marka araba bir kaza nedeni ile deneyden çıkarılmıştır. Aynı anda eşit şartlarda deneye alınan arabaların 10 km başına yakıt tüketimleri litre olarak Tablo 10.1'de verilmiştir. Yakıt tüketimi yönünden araba markaları arasında önemli bir fark var olup olmadığı araştırılmaktadır.

Tablo 10.1 Araba Markalarına İlişkin Yakıt Tüketimleri

(A) Marka Araba	(B) Marka Araba	(C) Marka Araba
1,52	1,96	1,5
1,39	2,00	1,62
1,60	1,85	1,52
0,96	1,80	0,97
1,57	1,20	0,98
1,80	1,50	0,99
0,99	1,98	--
Toplam 9,83	12,29	7,58

Tablo 10.1'de verilen araba markalarının 10 km için yakıt tüketim ortalamaları sırasıyla;

A marka araba için $= \frac{9{,}83}{7} = 1{,}40$

B marka araba için $= \frac{12{,}29}{7} = 1{,}75$

C marka araba için $= \frac{7{,}58}{6} = 1{,}26$

dir. Burada test edilecek nokta bulunan bu ortalamalar arasındaki farkların rasgelelik sınırları içinde olup olmayacağıdır. Bu ortalamalar birbirlerinden ne kadar farklı ise, o kadar çok aynı kitleden çekilmeme olasılığına sahip olacaklardır. Bu bölümde amaç örneklem ortalamalarının farklılığının araştırılmasıdır. Kullanılacak test için veriler ile ilgili dağılımın aşağıdaki şartları sağlaması gerekir:

1- Gruplar varyansları eşit olan kitlelerden çekilmiş olmalıdır.
2- Verilerin normal dağılım özelliği göstermeleri (Bu şart için uygulamada, veriler aralıklı ya da orantılı ölçek türü ile ölçülmüş olmalılar) gerekmektedir.

Varyans analizi problemlerinde önemli noktalardan biri de, farklılığı araştırılacak grupların yani etkenin düzeyleri için farklılık araştırılıyor ise, bu grupların (düzeylerin) nasıl oluşturulduğudur. Aşağıda bu konuda açıklama verilecektir.

k tane kitleden n_1, n_2, …, n_k gözlemli bağımsız rasgele örneklemlerin alınmış olduğu varsayılsın ve örneklemler yukarıda verilen varsayımları sağlasın. Farklılıkları incelenecek olan grupların (etkenin) düzeylerinin seçimi iki şekilde olabilir. Eğer farklılığı incelenecek olan gruplar araştırıcı için özellikle belirlenmiş ise örneğin cinsiyetler grupları oluşturuyorsa ya da belli yaş grupları arasında incelenen özellik için fark aranıyorsa ya da A, B, C gibi belli markalar karşılaştırılıyor ise, bu grupların çekildikleri kitlelerin ortalamaları sırasıyla $\mu_1, \mu_2, ..., \mu_k$ gibi kesin değerlerdir. Bu tür varyans analizi problemlerine özel seçilmiş modeller denir. Bu ortalamaların eşitliği;

H_o: $\mu_1 = \mu_2 = ... = \mu_k$ (Ortalamalar arasında fark yoktur.)

$H_{A:}$ En az bir ortalama diğerinden farklıdır (10.1)

hipotezinin testi ile araştırılır. Eğer grupları belirleyen özellikler çok farklılık gösteriyor ve bu farklı durumlar arasından k tane grup rasgele belirlenip, bu gruplar için sırasıyla $n_1, n_2, ..., n_k$ birim rasgele çekilir ise varyans analizi problemi rasgele seçimli model adını alır. Örneğin, yaşların bir TV programını tercihte etkisi araştırılmak istensin. Bu programı seyretme yaşı 15-50 olsun. Bu aralıkta genel karar verilecek ise, 15-50 yaş gruplarından rasgele k sayıda yaş grubu belirlenir ve araştırma o gruplar üzerinden yapılır. Bu durumlarda örneklem ortalamalarının birer rasgele değişken olduğu kabul edilir, yani $\mu_1, \mu_2, ..., \mu_k$ gibi kesin değerler yoktur. Örneklemlerin ortalamaları bir raslantı dağılımı gösterir. Bu dağılımın normal dağılım gösterdiği varsayılır. Dağılıma ait varyans σ_G^2'dir. Böyle problemlerde gruplar arasındaki farklılık gruplararası varyansın (σ_G^2) sıfıra eşit olup olmadığının incelenmesi ile test edilir. İlgili hipotez,

$$H_0 : \sigma_G^2 = 0$$
$$H_A : \sigma_G^2 > 0 \qquad (10.2)$$

dir. Testler için tek yönlü varyans analizi kullanılır. Tek yönlü varyans analizini ve yapılacak işlemleri genel olarak anlatabilmek için k tane örneklemi karşılaştıracak genel bir yapı ele alınsın. Test için kullanılacak kareler toplamları ve varyanslar Tablo 10.2 üzerinden yorumlanacaktır.

Kareler Toplamı

Tablo 10.2'de k tane grup, X_{ij} i.grubun j. gözleminin değeri ile gösterilmiştir.

Tablo 10.2 k Tane Grup İçin Veri Düzeni

Kitle	1	2		i		k
Örneklem Birimleri	X_{11}	X_{21}	…	X_{i1}	…	X_{k1}
	X_{12}	X_{22}		X_{i2}		X_{k2}
	.	.		.		.
	X_{1n_1}	X_{2n_2}		X_{in_i}		X_{kn_k}
Örneklem Birimlerinin Toplamı	$T_{1.}$	$T_{2.}$	…	$T_{i.}$	…	$T_{k.}$
Örneklem Ortalaması	$\overline{X}_{1.}$	$\overline{X}_{2.}$		$\overline{X}_{i.}$		$\overline{X}_{k.}$

Tablo 10.2'de, gözlem sayılarının toplamı N; n_1 birinci gruptaki gözlem sayısı, n_2 ikinci gruptaki gözlem sayısı,..,n_k k. gruptaki gözlem sayısı olmak üzere $n_1 + n_2 + ... + n_k = N$; i. gruptaki gözlem değerlerinin toplamı, $T_{i.} = \sum_{j=1}^{n_i} X_{ij}$; tüm gözlemlerin genel toplamı, $T_{..} = \sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}$; her bir gruba ilişkin ortalama, $\overline{X}_{i.} = \frac{T_{i.}}{n_i} = \frac{\sum_{j=1}^{n_i} X_{ij}}{n_i}$ ve genel ortalama, $\overline{\overline{X}}_{..} = \frac{T_{..}}{N}$'dır. Herhangi bir gözlemin değeri olan X_{ij}'nin genel ortalamadan ayrılışı iki parçaya ayrılabilir. Bunlar;

$$\left(X_{ij} - \overline{\overline{X}}_{..}\right) = \left(\overline{X}_{i.} - \overline{\overline{X}}_{..}\right) + \left(X_{ij} - \overline{X}_{i.}\right) \tag{10.3}$$

dir. Eşitlik (10.3)'ün sağ tarafında bulunan ilk terim grup ortalamaları arasındaki fark ile büyür. Araştırılacak olan da bu farkın büyüklüğüdür. Farka ait varyansları bulabilmek için eşitliğin her iki tarafının önce kareleri alınır. Sonrada her iki indis üzerinden toplama yapılır. Sonuçta:

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(X_{ij} - \overline{\overline{X}}_{..}\right)^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\overline{X}_{i.} - \overline{\overline{X}}_{..}\right)^2 + \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\overline{X}_{ij} - \overline{X}_{i}\right)^2 \tag{10.4}$$

özdeşliği elde edilir. Burada:

Her bir gözlemin genel ortalamadan ayrılışlarına ait kareler toplamı:

$$TKT = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(X_{ij} - \overline{\overline{X}}_{..}\right)^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}^2 - \frac{T_{..}^2}{N} \tag{10.5}$$

dır (TKT). Toplam kareler toplamı ya da genel kareler toplamı olarak adlandırılır.

Her bir gözlemin kendi grup ortalamasından olan ayrılışlarının kareler toplamı:

$$HKT=\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(X_{ij}-\overline{X}_{i.}\right)^2=\sum_{i=1}^{k}\sum_{j=1}^{n_i}X_{ij}^{\ 2}-\sum_{i=1}^{k}\frac{T_{i.}^2}{n_i} \tag{10.6}$$

hata kareler toplamıdır (HKT).

Her grubun ortalamasının genel ortalamadan ayrılışlarının kareler toplamı:

$$GKT=\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\overline{X}_{i.}-\overline{\overline{X}}_{..}\right)^2=\sum_{i=1}^{k}\frac{T_{i.}^2}{n_i}-\frac{T_{..}^2}{N} \tag{10.7}$$

gruplar arası kareler toplamıdır (GKT).

Bu kareler toplamları kendi serbestlik derecelerine bölünerek testte kullanılacak olan varyanslar elde edilir. TKT, HKT ve GKT'ye ilişkin serbestlik dereceleri (sd) sırasıyla aşağıda verilmiştir:

$TKT_{sd}=N-1$

$HKT_{sd}=N-k$

$GKT_{sd}=k-1$

Yukarıda açıklanan bilgilerin bir tabloda yer alması kolaylık sağlar. Tablo 10.3 bir varyans analizi tablosudur. Tablo 10.3'de yer alan GKO=GKT/(k-1) ve HKO=HKT/(N-k)'dır.

Tablo 10.3 Varyans Analizi Tablosu

Varyans Kaynağı	**Serbestlik Derecesi**	**Kareler Toplamı**	**Kareler Ortalaması**
Gruplar	k – 1	GKT	GKO
Hata	N – k	HKT	HKO
Toplam	N – 1	TKT	

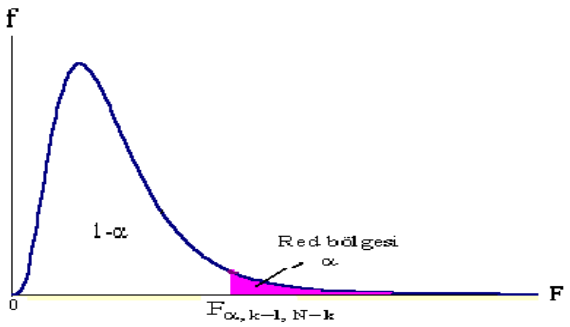
Varyans analizi tablosunda bulunan kareler toplamları arasında Eşitlik (10.4)'de verilen

$$TKT=GKT+HKT \tag{10.8}$$

özdeşliği vardır. Aynı özellik serbestlik dereceleri arasında da vardır. Varyans analizi, genel kareler toplamında gruplar arası kareler toplamı ve hata kareler toplamının paylarını inceler. Bu da GKO ile HKO'nın oranı ile bulunur. Bu iki varyansın oranı (k-1) ve (N-k) serbestlik dereceli F dağılımı gösterir ve aşağıda verildiği biçimde elde edilir.

$$F_H = \frac{GKO}{HKO} \qquad (10.9)$$

Eşitlik (10.9) ile bulunan F_H değeri, belirlenen α anlamlılık düzeyinde, GKO'nun serbestlik derecesi olan k–1 ve HKO'nun serbestlik derecesi olan N–k serbestlik derecesinde F tablo değeri ile karşılaştırılır. Şekil 10.1'den de görüldüğü gibi, eğer, $F_H \leq F_{\alpha, k-1, N-k}$ ise Eşitlik (10.1) ve Eşitlik (10.2) de verilen H_0 hipotezleri reddedilemez, $F_H > F_{\alpha, k-1, N-k}$ ise H_0 hipotezleri reddedilir. Özel seçimli modelde H_0 hipotezinin reddedilmesi grup ortalamalarından en az birinin diğerlerinden farklı olduğunu ifade eder. Bu durumda, hangi grup ortalamasının farklı olduğu bilinmek istenebilir. Farklı olan grubu bulabilmek için bazı testler önerilmiştir. k bağımsız grubun ortalamalarının birbirinden farklılığını araştırmak için geliştirilen bu testlere çoklu karşılaştırma testleri (son test) adı verilir. Çoklu karşılaştırmalarla ilgili birçok teknik olmasına rağmen, burada bu testlerden Tukey-Kramer testi ve Newman-Keuls testleri verilecektir. Rasgele seçimli modelde H_0 hipotezi reddedildiğinde σ_G^2'nin genel varyansdaki yüzdesi aranır. Daha ayrıntılı bilgi için bakınız (Muluk, Z. vd.,1982).



Şekil 10.1 Varyans Analizi İçin Kritik Bölge

Özel Seçimli Varyans Analizi Problemlerinde Çoklu Karşılaştırmalar

Tukey-Kramer Testi

Grup varyansları eşit kabul edilen k bağımsız grubun ortalamalarının birbirinden farklılığını araştırmak için kullanılan testtir. Bu testi gruplarda farklı gözlem olması durumunda kullanabiliriz. Testi kullanmak için ilk yapılması gereken her bir grup ortalamasının diğer gruplardan olan farklarının bulunmasıdır. Daha sonra bu farklar aşağıda verilen Tukey-Kramer kritik değeri (TK_{kd}) ile karşılaştırılarak farklılığı yaratan grup elde edilebilir.

$$TK_{kd} = q_{\alpha}\sqrt{\frac{HKO}{2}\left(\frac{1}{n_i}+\frac{1}{n_{i'}}\right)} \qquad (10.10)$$

q_{α}: α anlamlılık düzeyinde Ek 8'de verilen tablodan yararlanarak, k bağımsız grup sayısı sütundan ve N-k serbestlik derecesi satırlarının kesişim noktasındaki değer q_{α}'dır.

n_i ve $n_{i'}$: Sırasıyla i. ve i'. grubun gözlem sayısı, $i = i'$ 'dür.

HKO: Tablo 10.3'de verilen hata kareler ortalamasıdır.

Eğer i. ve i'. grup ortalama farkı, i. ve i'. gruba ilişkin TK_{kd} değerinden büyük ise, i. ve i'. grup ortalamalarının farklı olduğu, eğer i. ve i'. grup ortalama farkı, i. ve i'. gruba ilişkin TK_{kd} değerinden küçük ise, i. ve i'. grup ortalamalarının farklı olmadığı söylenir. Konuya ilişkin bir örnek Örnek 10.2'de verilmiştir.

Newman-Keuls Testi

Farklı olan grubun bulunması için tüm ikili karşılaştırmaları test eden Newman-Keuls testinde aşağıdaki adımlar uygulanır:

1. k bağımsız gruba ilişkin ortalamaların büyükten küçüğe doğru sıralanması ile bir matrisin sütunları ve satırlarının elemanları oluşturulur. Bu matris
   $\overline{X}_A > \overline{X}_B > \overline{X}_C > \overline{X}_D \ldots$

olmak üzere Tablo 10.4'de verilmiştir:

Tablo 10.4 Sıralanmış Matris

	$\overline{X}_A$	$\overline{X}_B$	$\overline{X}_C$	$\overline{X}_D$	. . .
$\overline{X}_A$					
$\overline{X}_B$					
$\overline{X}_C$					
$\overline{X}_D$					
. . .					

2. Varyans analizi tablosundaki hata kareler ortalaması ile serbestlik derecesi bulunur.
3. Grup ortalamalarının standart hataları Eşitlik (10.11)'e göre bulunur:

$$S_{\overline{X}} = \sqrt{\frac{HKO}{\overline{n}}} \tag{10.11}$$

Burada, $\overline{n} = \dfrac{N^2 - \sum_{i=1}^{k} n_i^2}{(k-1)N}$ 'dir. $n_1=n_2=...=n_k=n$ olduğunda $\overline{n} = n$ olur.

4. Birinci adımda oluşturulan matrisin elemanları için; ortalamalar karşılaştırılırken büyükten küçüğe yapılan sıralamada yan yana gelen grup sayısı p=2,3,4,...k dikkate alınarak, hata serbestlik derecesi ve α anlamlılık düzeyinde $q_2,q_3,...,q_k$ değerleri Ek 8'deki tablodan alınır. q_2 yan yana iki grup için kullanılır. Bu değer Ek 8'deki tabloda HKT_{sd} ilk sütundan, grup sayısı 2 ilk satırdan alınarak bakılır. q_3 yan yana 3 grup içindir. İlk satırdan 3 ilk sütundan HKT_{sd} alınarak bakılır ve bu şekilde devam edilir.
5. Dördüncü adımda Ek 8 tablosundan bulunan q değerleri ile $S_{\overline{X}}$ çarpılarak Tablo 10.5'de verilen en küçük önem genişlikler matrisi bulunur.

Tablo 10.5 En Küçük Önem Genişlik Matrisi

	$\overline{X}_A$	$\overline{X}_B$	$\overline{X}_C$	$\overline{X}_D$	. . .
$\overline{X}_A$	0				
$\overline{X}_B$	$q_2 \times S_{\overline{X}}$	0			
$\overline{X}_C$	$q_3 \times S_{\overline{X}}$	$q_2 \times S_{\overline{X}}$	0		
$\overline{X}_D$	$q_4 \times S_{\overline{X}}$	$q_3 \times S_{\overline{X}}$	$q_2 \times S_{\overline{X}}$	0	
. . .	. . .	. . .	. . .	. . .	. . .

6. Birinci adımda oluşturulan matrisin elemanları bir kez de satır, sütun başlarında bulunan ortalamaların birbirinden çıkartılması ile oluşturulur. Bu şekilde Tablo 10.6'da verilen fark matrisi elde edilir.

Tablo 10.6 Fark Matrisi

	$\overline{X}_A$	$\overline{X}_B$	$\overline{X}_C$	$\overline{X}_D$	. . .
$\overline{X}_A$	0				
$\overline{X}_B$	$(\overline{X}_A - \overline{X}_B)$	0			
$\overline{X}_C$	$(\overline{X}_A - \overline{X}_C)$	$(\overline{X}_B - \overline{X}_C)$	0		
$\overline{X}_D$	$(\overline{X}_A - \overline{X}_D)$	$(\overline{X}_B - \overline{X}_D)$	$(\overline{X}_C - \overline{X}_D)$	0	
. . .	. . .	. . .	. . .	. . .	0

7. Beşinci adımda elde edilen en küçük önemli genişlikler matrisi ile altıncı adımda elde edilen fark matrisinin elemanları birebir karşılaştırılır. Örneğin Tablo 10.5'in 1. sütun, 2. satır elemanı $q_2 \times S_{\overline{X}}$ ile Tablo 10.6'nın 1. sütun, 2. satır elemanı $(\overline{X}_A - \overline{X}_B)$ karşılaştırılır. $q_2 \times S_{\overline{X}} > (\overline{X}_A - \overline{X}_B)$ ise bu iki ortama arasındaki farkın önemli olmadığı, $q_2 \times S_{\overline{X}} < (\overline{X}_A - \overline{X}_B)$ ise α anlamlılık düzeyinde bu iki ortama arasındaki farkın önemli olduğu söylenebilir.

Örnek 10.1: Yukarıda verilen bilgiler doğrultusunda Tablo 10.1'de verileri olan yakıt tüketim probleminde araba markaları arasında yakıt tüketimleri açısından fark olup olmadığını 0,05 anlamlılık düzeyinde test ediniz. Gruplar arasında fark var ise hangi grubun farklı olduğunu Newman-Keuls testi ile bulunuz. Araba markaları bir çok marka içinden rasgele seçilmedikleri için model özel seçimlidir. Hipotez;

H_o: $\mu_1 = \mu_2 = \mu_3$

H_A: En az bir ortalama diğerinden farklıdır.

$T_{1.}$=9,83, $T_{2.}$=12,29, $T_{3.}$=7,58, $T_{..}$=29,7

$\sum\sum X_{ij}^2 = 46,5898$ N = 7 + 7 + 6 = 20

$$TKT = 46,5898 - \frac{(29,7)^2}{20} = 2,4853$$

$$HKT = 46,5898 - \left[\frac{(9,83)^2}{7} + \frac{(12,29)^2}{7} + \frac{(7,58)^2}{6}\right] = 1,632$$

$$GKT = 44,9578 - 44,1045 = 0,8533$$

Tablo 10.7 Araba Yakıt Tüketimlerine İlişkin Varyans Analiz Tablosu

Varyans Kaynağı	**Serbestlik Derecesi**	**Kareler Toplamı**	**Kareler Ortalaması**
Gruplar	2	0,8533	0,4266
Hata	17	1,632	0,096
Toplam	19	2,4853	

$$F_H = \frac{GKO}{HKO} = \frac{0,4266}{0,096} = 4,44 \qquad F_{0,05,\,2,\,17} = 3,592$$

4,44>3,592 olduğundan H_o reddedilir.

Yorum: 0,05 anlamlılık düzeyinde, en az bir ortalama diğerinden farklıdır.

Farklı olan grubu bulabilmek için Newman-Keuls testi yapılabilir.

Araba marka tüketim ortalamaları büyükten küçüğe aşağıda verilmiştir:

$$\overline{X}_B = 1,75,\ \overline{X}_A = 1,40,\ \overline{X}_C = 1,26$$

$$\overline{n} = \frac{20^2 - (7^2 + 7^2 + 6^2)}{(3-1).20} = 6,65 \approx 7$$

$$S_{\overline{X}} = \sqrt{\frac{0,096}{7}} = 0,1171$$

Ek 8 tablosundan α=0,05, $q_{2,17}$ =2,98, $q_{3,17}$ =3,62 olup, verilere ilişkin en küçük önem genişliği tablosu,

	$\overline{X}_B$	$\overline{X}_A$	$\overline{X}_C$
$\overline{X}_B$	0		
$\overline{X}_A$	0,1171×2,98=0,34	0	
$\overline{X}_C$	0,1171× 3,62=0,42	0,1171×2,98=0,34	0

Verilere ilişkin fark matrisi tablosu aşağıda verilmiştir.

	$\overline{X}_B$	$\overline{X}_A$	$\overline{X}_C$
$\overline{X}_B$	0		
$\overline{X}_A$	1,75-1,40=0,35	0	
$\overline{X}_C$	1,75-1,26=0,49	1,40-1,26=0,14	0

0,34<0,35 olduğundan, B ile A araba markalarının yakıt tüketim ortalamalarının farkının önemli olduğu söylenebilir. 0,42<0,49 olduğundan, B ile C araba markalarının yakıt tüketim ortalamalarının farkının da önemli olduğu söylenebilir. 0,34>0,14 olduğundan, A ile C araba markalarının yakıt tüketim ortalamalarının farkının önemli olmadığı söylenebilir. 0,05 anlamlılık düzeyinde, B markası diğer iki markaya göre daha fazla yakıt tüketmektedir.

Örnek 10.2: Bir aracı kurumda üç farklı portföyün aylık getirisinin aynı olup olmadığı araştırılmak istenmektedir. Bu amaçla portföylerin haftalık getirileri aşağıdaki tabloda verilmiştir. 0,05 anlamlılık düzeyinde getirilerin önemli ölçüde farklı olup olmadığını test ediniz. Gruplardan birisi diğerlerinden farklı ise Tukey-Kramer testini kullanarak bulunuz.

Hafta	Portföy		
	I	II	III
1	0,6	1,2	1,5
2	1,1	1,0	1,2
3	1,0	0,9	1,3
4	0,7	0,9	1,2

H_o: $\mu_1 = \mu_2 = \mu_3$

H_A: En az bir ortalama diğerinden farklıdır.

$$\overline{X}_{i.} = \frac{\sum_{J=1}^{n_i} X_{ij}}{n_i} \Rightarrow \overline{X}_{1.} = 0{,}85;\ \overline{X}_{2.} = 1;\ \overline{X}_{3.} = 1{,}3\ ;\ \overline{\overline{X}}_{..} = \frac{\sum_{j=1}^{k}\sum_{i=1}^{n_j} X_{ij}}{N} = \frac{12{,}6}{12} = 1{,}05$$

$T_{1.} = 3{,}4$ $\quad T_{2.} = 4{,}0$ $\quad T_{3.} = 5{,}2$ $\quad T_{..} = 12{,}6$

$$GKT = \left(\frac{3{,}4^2}{4} + \frac{4{,}0^2}{4} + \frac{5{,}2^2}{4}\right) - \frac{12{,}6^2}{12} = 13{,}65 - 13{,}23 = 0{,}42$$

$TKT = 13{,}94 - 13{,}23 = 0{,}71$

$HKT = TKT - GKT = 0{,}71 - 0{,}42 = 0{,}29$

$$GKO = \frac{GKT}{k-1} = \frac{0{,}42}{3-1} = 0{,}21$$

$$HKO = \frac{HKT}{N-k} = \frac{0{,}29}{12-3} = 0{,}03222$$

Tablo 10.8 Portföylerin Haftalık Getirilerine İlişkin Varyans Analiz Tablosu

Varyans Kaynağı	Serbestlik Derecesi	Kareler Toplamı	Kareler Ortalaması
Portföyler	2	0,42	0,21
Hata	9	0,29	0,032
Toplam	11	0,71	

$$F_H = \frac{GKO}{HKO} = \frac{0,21}{0,032} = 6,517 \qquad F_{0,05,2,9} = 4,2565$$

6,51724 >4,2565 olduğundan H_o reddedilir.

Yorum: 0,05 anlamlılık düzeyinde, en az bir portföy getirisinin diğerinden farklı olduğu söylenebilir.

Farklı olan grubu bulabilmek için Tukey-Kramer testi uygulanabilir.

HKO=0,032 ve k=3 ve N-k=9 serbestlik derecesindeki $q_{0,05}$=3,95'dir. $n_1=n_2=n_3=4$; olduğundan (denek sayıları eşit) tek bir TK kritik değeri bulunur.

$$TK_{kd} = q_\alpha \sqrt{\frac{HKO}{2}\left(\frac{1}{n_i}+\frac{1}{n_i'}\right)} = 3,95\sqrt{\frac{0,032}{2}\left(\frac{1}{4}+\frac{1}{4}\right)} = 0,35$$

$\left|\overline{X}_1 - \overline{X}_2\right| = |0,85 - 1| = 0,15 < 0,35$ olduğundan, 1. ve 2. grup ortalamaları farklı değil

$\left|\overline{X}_2 - \overline{X}_3\right| = |1 - 1,3| = 0,3 < 0,35$ olduğundan, 2. ve 3. grup ortalamaları farklı değil

$\left|\overline{X}_1 - \overline{X}_3\right| = |0,85 - 1,3| = 0,45 > 0,35$ olduğundan, 1. ve 3. grup ortalamaları farklıdır.

Yorum: 1. ve 3. grup birbirlerinden farklı, 2. grup hem 1. hem de 3. grupta kabul edilebilir.

Örnek 10.3: Bir bankanın üç şubesindeki vadesiz mevduat hesaplarının işlem sayıları çeşitli günlerde belirlenmiş ve aşağıdaki tabloda verilmiştir. Şubeler arasında vadesiz işlem yükü açısından farklılık olup olmadığını 0,05 anlamlılık düzeyinde test ediniz. Gruplar arasında fark var ise farklı olan grubu bulunuz.

Şubeler	**A**	61	53	61	58	50	49	55	58	59		
	B	42	46	40	38	44	45	42	40	52	54	49
	C	56	52	57	48	55	57	53	52			

H_o: $\mu_1 = \mu_2 = \mu_3$

H_A: En az bir ortalama diğerinden farklıdır.

$N = 28$, $n_1 = 9$, $n_2 = 11$, $n_3 = 8$

$$\overline{X}_i = \frac{\sum_{j=1}^{n_i} X_{ij}}{n_i} = \frac{504}{9} \Rightarrow \overline{X}_{1.} = 56 \; ; \; \overline{X}_{2.} = \frac{492}{11} = 44{,}72 \; ; \; \overline{X}_{3.} = \frac{430}{8} = 53{,}75$$

$$\overline{\overline{X}}_{..} = \frac{\sum_{j=1}^{k}\sum_{i=1}^{n_j} X_{ij}}{N} = \frac{1426}{28} = 50{,}92$$

$$GKT = \left(\frac{504^2}{9} + \frac{492^2}{11} + \frac{430^2}{8}\right) - \frac{1426^2}{28} = 718{,}17$$

$$TKT = 73836 - \frac{1426^2}{28} = 1211{,}85$$

$$HKT = TKT - GKT = 493{,}68$$

$$GKO = \frac{GKT}{k-1} = \frac{718{,}1753}{3-1} = 359{,}08$$

$$HKO = \frac{HKT}{N-k} = \frac{493{,}6818}{28-3} = 19{,}74$$

Tablo 10.9 İşlem Sayılarına İlişkin Varyans Analiz Tablosu

Varyans Kaynağı	**Serbestlik Derecesi**	**Kareler Toplamı**	**Kareler Ortalaması**
Şubeler	2	718,17	359,08
Hata	25	493,68	19,74
Toplam	27	1211,85	

$$F_H = \frac{GKO}{HKO} = 18{,}18 \quad F_{0,05,2,25} = 3{,}3852$$

$18{,}18 > 3{,}3852$ olduğundan H_o reddedilir.

Yorum: 0,05 anlamlılık düzeyinde şubeler arasında önemli fark olduğu söylenebilir.

Farklı olan grupları bulabilmek için Tukey-Kramer testi uygulanabilir.

HKO=19,74 ve k=3 ve N-k=25 serbestlik derecesindeki $q_{0,05} \approx 3{,}52$'dir.

1. ve 2. grup için, $TK_{kd} = 3{,}52\sqrt{\frac{19{,}74}{2}\left(\frac{1}{9} + \frac{1}{11}\right)} = 4{,}97$

2. ve 3. grup için, $TK_{kd} = 3{,}52\sqrt{\frac{19{,}74}{2}\left(\frac{1}{11} + \frac{1}{8}\right)} = 5{,}13$

1. ve 3. grup için, $TK_{kd} = 3{,}52\sqrt{\frac{19{,}74}{2}\left(\frac{1}{9}+\frac{1}{8}\right)} = 5{,}37$

$\left|\overline{X}_1 - \overline{X}_2\right| = |56 - 44{,}72| = 11{,}28 > 4{,}97$ olduğundan, 1. ve 2. grup ortalamaları farklıdır.

$\left|\overline{X}_2 - \overline{X}_3\right| = |44{,}72 - 53{,}75| = 9{,}03 > 5{,}13$ olduğundan, 2. ve 3. grup ortalamaları farklıdır.

$\left|\overline{X}_1 - \overline{X}_3\right| = |56 - 53{,}75| = 2{,}25 < 5{,}37$ olduğundan, 1. ve 3. grup ortalamaları farklı değildir.

Yorum: 1. ve 3. grup ortalamaları aynı olup, bu iki grubun aynı kitleden çekildiği 2. grubun bunlardan farklı olduğu söylenebilir.

Örnek 10.4: Bisküvi ihracatı yapan bir firma bisküvilerinin bazı özelliklerini belirlemek ve bu özellikler yönünden değişiminin önemli olup olmadığını saptamak için ihraç edilecek ürünlerden rasgele 5 koli seçip; ağırlık, tuz oranı, nemlilik vb. özellikleri belirlemiştir. Aşağıda bu bilgilerden ağırlıklar verilmiştir. 0,05 anlamlılık düzeyinde ihraç edilecek ürünlerin ağırlık yönünden değişimleri önemli midir?

Rasgele Seçilen Koliler

1	2	3	4	5
150	148	149	148	149
150	149	149	150	148
150,2	148	149	150	149
150,3	148,2	150	151	149
151	148	150	151,2	149
149,8	150,1	151	149,2	150
149,5	150	150,4	151	149
149	150,1	150	150,1	148
150	149	149,2	150	148
151,1	149,2	149,9	149,6	148,7

$H_0 : \sigma_G^2 = 0$

$H_A : \sigma_G^2 > 0$

$T_{1.}=1500{,}9$ $T_{2.}=1489{,}6$ $T_{3.}=1497{,}5$ $T_{4.}=1500{,}1$ $T_{5.}=1487{,}7$

Tablo 10.10 Kolilere İlişkin Varyans Analiz Tablosu

Varyans Kaynağı	**Serbestlik Derecesi**	**Kareler Toplamı**	**Kareler Ortalaması**
Koliler	4	14,939	3,735
Hata	45	26,628	0,592
Toplam	49	41,567	

$$F_H = \frac{GKO}{HKO} = 6{,}312 \quad F_{0,05,4,45} = 2{,}575$$

6,312>2,585 olduğundan H_o reddedilir.

Yorum: 0,05 anlamlılık düzeyinde ihraç edilecek bisküvi partisinde koliler arasında değişimin önemli olduğu söylenebilir.

PROBLEMLER

1- Varyans analizi modelinin özel seçimli olması ile rasgele seçimli olması ayırımını açıklayınız.

2- Varyans analizi modeli özel seçimli ise hipotezlerin nasıl olacağını, yorumun nasıl yapılacağını yazınız.

3- Varyans analizi modeli rasgele seçimli ise hipotezlerin nasıl olacağını, yorumun nasıl yapılacağını yazınız.

4-Bir tekstil fabrikasında çok sayıda dokuma tezgahı vardır. Her tezgahın saniyede aynı uzunlukta kumaş dokuduğu düşünülüyor. Rasgele seçilen 3 tezgahın farklı zamanlarda çıktıları ölçülüyor. Aşağıda verilen tablodaki eksik bilgileri tamamlayınız ve 0,05 anlamlılık düzeyinde kitle ortalamaları arasında fark olup olmadığını araştırınız.

VK	sd	KT	KO
Kumaşlar	-	0,634	-
Hata	12	-	0,182
Toplam	-	2,818	

5- Bir mağazada çalışan 3 grubun rasgele olarak belirlenen 5 gün içinde sattıkları parça sayılarına ilişkin verilen varyans çözümleme tablosu aşağıdaki biçimde oluşturulmuştur. Eksik yerleri tamamlayınız, gruplar arası fark kontrolüne ilişkin hipotezi kurunuz ve 0,05 anlamlılık düzeyinde testi yapınız.

VK	sd	KT	KO
Gruplar	-	240,94	-
Hata	-	-	80,53
Toplam	-	1207,34	

6- Üç ambalaj türü yiyecek maddelerinin saklanmasında denenmiş her tip ambalaj ile 10 örnek sarılmış ve belirli süre sonra nem kaybı ölçülmüştür. Aşağıdaki bilgilerden yararlanarak varyans çözümleme tablosunu oluşturunuz. Paketlemeler arasında 0,05 anlamlılık düzeyinde fark olup olmadığını test ediniz.

	Ambalaj Türleri		
	A	B	C
Nem Kaybı Ortalaması	200	250	300
$\sum_{i=1}^{10}\left(X_{ij}-\overline{X}_{.i}\right)^2$	30	40	20

7- Aşağıda verilen bilgilere göre, varyans çözümleme tablosunu düzenleyiniz. 0,05 anlamlılık düzeyinde denemeler arasında fark olup olmadığını araştırınız.

	Gruplar			
	A	B	C	
Ortalaması	4	3	6	
Gözlem Sayısı	6	6	6	HKT=8

8- Aynı malı üreten üç farklı firmadan 5'er örnek alınıp istenmeyen madde oranları saptanmıştır. İstenmeyen madde bakımından firmalar arasında 0,05 anlamlılık düzeyinde fark var mıdır?

Firmalar	A	B	C	
İstenmeyen Madde Oran Toplamı	28	14	32	TKT=74

9- Bir iş yerinde aynı işi yapan üç kişinin rasgele olarak belirlenen 4 iş gününde ürettikleri parça sayıları aşağıda verilmiştir. Bu üç işçinin ortalama verimliliklerinin farklı olup olmadığını 0,01 anlamlılık düzeyinde test ediniz.

	İşçiler		
Günler	A	B	C
1	80	78	80
2	90	82	82
3	79	85	82
4	80	76	84

10- Bir kozmetik firmasında çalışan 4 satış elemanının 5 ay boyunca sattıkları ürünlere ilişkin tablo aşağıda verilmiştir. 0,05 anlamlılık düzeyinde satış elemanlarının verimlilikleri arasında fark var mıdır?

Satış elemanı	Satış Sayıları				
1	60	65	70	75	80
2	80	85	86	79	75
3	70	75	86	84	76

11- Üç yağın kızartmada emilmelerinin farklı olup olmadığını araştıran bir deney hazırlanmış ve alınan bilgiler aşağıdaki tabloda verilmiştir. Tabloya göre varyans analiz tablosunu hazırlayınız ve gruplar arasında fark olup olmadığını 0,05 anlamlılık düzeyinde test ediniz.

	Yağ Türleri	
A	**B**	**C**
5	6	8
8	2	6
7	3	8
7	5	9
	3	5
	4	

12- Yaşları ve hastalık dereceleri aynı olan hastalar arasından 25'i rasgele seçiliyor. İlk 8 kişiye A tedavisi, sonraki 8 kişiye B tedavisi ve son 9 kişiye C tedavi yöntemi uygulanıyor ve iyileşme sürelerine bakılıyor. Sonuçlar aşağıdaki tabloda verilmiştir. 0,05 anlamlılık düzeyinde iyileşme süreleri açısından tedaviler arasında fark var mıdır? Fark var ise farklı olan grubu bulunuz.

	Tedavi Türleri	
A	**B**	**C**
4	6	8
7	2	6
7	3	8
6	5	7
6	3	5
5	3	6
5	3	5
6	3	5
		7

13-http://www.adli-sicil.gov.tr/ISTATIST.HTM web sayfasından yıllara göre bölgelerde açılan davalar ile ilgili bilgileri elde ediniz, bu sayıların logaritmaları alındığında varyans çözümleme koşullarını sağlayacağını gösteriniz ve bölgeler arası farkın ortalamaları arasında 0,01 anlamlılık düzeyinde fark olup olmadığını varyans analizi kullanarak tartışınız.

14- Bir bitki türünün boylarının yetiştiği yöreye göre değişip değişmediğini araştırmak amacıyla 4 yörede bu bitki türünden 5'er ölçüm yapılmış aşağıdaki sonuçlar elde edilmiştir.

	Yöreler			
	A	B	C	D
Ortalama	11	20	21	13
Gözlem Sayısı	5	5	5	5
Varyans	4	6,25	9	2,89

a- Bu problemde varyans çözümlemesi uygulanabildiğine göre hangi varsayımlar sağlanmıştır.

b- Varyans çözümleme tablosunu hazırlayınız. Gerekli hipotezi kurup 0,05 anlamlılık düzeyinde test ediniz.

15- Fasulye verim denemesi için rasgele tarlalardan ürün miktarlarına ilişkin aşağıdaki bilgiler verilmiştir.

	Fasulye Çeşitleri		
	A	B	C
Gözlem Sayısı	6	5	3
Toplam	40,8	31	21,6

TKT=14,36

Varyans çözümleme tablosunu düzenleyiniz. 0,05 anlamlılık düzeyinde test ediniz

16- Patates bitkisinde hastalıktan korunmak için kükürt kullanılmaktadır. Bir grup kontrol grubu olarak bırakılmış yani kükürt uygulanmıştır. Üç farklı düzeyde kükürt uygulanması sonucu bitkilerdeki hastalık yüzdeleri saptanmış ve aşağıdaki özet bilgiler verilmiştir.

	Kükürt Düzeyleri			
	S_0	S_1	S_2	S_3
Ortalama	48	43	21	34
Gözlem Sayısı	6	6	6	6
Varyans	4,7	5,7	4	4,5

a- Bu problemde varyans çözümlemesi uygulanabildiğine göre hangi varsayımlar sağlanmıştır.

b- Varyans çözümleme tablosunu hazırlayınız. Gerekli hipotezi kurup 0,05 anlamlılık düzeyinde test ediniz.

c- Newman Keuls testini kullanarak en iyi kükürt düzeyini bulunuz.

17- Gözlem sayılarının farklı olması durumunda kullanılan $\overline{n}$ eşitliğinin $n_1=n_2=...=n_k$ olması durumunda n’e eşit olacağını gösteriniz.

18- 3 bayan tezgahtarın rasgele seçilen 5 ay içinde sattıkları ürün miktarları aşağıdadır. 0,01 anlamlılık düzeyinde ürün satış sayıları arasında fark olup olmadığını test ediniz. Eğer fark var ise, farklı olan grubu bulunuz.

Ürün Satış Sayıları

A	B	C
60	52	46
56	51	50
58	65	56
59	50	52
59	62	53

19- 3 farklı ehliyet kursuna devam eden 6'şar öğrencinin 10 üzerinden aldıkları yazılı sınav sonuçları aşağıdaki tabloda verilmiştir. 0,05 anlamlılık düzeyinde sınav sonuçları arasında fark var mıdır?

Yazılı Sınav Sonucu

1	2	3
6	6	8
7	7	6
7	5	8
6	5	7
6	4	5
5	4	6

20- Bir otelin ön büro elemanı otele gelen müşterilerin kalma süreleri ile ilgili bir çalışma yapmak istemektedir. Bu amaçla 1 yıl içerisinden rasgele 44 müşteriyi seçip kaldıkları ay (haziran, temmuz, ağustos) ve kalma sürelerini incelemiştir. Bu üç ay içerisinde müşterilerin kalma süreleri bakımından 0,05 anlamlılık düzeyinde farkının olup olmadığını test ediniz. Hangi ayın daha yoğun olduğu konusunda bir bilgi söylenebilir mi?

Haziran	Temmuz	Ağustos
10	16	21
4	13	20
6	20	23
12	11	13
7	12	13
10	12	21
7	15	14
13	13	12
14	14	24
6	11	19
0	17	26
0	14	23
6	13	19
	15	24
	13	18
		18

11. ONBİRİNCİ BÖLÜM

11.1. GİRİŞ

11.2. EŞİT OLASILIKLI SINIFLAR İÇİN UYUM İYİLİĞİ TESTİ

11.3. EŞİT OLASILIKLI OLMAYAN SINIFLAR İÇİN UYUM İYİLİĞİ TESTİ

11.4. ÇAPRAZ TABLOLAR

PROBLEMLER

11.1. GİRİŞ

Uygulamada, bazı problemler gözlemlerin sınıflara sayısal dağılmasının incelenmesi ile ortaya çıkar. Örneğin satış sayılarının mevsimlere dağılımı, okur yazar sayılarının bölgelere dağılımı gibi. Gözlemlerin sınıflara dağılımı bazen eşit olasılık ile gerçekleşmesi beklenirken, bazen de her sınıf için belli olasılıklarla dağılması söz konusudur. Uygulamada sınıflara dağılımın belli olasılık dağılımlarına benzemesi de karşılaşılan durumlardandır. Bu dağılımlar normal, poisson, binom vb. gibi dağılımlar olabilir. Gözlenen dağılımın belirlenen olasılık dağılımlarına uyup uymadığı uyum iyiliği testi ile test edilir. Normal dağılıma uyum testi için geliştirilmiş çeşitli testler vardır. Bilgisayar paket programlarında bu testler normal dağılıma uyum testleri olarak yer alır (goodness of fit tests). Uyum iyiliği testi için sınıflara dağılımın eşit olasılık ile, belli olasılıklar ile ya da bir olasılık dağılımına uygun olarak dağılımın varsayımı ile kuramsal sıklıklar bulunur. Gözlenen sıklıklar ile kuramsal sıklıkların birbirlerine uygunluğu, problemin durumuna göre aşağıda verilen hipotezlerin test edilmesi ile gerçekleştirilir. Eğer gözlemlerin sınıflara dağılımının eşit olasılıklı olduğu varsayılıyor ise kurulacak hipotez,

H_0: Gözlemler sınıflara eşit olasılıklı dağılmıştır.

H_A: Gözlemler sınıflara eşit olasılıklı dağılmamıştır.

biçimindedir. Eğer gözlemlerin sınıflara dağılımı belli olasılıklar ile ise hipotez,

H_0: Gözlemlerin sınıflara dağılımı verilen olasılıklarladır.

H_A: Gözlemlerin sınıflara dağılımı verilen olasılıklarla değildir.

biçimindedir. Eğer gözlemlerin sınıflara dağılımı normal, poisson, binom vb. ise,

H_0: Gözlemlerin dağılımlarının normal dağılımdan (poisson dağılımından, binom dağılımından) ayrılışları önemsizdir.

H_A: Gözlemlerin dağılımlarının normal dağılımdan (poisson dağılımından, binom dağılımından) ayrılışları önemlidir.

şeklinde kurulur.

Yukarıda verilen H_0 hipotezlerinin doğruluğu varsayılarak, kullanılacak istatistik Eşitlik (11.1)'de verilmiştir:

$$\chi_H^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} = \sum_{i=1}^{k} \frac{O_i^2}{E_i} - n \qquad (11.1)$$

Burada k: Grup sayısı, O_i: i. gözlenen sıklık, E_i: i. beklenen sıklık, n=toplam gözlem sayısıdır. Eşitlik (11.1), k–1 serbestlik dereceli χ^2 dağılımı göstermektedir. Eğer $\chi^2_H > \chi^2_{\alpha;k-1}$ ise H_0 hipotezi reddedilir, diğer durumda reddedilemez.

11.2. EŞİT OLASILIKLI SINIFLAR İÇİN UYUM İYİLİĞİ TESTİ

H_0 hipotezine dayanarak E_i beklenen sıklıkların sınıflara düşme olasılıkları p_i'ler aracılığı ile bulunur. p_i'ler eşit olasılıklı ($p_i = \frac{1}{k}$) ve $p_1 + p_2 + ... + p_k = 1$ olmalıdır.

Beklenen sıklıklar,

$$E_i = p_i \times n \tag{11.2}$$

eşitliğinden bulunur. Burada $n=O_1+O_2+...+O_k$ olup, $E_1 + E_2 + . . . + E_k = n$ olmalıdır. E_i 'ler bulunur. Eşitlik (11.1)'de uygulanarak hipotez test edilir.

Örnek 11.1: Üç farklı içeceğin tercihlerini belirlemek amacıyla rasgele 33 kişi seçilmiştir. Kişilerin tercihlerine göre dağılımı Tablo 11.1'de verilmiştir. 0,05 anlamlılık düzeyinde çeşit tercihleri arasında önemli bir fark var mıdır?

Tablo 11.1 İçecek Tercih Gözlenen Sıklık Değerleri

	Gözlenen Sıklık (O_i)
1. Çeşit	8
2. Çeşit	10
3. Çeşit	15
Toplam	33

H_o: İçecek tercihlerinin sınıflara dağılımı eşit olasılıkladır (İçecekler, tercih açısından aynıdır).

H_A: İçecek tercihlerinin sınıflara dağılımı eşit olasılıkla değildir (İçecekler, tercih açısından aynı değildir).

Örnek 11.1 için p_i olasılıkları aşağıda verilmiştir.

$$p_1 = \frac{1}{3},\ p_2 = \frac{1}{3},\ p_3 = \frac{1}{3}$$

Tablo 11.2 İçecek Tercih Gözlenen ve Beklenen Sıklık Değerleri

	Gözlenen Sıklık (O_i)	Beklenen Sıklık (E_i)
1. Çeşit	8	$\frac{1}{3} \times 33 = 11$
2. Çeşit	10	$\frac{1}{3} \times 33 = 11$
3. Çeşit	15	$\frac{1}{3} \times 33 = 11$
Toplam	33	33

Khi-kare hesap değeri,

$$\chi^2_H = \frac{(8-11)^2}{11} + \frac{(10-11)^2}{11} + \frac{(15-11)^2}{11} = 2{,}36$$

olarak bulunur. χ^2 için tablo değeri, k-1=3-1=2 serbestlik derecesinde ve 0,05 anlamlılık düzeyinde, $\chi^2_{0,05;2} = 5{,}99$'dur. 2,36<5,99 olduğundan, H_0 hipotezi reddedilemez. 0,05 anlamlılık düzeyinde kişilerin üç tür içecek tercihleri arasında önemli bir fark olduğu söylenemez ($p > 0{,}05$).

Örnek 11.2: Bir madeni paranın 200 kez atılmasında 115 tura, 85 yazı gözlenmiştir. Paranın dengeli olduğu hipotezini 0,01 ve 0,05 anlamlılık düzeyinde test ediniz.

Tablo 11.3 Para Örneği Tablosu

	O_i	E_i
Yazı	85	100
Tura	115	100
Toplam	200	200

H_o: Para dengelidir (Tura gelmesi ile yazı gelmesi aynı şansa sahiptir).

H_A: Para dengeli değildir (Tura gelmesi ile yazı gelmesi aynı şansa sahip değildir).

$$p_1 = \frac{1}{2};\ p_2 = \frac{1}{2};\ E_1 = \frac{1}{2} \times 200 = 100;\ E_2 = \frac{1}{2} \times 200 = 100$$

$$\chi^2_H = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} = \frac{(85-100)^2}{100} + \frac{(115-100)^2}{100} = 2,25 + 2,25 = 4,50$$

0,01 anlamlılık düzeyinde ve 1 serbestlik derecesinde, $\chi^2_{0,01;1} = 6,63$'dir. 4,50<6,63 olduğundan H_o hipotezi reddedilemez. 0,01 anlamlılık düzeyinde, paranın yazı ya da tura gelmesinin aynı şansa sahip olduğu söylenebilir.

0,05 anlamlılık düzeyinde ve 1 serbestlik derecesinde, $\chi^2_{0,05;1} = 3,84$'dir. 4,50>3,84 olduğundan H_o hipotezi reddedilir. 0,05 anlamlılık düzeyinde, paranın yazı ya da tura gelmesinin aynı şansa sahip olduğu söylenemez.

11.3. EŞİT OLASILIKLI OLMAYAN SINIFLAR İÇİN UYUM İYİLİĞİ TESTİ

Örnek 11.1 ve Örnek 11.2, sıklıkların gruplara eşit olasılık ile dağıldığı durumlar için verilmiştir. Bazı problemlerde grup olasılıkları farklı olabilir. Burada test edilecek hipotez, gözlenen sıklıkların sınıflarda belirlenen bu olasılıklar ile dağıldığıdır. Örnek 11.3'de bu tür problem verilmiştir.

Örnek 11.3: Bir borsa aracı kurumu oluşturduğu portföyler ile ilgili aşağıdaki tabloyu düzenlemiştir. Tablo 11.4'e bakarak 2003 yılının ilk üç ayında oluşturulan portföy getirilerinin 0,05 anlamlılık düzeyinde geçen yılın aynı dönemindeki gibi olduğu söylenebilir mi?

Tablo 11.4 2003 ve 2004 Yılının Portföy Getirileri

Getiri %	2003 ilk üç ay (adet)	2004 ilk üç ay (adet)
$-15 > X$	50	45
$-15 \le X < 0$	45	40
$0 \le X < 10$	125	140
$10 \le X < 15$	115	170
$15 \le X$	100	130
Toplam	435	525

H_o: 2004'ün ilk üç ayına ait portföylerin getirisi geçen yılın aynı dönemindeki getirisine eşittir.

H_A: 2004'ün ilk üç ayına ait portföylerin getirisi geçen yılın aynı dönemindeki getirisine eşit değildir.

Önce H_0 hipoteze göre olasılıklar elde edilmelidir:

P(–15>X)=50/435=0,115 ; P(–15 ≤X<0)=45/435=0,103 ; P(0≤X<10)=125/435=0,288
P(10≤X<15)=115/435=0,266 ; P(15≤ X)=100/435=0,230

Eğer 2004 yılının ilk üç ayına ait getiriler 2003 yılının getirileri ile aynı ise, getiri gruplarına ait 2003 yılı için bulunan olasılıklar 2004 yılında da aynı olmalıdır. Bu görüş doğrultusunda, her bir sınıf için beklenen sıklıklar sırasıyla;

E_1= 0,115× 525=60,37 ; E_2= 0,103×525=54,07 ; E_3= 0,288× 525=151,20
E_4= 0,264× 525=138,60 ; E_5=0,230× 525=120,75

olur ve khi-kare hesap değeri,

$$\chi^2_H = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} = \frac{(45-60,37)^2}{60,37} + \frac{(40-54,07)^2}{54,07} + \ldots + \frac{(130-120,75)^2}{120,75}$$
$$= 3,91 + 3,66 + 0,82 + 7,11 + 0,70 = 16,2$$

olarak elde edilir. 4 serbestlik derecesinde ve 0,05 anlamlılık düzeyinde tablo değeri, $\chi^2_{0,05;4} = 9,49$ 'dur. 16,2>9,49 olduğundan H_o hipotezi reddedilir. 0,05 anlamlılık düzeyinde, 2004'ün ilk üç ayına ait portföylerin getirisi geçen yılın aynı dönemindeki getirisine eşit olmadığı söylenebilir.

Örnek 11.4: Bir işletmede, satışlardan sağlanan karı artırmak amacı ile yapılan çalışmada, satışların mevsimlere göre dağılımının Tablo 11.5'de verildiği biçimde olması gerektiği ifade edilmiştir. Bu yıl içinde yapılan toplam 7414 adet satış işleminin adet olarak dağılımı da aynı tabloda verilmiştir. Yapılan satışların, istenen yapıda olup olmadığını 0,05 anlamlılık düzeyinde test ediniz.

Tablo 11.5 Satışların Mevsimlere Göre Dağılım

Mevsimler	İstenen Satış Oranı (%)	Gerçekleşen Satış (O_i)	Beklenen Satış (E_i)
Kış	40	2521	0,40×7414=2965,6
Yaz	15	1040	15×7414=1112,1
Bahar	20	1731	20×7414=1482,8
Güz	25	2122	25×7414=1853,5
Toplam	100	7414	7414

H_o: Beklenen oranlar doğrultusunda satışlar gerçekleşmiştir.

H_A: Beklenen oranlar doğrultusunda satışlar gerçekleşmemiştir.

H_0 hipotezinin doğruluğu kabul edilirse; kış için beklenen sıklık 0,40×7414 = 2965,6 bulunur. Diğerleri de benzer şekilde elde edilir ve khi-kare hesap değeri,

$$\chi^2_H = \frac{(2521-2965{,}6)^2}{2965{,}6} + \frac{(1040-1112{,}1)^2}{1112{,}1} + \frac{(1731-1482{,}8)^2}{1482{,}8} + \frac{(2122-1853{,}5)^2}{1853{,}5}$$

$$=66{,}65+4{,}67+41{,}57+38{,}89=151{,}75$$

olur. 3 serbestlik derecesinde ve 0,05 anlamlılık düzeyinde tablo değeri, $\chi^2_{0,05;3} = 7{,}814$ 'dir. 151,75>7,814 olduğundan H_o hipotezi reddedilir. 0,05 anlamlılık düzeyinde satışlar belirlenen oranlarda gerçekleşmemiştir denebilir.

11.4. ÇAPRAZ TABLOLAR

11.4.1. Çapraz Tablo Oluşumu ve Yüzdelerinin Yorumları

Çapraz tablolar bir olaya etki eden iki değişkeni birlikte ele alabilmek için kullanılan tablolardır. Özellikle, anket uygulanarak bilgi toplandığında, bazı değişkenlerin birlikte düşünülmesi araştırıcıya önemli bilgiler verir. Örneğin bir araştırmada uygulanan ankette, yerleşim yeri ile ilk doğum yaşını birlikte düşünerek, ilk doğum yaşının hangi yerleşim yerinde daha küçük olduğu bilinmek istenebilir. Bu tür tablolarda satır, sütun yüzdelerinin verilmesi gerekir. Bu bölümde, oluşturulacak çapraz tablolarda yüzdeler ve hipotez testleri ele alınacaktır. Bunları bir örnek ile açıklayalım.

Örnek 11.5: Kırsal ve kentsel alanda yaşayan 0-5 yaş arasında çocuğu olan 50 annenin bebek beslenmesi ile ilgili davranışlarını belirlemek için yapılan araştırmada, yerleşim yeri ve doğumda kimden yardım aldığı sorulmuştur. Yerleşim yeri sorusunda; 1: kentsel alan, 2: kırsal alan yerleşim yerlerini gösteren kodlar ve doğumda kimden yardım aldınız sorusunda; 1: doktorlardan, 2: hemşireden, 3: diplomalı ebeden, 4: sağlık eğitimi olmayan önceden doğum yaptırmış kişi olarak kodlanmıştır. Bu sorular için anket sonuçları aşağıda verilmiştir.

Anket Numarası	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
Yerleşim yeri	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2
Doğumda kimden yardım aldığı	1	1	1	1	1	1	3	1	1	1	2	1	3	3	1	1	3	1	1	3	1	1	1	1	1	3

Anket Numarası	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
Yerleşim yeri	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Doğumda kimden yardım aldığı	1	2	2	2	1	2	3	3	3	3	4	1	3	3	3	3	3	2	2	3	3	2	4	1

Yerleşim yeri ve doğuma yardım eden kişi için bir çapraz tablo oluşturulursa aşağıdaki tabloya ulaşılır.

Tablo 11.6 Yerleşim Yeri ve Doğuma Yardım Eden Kişi için Çapraz Tablo

			Doğumunuza yardım eden kişi				Toplam
			Doktor	Hemşire	Ebe (diplomalı)	Önceden Doğum Yaptıran Kişi (sağlık eğitimi yok)	
Yerleşim yeri	kentsel	**sayı**	**19**	**1**	**5**		**25**
		% yerleşim yeri	76	4	20		100
		% doğumunuza yardım eden kişi?	82,6	12,5	29,4		50
		% Toplam	38	2	10		50
	kırsal	**sayı**	**4**	**7**	**12**	**2**	**25**
		% yerleşim yeri	16	28	48	8	100
		% doğumunuza yardım eden kişi?	17,4	87,5	70,6	100	50
		% Toplam	8	14	24	4	50
Toplam		**sayı**	**23**	**8**	**17**	**2**	**50**
		% yerleşim yeri	46	16	34	4	100
		% doğumunuza yardım eden kişi?	100	100	100	100	100
		% Toplam	46	16	34	4	100

Tablo 11.6’da, satırlar yerleşim yerlerini, sütunlar ise doğuma yardım eden kişileri belirtmektedir. Anne, kentte oturuyor ve doğumunda doktor yardımcı olmuş ise bu iki özelliğin kesiştiği göze de sayılacaktır. Tablo 11.6’da görülen ilk yüzde satırlara, ikinci yüzdeler sütunlara ve üçüncü yüzde toplama göre alınmıştır. Örneğin satır yüzdeleri için kent yerleşim yerlerinde 19/25=0,76’yı, sütun yüzdeleri için 19/23=0,826’yı, toplama göre yüzdeler için ise 19/50=0,38 değerlerinin bulunmasını örnek olarak verebiliriz.

Satır yüzdelerine göre yorum: Tabloda yerleşim yerine göre yüzdelerde %76’nın yorumu; Bu çalışmada kentlerde yaşayan annelerin %76’sı doğumlarında doktordan yardım almıştır. Kentlerde yaşayan annelerin %4’ü doğumlarında hemşireden yardım almıştır. Kentlerde yaşayan annelerin %20’si doğumlarında diplomalı ebeden yardım almışlardır. Diğer satır yüzdelerine ilişkin yorumlar da benzer şekilde yapılabilir.

Sütun yüzdelerine göre yorum: İlk sütun yüzdeleri için yorum; Araştırmaya göre, doğumlarında doktordan yardım almış annelerin %82,6’sı kent yerleşim yerlerinde, %17,4’ü kırsal yerleşim bölgelerinde oturmaktadır.

Toplama Göre Yüzdelerin Yorumu: Araştırmaya dahil edilen annelerin %38'i kentte oturup, doğumlarında doktordan yardım almaktadır.

11.4.2. Çapraz Tablolarda Hipotez Testleri

Bu bölümde, oluşturulacak çapraz tablolarda hipotez testleri ele alınacaktır. İki değişkenli tablolarda test edilecek hipotez, gruplar arası fark kontrolü ya da iki değişken arasındaki ilişkinin önem kontrolüdür. Bu nedenle aşağıda verilen iki farklı H_o hipotezi kullanılır.

H_o: Değişkenler arasında ilişki yoktur.

H_A: Değişkenler arasında ilişki vardır.

ya da

H_o: Gruplar arasında fark yoktur (Gruplar homojendir).

H_A: Gruplar arasında fark vardır (Gruplar homojen değildir).

Her iki hipotezin testi için kullanılacak test istatistiği H_o hipotezinin doğruluğu koşulu altında, hesaplanan beklenen sıklıklar E_i ile gözlenen sıklıklar O_i'ler arasında olan farkın büyüklüğüne dayanır. Uyum iyiliği testlerinde de görüldüğü gibi bu fark,

$$\chi^2_H = \sum_{j=1}^{c}\sum_{i=1}^{r}\frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}} \tag{11.3}$$

eşitliği ile test edilir. Burada;

O_{ij} = i. satır j. sütunda bulunan gözlenen sıklık,

E_{ij} = i. satır j. sütunda bulunan beklenen sıklıktır.

χ^2 dağılımı serbestlik derecesine bağlı bir dağılım olduğundan, tablo değerleri de serbestlik dereceleri kullanılarak bulunur. Burada serbestlik derecesi sd=(r–1)(c–1)'dir. r satır sayısını, c sütun sayısını göstermektedir. Eğer $\chi^2_H > \chi^2_{\alpha;(r-1)(c-1)}$ ise H_0 hipotezi reddedilir, diğer durumda reddedilemez.

İlişki (Bağıntı) Kontrolü

Tablo 11.7'de verildiği gibi A ve B gibi iki değişken olduğunu, A değişkeninin r, B değişkeninin c sınıflı olduğu düşünülüp, r x c tane sınıflama yapıldığı kabul edilsin. A ve B değişkenleri nitel özellik taşımaktadır. Örneklemdeki deneklerin bağlı olduğu bu iki değişken arasında ilişki kontrolü yapılmak istenmektedir.

Tablo 11.7 A ve B Değişkenlerine Bağlı Tablolar

	Özellik B							
Özellik A	1	2	3	...	j	...	c	Toplam
1	O_{11}	O_{12}	O_{13}	...	…	...	O_{1c}	R_1
2	O_{21}	O_{22}	O_{23}	...	…	...	O_{2c}	R_2
3	O_{31}	O_{32}	O_{33}	...	…	...	O_{3c}	R_3
...	...	...	...	...	...	...	...	...
i	...	...	...	...	O_{ij}	....	...	R_i
...	...	...	...	...	…	...	...	...
r	O_{r1}	O_{r2}	O_{r3}	...	...	...	O_{rc}	R_r
Toplam	C_1	C_2	C_3	...	C_j	...	C_c	N

Tablo 11.7'de rxc çapraz gözede, rxc tane O_{ij} gözlem değerleri bulunmaktadır. Eğer değişkenler arasındaki ilişkinin önemsiz olduğu hipotezinin doğruluğu kabul edilirse, her gözede olması gereken sayı ne olacaktır sorusuna cevap aranır. Bu soruya, satır ve sütun toplamlarının sabit olması durumunda anlamlı bir cevap bulunabilir. Bu durumda gözlemlerin 1. satırda bulunma olasılığı R_1/N, i. satırda bulunma olasılığı R_i/N'dir. 1. sütunda C_1 kadar olan gözlemin, 1. satırda bulunma olasılığı R_1/N olup, bu gözede beklenen sıklık,

$$E_{11} = \frac{C_1 \times R_1}{N}$$

dir. i. satır j. sütun için,

$$E_{ij} = \frac{C_j \times R_i}{N} \qquad (11.4)$$

olacaktır.

Not: Burada dikkat edilecek nokta E_{ij}'lerin 5'den küçük olmamasıdır. Eğer 5'den küçük beklenen sıklık varsa, bu sıklığın yer aldığı satır ya da sütun tablodaki uygun bir satır ya da sütun ile birleştirilmelidir. Ancak uygulamada, bu konuda bir esneklik getirilmiştir. Eğer rxc sayıdaki gözenin %20'sinden azında 5 ve 5'in altında gözlem varsa Eşitlik (11.3)'ün kullanılabileceği belirtilmektedir.

Örnek 11.6: Bir şirket üst düzey yöneticilerinden 190 kişilik bir örneklem seçerek şirket düzeyindeki stratejiyi ve mali başarımı değerlendirmelerini istemiştir. Elde edilen bilgiler Tablo 11.8'de verilmiştir. Mali başarım ve şirket düzeyindeki strateji arasında ilişki yoktur hipotezini 0,01 anlamlılık düzeyinde test ediniz.

Tablo 11.8 Şirkete ait Strateji ve Mali Başarım Değerleri

Strateji	Mali Başarım						Toplam
	Düşük		Orta		Yüksek		
Düşük	O_{11}=66	E_{11} = 63,79	O_{12}=40	E_{12} = 42,95	O_{13}=14	E_{13} = 13,26	120
Yüksek	O_{21}=35	E_{21} = 37,21	O_{22}=28	E_{22} = 25,05	O_{23}=7	E_{23} = 7,74	70
Toplam	101		68		21		190

H_o: Mali başarım değerlendirmeleri ile şirketin stratejisini değerlendirme arasında ilişki yoktur.

H_A: Mali başarım değerlendirmeleri ile şirketin stratejisini değerlendirme arasında ilişki vardır.

H_0 hipotezi, üst düzey yöneticilerinin mali başarımı değerlendirirken şirketin stratejisinden bağımsız karar verdikleri anlamına gelir.

Örnek 11.6 için Tablo 11.8 ile verilen çapraz tablo 2x3'lük bir çapraz tablodur. H_o hipotezinin doğruluğu altında Eşitlik (11.4) kullanılarak beklenen sıklıklar sırasıyla;

$$E_{11} = \frac{120\times 101}{190} = 63{,}79 \quad E_{12} = \frac{120\times 68}{190} = 42{,}95 \quad E_{13} = \frac{120\times 21}{190} = 13{,}26$$

$$E_{21} = \frac{70\times 101}{190} = 37{,}21 \quad E_{22} = \frac{70\times 68}{190} = 25{,}05 \quad E_{23} = \frac{170\times 21}{190} = 7{,}74$$

olarak hesaplanıp Tablo 11.8'de yazılmıştır.

$$\chi^2_H = \sum_j \sum_i \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = \frac{(66-63{,}79)^2}{63{,}79} + \frac{(40-42{,}95)^2}{42{,}95} + \frac{(14-13{,}26)^2}{13{,}26} + \frac{(35-37{,}21)^2}{37{,}21}$$

$$+ \frac{(28-25{,}05)^2}{25{,}05} + \frac{(7-7{,}74)^2}{7{,}74} = 1{,}051$$

sd=(2–1)(3–1)=2'dir. α=0,05 anlamlılık düzeyi için tablo değeri, $\chi^2_{0,05;2} = 5{,}99$'dur. 1,051<5,99 olduğundan H_o hipotezi reddedilemez. 0,05 anlamlılık düzeyinde, yüksek düzey yöneticilerin mali başarım değerlendirmesi ile strateji değerlendirmeleri arasında ilişki yoktur.

Örnek 11.7: Bir bölgede yaşayan 220 kişi rasgele seçilip, sigara içip içmedikleri ve sağlıklarından şikayetçi olup olmadıkları sorulmuştur. Cevaplara ilişkin aşağıda verilen tabloya göre 0,05 anlamlılık düzeyinde ilişki olup olmadığını test ediniz.

Tablo 11.9 Sigara İçme ve Sağlığından Şikayetçi Olma Değerleri

Sağlığınızdan şikayetçi misiniz?	Sigara içer misiniz?				Toplam
	Evet		Hayır		
Evet	O_{11}=100	E_{11} = 76	O_{12}=40	E_{12} = 64	140
Hayır	O_{21}=20	E_{21} = 44	O_{22}=60	E_{22} = 36	80
Toplam	120		100		220

H_o: Sigara içmek ile sağlığından şikayetçi olmak arasında ilişki yoktur.
H_A: Sigara içmek ile sağlığından şikayetçi olmak arasında ilişki vardır.

H_o doğru ise, $E_{11} = \frac{120 \times 140}{220} = 76{,}36 \approx 76$ ve diğer beklenen sıklıklar elde edilir. Elde edilen değerlerden khi-kare değeri,

$$\chi^2_H = \frac{(100-76)^2}{76} + \frac{(40-64)^2}{64} + \frac{(20-44)^2}{44} + \frac{(60-36)^2}{36} = 45{,}6$$

olarak elde edilir. sd=(2–1)(2–1)=1 ve 0,05 anlamlılık düzeyinde tablo değeri $\chi^2_{0,05;1} = 3{,}841$'dır. 45,6>3,841 olduğundan H_o hipotezi reddedilir. 0,05 anlamlılık düzeyinde, sigara içmekle sağlığından şikayetçi olmak arasında ilişki vardır.

Örnek 11.7'nin sonucu, ikinci bir soruyu akla getirir. İlişki olduğuna göre bu ilişki ne kadardır? Çapraz tablolarda kullanılan çeşitli ilişki katsayıları vardır. Burada kısaca önemli üç tanesi verilecektir.

Bağımlılık Katsayıları

Pearson'un C Katasayısı

Satır ve sütun sayısı ikiden çok olan çapraz tablolarda tercih edilen ölçü,

$$C = \sqrt{\frac{\chi_H^2}{\chi_H^2 + N}} \qquad (11.5)$$

olan Pearson'un C katsayısıdır.

Örnek 11.7'de verilen sigara problemi için ilişki ölçüsü aşağıda verilmiştir:

$$C = \sqrt{\frac{45,6}{45,6 + 220}} = 0,41$$

Yorum: Pearson C katsayısına göre sigara içmek ile sağlığından şikayetçi olmak arasında %41'lik bir ilişki vardır.

Cramer'in V Katsayısı

χ^2 değeri önemli olduğunda kullanılan katsayı aşağıda verilen Cramer'in V katsayısıdır.

$$V = \sqrt{\frac{\chi_H^2}{N(k-1)}} \qquad (11.6)$$

Eşitlik (11.6)'da k, satır ya da sütundan küçük olanın sayısına eşittir. Cramer'in V katsayısı, herhangi boyuttaki tabloda kullanılabilir olduğundan, çok tercih edilen bir katsayıdır.

Phi-Katsayısı

Sayımla belirtilen kitlelerde, iki değişken arasındaki bağımlılığın yüzde olarak ölçüsüdür. 2×2 çapraz tablolarda tercih edilir. Phi-katsayısı Eşitlik (11.7)'de verilmiştir:

$$\phi = \sqrt{\frac{\chi_H^2}{N}} \qquad (11.7)$$

Aynı problem için yukarıda verilen katsayılar farklı değerler verebilir. Araştırıcı bu katsayılardan hangisini tercih ettiğini belirtmeli ve yorumunu ona göre yapmalıdır.

Gruplar Arası Fark Kontrolü

Nitel değişkenlerde farklı gruplarda alınan bilgilerin birbirlerinden farklı olup olmadığının testi için kullanılır.

Örnek 11.8: Bir sigorta şirketi 2004 yılının Ocak ayında, dört farklı alanda sigorta yaptıran müşterilerinden tabakalı basit rasgele örnekleme kullanarak 260 müşterisini seçip araştırma yapmıştır. İncelediği konulardan biri de bu müşterilerinin poliçelerini Ocak ayında yenileyip yenilemedikleridir. Sigorta alanları ve gözlenen bilgiler Tablo 11.10'da verilmiştir. Sigorta alanları arasında yeniden sigortalama bakımından önemli fark olduğu söylenebilir mi?

Tablo 11.10 2004 Yılında Sigortalarını Yenileyen Müşterilerin Dağılımı

Sigorta Alanları	**2004 yılında poliçelerini yenileyenler**	**%**	**2004 yılında poliçelerini yenilemeyenler**	**%**	**Toplam Poliçe Sayısı**
Yangın Sigortası	38	68	18	32	56
Ulaşım	40	71	16	29	56
DASK	90	86	15	14	105
Kaza	16	37	27	63	43
Toplam	184	71	76	29	260

H_o: Sigorta poliçesi alan müşterilerin poliçelerini yeniletmesi bakımından sigorta dalları arasında önemli fark yoktur.

H_A: Sigorta poliçesi alan müşterilerin poliçelerini yeniletmesi bakımından sigorta dalları arasında önemli fark vardır.

H_o hipotezine dayanarak, ilişki kontrolünde anlatıldığı gibi beklenen sıklıklar bulunur. Sigorta müşterilerinin sigortalarını yenileme oranları;

$$\hat{P}_{.1} = \frac{R_1}{N} = \frac{184}{260} = 0{,}71 \text{ ve } \hat{P}_{.2} = \frac{R_2}{N} = \frac{76}{260} = 0{,}29$$

olarak bulunmuştur. Eğer sigorta poliçeleri arasında yenileme açısından önemli fark yok ise bu oranın her grup için geçerli olması gerekir. Böylece beklenen sıklıklar bulunabilir.

1. satır 1. sütun $E_{11} = \frac{R_1}{N} \times C_1 = \frac{184}{260} \times 56 = 0{,}71 \times 56 \cong 40$ olur.

1. satır 2. sütun $E_{12} = \frac{R_2}{N} \times C_1 = \frac{76}{260} \times 56 = 0{,}29 \times 56 \cong 16$ olur.

Eğer E_{11} bulunmuş ise $E_{12} = C_1 - E_{11} = 56 - 40 = 16$ olarak daha kolay bulunabilir. Diğer beklenen sıklıklar aynı düşünce ile hesaplanır ve aşağıda verilen Tablo 11.11 oluşturulur.

Tablo 11.11 2004 Yılında Sigortalarını Yenileyen Müşterilerin Gözlenen ve Beklenen Sıklıklarla Birlikte Dağılımı

Sigorta Alanları	**2004 yılında poliçelerini yenileyenler**		**2004 yılında poliçelerini yenilemeyenler**		**Toplam Poliçe Sayısı**
Sıklıklar	O_{ij}	E_{ij}	O_{ij}	E_{ij}	R_i
Yangın Sigortası	38	40	18	16	56
Ulaşım	40	40	16	16	56
DASK	90	75	15	30	105
Kaza	16	29	27	14	43
Toplam	184	184	76	76	260

Gruplar arası fark kontrollerinde, H_o hipotezi reddedildiğinde farklı olan grubu ayırmak için khi-kare hesap değerinin her grup için tek tek hesaplanması yararlı olur. Bu problemde;

$$\chi_1^2 = \frac{(38-40)^2}{40} + \frac{(18-16)^2}{16} = 0{,}35$$

$$\chi_2^2 = \frac{(40-40)^2}{40} + \frac{(16-16)^2}{16} = 0$$

$$\chi_3^2 = \frac{(90-75)^2}{75} + \frac{(15-30)^2}{30} = 10{,}5$$

$$\chi_4^2 = \frac{(16-29)^2}{29} + \frac{(27-14)^2}{14} = 17{,}90$$

olup, $\chi^2_H = \sum_{i=1}^{4} \chi^2_i = 28{,}75$'dir. 0,05 anlamlılık düzeyinde, (4–1)(2–1)=3 serbestlik derecesinde tablo değeri, $\chi^2_{0,05;3} = 7{,}81$ bulunur.

28,75>7,81 olduğundan H_o hipotezi reddedilir. 0,05 anlamlılık düzeyinde poliçe yeniletme açısından sigorta dalları arasında önemli fark vardır.

Khi-kare değeri en büyük olan poliçe, 4. grup olan kaza poliçeleridir ($\chi^2_4 = 17{,}90$). Bu grubu çıkarıp, işlemleri kalan gruplar arasında yineleyerek devam edebiliriz.

Ho :Yenilenen poliçeler arasında fark yoktur.

H_A :Yenilenen poliçeler arasında fark vardır.

Tablo 11.12 2004 Yılında Sigortalarını Yenileyen Müşterilerin Dağılımı (Kaza Poliçesi hariç)

Sigorta Alanları	**2004 yılında poliçelerini yenileyenler**		**2004 yılında poliçelerini yenilemeyenler**		**Toplam Poliçe Sayısı**
Sıklıklar	O_{ij}	E_{ij}	O_{ij}	E_{ij}	R_i
Yangın Sigortası	38	43	18	13	56
Ulaşım	40	44	16	12	56
DASK	90	81	15	14	105
Toplam	168	168	49	49	217

$\hat{P}_{.1} = \frac{168}{217} = 0{,}77$ $E_{11} = 0{,}77 \times 56 = 43{,}12 \cong 43$ ve diğer beklenen sıklıklar aynı biçimde elde edilir. Khi-kare değerleri sırasıyla,

$$\chi^2_1 = \frac{(38-43)^2}{43} + \frac{(18-13)^2}{13} = 2{,}5$$

$$\chi^2_2 = \frac{(40-44)^2}{44} + \frac{(16-12)^2}{12} = 1{,}69$$

$$\chi^2_3 = \frac{(90-81)^2}{81} + \frac{(15-24)^2}{24} = 4{,}375$$

olup, $\chi^2_H = \sum_{i=1}^{3} \chi^2_i = 8{,}56$'dır. 0,05 anlamlılık düzeyi ve (3–1)(2–1)=2 serbestlik derecesinde tablo değeri $\chi^2_{0,05;2} = 5{,}99$ olarak elde edilir.

8,56>5,99 olduğundan, H_0 hipotezi reddedilir. 0,05 anlamlılık düzeyinde yangın sigortası, ulaşım, DASK sigorta türleri arasında yeniden sigortalama yönünden önemli fark bulunmaktadır.

Khi-kare değeri en büyük olan poliçe, 3. grup olan DASK poliçeleridir. Bu grubu çıkarıp işlemleri kalan gruplar arasında yineleyerek devam edebiliriz.

Ho :Yenilenen poliçeler arasında fark yoktur.

H_A :Yenilenen poliçeler arasında fark vardır.

Tablo 11.13 2004 Yılında Sigortalarını Yenileyen Müşterilerin Dağılımı (Kaza ve DASK Poliçesi hariç)

Sigorta Alanları	**2004 yılında poliçelerini yenileyenler**		**2004 yılında poliçelerini yenilemeyenler**		**Toplam Poliçe Sayısı**
Sıklıklar	O_{ij}	E_{ij}	O_{ij}	E_{ij}	R_i
Yangın Sigortası	38	39	18	17	56
Ulaşım	40	39	16	17	56
Toplam	78	78	34	34	112

$$\chi_1^2 = \frac{(38-39)^2}{39} + \frac{(18-17)^2}{17} = 0{,}0844$$

$$\chi_2^2 = \frac{(40-39)^2}{39} + \frac{(16-17)^2}{17} = 0{,}0844$$

$\chi_H^2 = \sum_{i=1}^{2} \chi_i^2 = 0{,}1688$’dir. 0,05 anlamlılık düzeyi ve (2–1)(2–1)=1 serbestlik derecesinde tablo değeri $\chi_{0,05;1}^2 = 3{,}84$ olarak elde edilir.

0,1688<3,84 olduğundan H_0 reddedilemez. 0,05 anlamlılık düzeyinde yangın sigortası ve ulaşım arasında yeniden sigortalama yönünden önemli bir fark bulunmamaktadır.

Problem için verilen ilk tabloda poliçelerini yenileyenlerin oranı verilmektedir. Bu oranlar ve elde edilen test sonuçları birleştirildiğinde, en fazla yenilemenin DASK’ta olduğu (0,86); yangın (0,68) ve ulaşım (0,71) sigortalarının arasında önemli fark olmadığı, en az yenilemenin kaza poliçelerinde gerçekleştiği söylenebilir.

PROBLEMLER

1- Yüksek tansiyonun sigara içme ile ilişkisini araştırmak amacıyla bilgiler toplanmış ve aşağıdaki çapraz tabloda verilmiştir. Verilen bilgilerden yararlanarak, yüksek tansiyonun sigara içme ile ilişkisinin olup olmadığı 0,05 anlamlılık düzeyinde test ediniz.

	Sigara		
	İçmeyen	Az içen	Çok içen
Yüksek Tansiyon	40	50	30
Normal Tansiyon	60	10	10

2- Limonlu ve portakallı bulaşık deterjanlarından iki marka bir sene boyunca izlenmiş, her hafta bu markaların indirimli satış yapıp yapmadığı kaydedilmiştir. Aşağıdaki çapraz tabloya göre, iki marka arasında ilişkinin olup olmadığını 0,05 anlamlılık düzeyinde test ediniz.

B marka

A marka	İndirim var	İndirim yok
İndirim var	18	23
İndirim yok	14	5

3- Aşağıdaki tablo, kadın ve erkekler için bağımsız rasgele örneklemlerde günde 3 saatten az ya da çok televizyon izleyenlerin sayısını göstermektedir. Bir kimsenin cinsiyetiyle televizyon seyretme süresi arasında ilişkinin olup olmadığını 0,01 anlamlılık düzeyinde test ediniz.

Günlük seyredilme süresi

Cinsiyet	3 saatten az	3 saatten çok
Erkek	20	12
Kadın	15	14

4- Bir zar 200 kez atılmış ve üste gelen yüzün sayıları aşağıdaki tabloda verilmiştir. Tabloya göre, 0,05 anlamlılık düzeyinde zarın hilesiz olduğu söylenebilir mi?

1	2	3	4	5	6	Toplam
29	26	30	40	35	40	200

5- İki farklı diyet yöntemi, zayıflamak isteyen 50'şer kişide uygulanmıştır. Aşağıda verilen tabloyu kullanarak diyet yöntemleri arasında önemli bir farkın olup olmadığını 0,01 anlamlılık düzeyinde test ediniz.

Diyet Yöntemleri	İstenen Sonuç Alınanlar	İstenen Sonuç Alınmayanlar	Toplam
A	40	10	50
B	30	20	50

6- Aynı tür malzeme, iki farklı firmanın ürettiği aynı büyüklükteki tencerelerde aynı sürede pişirilmiştir. A firmasının tencereleri ile 80 deneme yapılıyor ve bunlardan 40 tanesinin iyi piştiği, B firmasının tencereleri ile 120 deneme yapılıyor ve bunlardan 110 tanesinin iyi piştiği belirleniyor. İyi pişirmede firmaların ürettiği tencereler arasında 0,05 anlamlılık düzeyinde fark olup olmadığını test ediniz.

7- Bir firma ürettiği fırınlarda kalite kontrolü için, satış yaptığı müşterilerden altı ay kullanım sonunda gelen şikayetleri kaydetmiştir. Fırınların çalıştığı ısılar 350^0 C ve 450^0 C olmak üzere sabittir. Sonuçta aşağıdaki tablo elde edilmiştir. Şikayet ile yüksek sıcaklıkta çalışma arasında 0,05 anlamlılık düzeyinde ilişki önemli midir?

Şikayetçi	350^0 C	450^0 C	Toplam
Olan	5	20	25
Olmayan	45	30	75

8- İki farklı katkı maddesi kullanılan hazır dondurma paketlerinden rasgele 54 tane seçilip, bozuk olup olmadıkları kontrol edilmiştir. Koruma yönünden katkı maddeleri arasında 0,01 anlamlılık düzeyinde fark var mıdır?

Katkı Maddesi	Kontrol Sonucu		
	Bozuk	Sağlam	Toplam
A	6	12	18
B	8	28	36

9- Beslenme ile çocukların başarıları arasında ilişki aranmaktadır. Bunun için aşağıdaki tablo hazırlanmıştır. 0,05 anlamlılık düzeyinde ilişki var mıdır? Pearson C katsayısını hesaplayınız ve ne anlama geldiğini açıklayınız.

Beslenme Durumu	Başarılı	Başarısız	Toplam
İyi	70	20	90
Kötü	20	80	100

10- Piyasaya sürülen yeni bir içki türünün beğenilip beğenilmediği üç yaş grubuna soruluyor. Alınan sonuçlar aşağıda verilmiştir. 0,01 anlamlılık düzeyinde ürünün beğenilmesi ile yaş arasında bir ilişki var mıdır? Ne kadardır?

Yaş Grubu	Beğenilme Durumu		
	Beğeniliyor	Beğenilmiyor	Toplam
Genç	70	20	90
Orta	40	50	90
Yaşlı	20	80	100

11- Kız ve erkek öğrencilerin sigara içme alışkanlıklarına göre dağılımı aşağıdaki tabloda verilmiştir. Sigara içme alışkanlığı ile cinsiyet arasındaki ilişki katsayısını hesaplayınız.

Sigara	Kız	Erkek	Toplam
İçiyor	28	62	90
İçmiyor	40	30	70

12- Demografik araştırma yapan bir araştırıcının ilgilendiği bir konu, annelerin eğitim düzeyi ile çocuk sayısı arasında bir ilişkinin var olup olmadığıdır. Bu amaçla yaptığı anket çalışmasında, anne eğitimi ile çocuk sayısını çaprazlayarak aşağıda verilen tabloyu elde etmiştir. Verilen bilgileri kullanarak araştırıcının çalışmasından elde edeceği sonucu 0,05 anlamlılık düzeyinde test ederek veriniz. İlişki varsa bu ilişkinin ne kadar olduğunu bulunuz.

	Çocuk Sayısı		
Annenin Eğitimi	0-1	2-3	4 ve daha çok
Temel Eğitim	19	42	25
Lise	14	37	40
Yüksek	37	17	4

13- Bir üniversitenin kütüphane sorumlusu ders dışı kitap isteklerinin fakültelere göre dağılımını araştırmak istiyor. Bu amaçla bir yarı yılda kütüphaneden kitap alanların fakültelerini, aldığı kitabın ders ile ilgisinin olup olmadığını kaydediyor. Sonuçta aşağıda verilen tablo elde edilmiştir. Bu bilgileri kullanarak 0,01 anlamlılık düzeyinde ders dışı kitap okuma bakımından fakülteler arası fark var mıdır? Ders dışı kitap okuma yönünden hangi fakülte daha öndedir?

Fakülteler	Alınan kitap ders dışı	Alınan kitap ders ile ilgili
Tıp	20	80
Fen	45	55
Sosyal Bil.	75	25
Sağlık Bil.	40	60
Mühendislik	30	80

14- Bir sigorta acentesi belli bir yılı baz alıp, bu yıl için yürürlükte olan poliçe sayılarını üçer aylık devrelerde kaydetmiştir. Bir sonraki ve iki yıl sonraki kayıtlar aşağıda verilmiştir. 0,05 anlamlılık düzeyinde indirimsiz poliçelerde baz alınan yıla göre diğer iki yıl farklı mıdır?

Baz alınan yıl sonuçları:

Tarih	Yürürlükteki Poliçe		
	İndirimsiz	%20 İndirimli	%40 indirimli
1 ocak	2460	2045	4001
1 nisan	2527	2147	4216
1temmuz	2735	2280	4227
1ekim	2877	2546	4375
1ocak	2900	2697	4486

Bir sonraki yıl sonuçları:

Tarih	Yürürlükteki Poliçe		
	İndirimsiz	%20 İndirimli	%40 indirimli
1 ocak	2910	2698	4488
1 nisan	2913	2699	4490
1temmuz	2899	2699	4499
1ekim	2900	2712	4515
1ocak	2955	2737	4525

İki yıl sonraki sonuçları:

Tarih	Yürürlükteki Poliçe		
	İndirimsiz	%20 İndirimli	%40 indirimli
1 ocak	2955	2730	4425
1 nisan	2957	2735	4426
1temmuz	2965	2740	4435
1ekim	2966	2742	4445
1ocak	2970	2746	4449

15- On dördüncü soruda baz alınan yılda üçer aylık dönemlere olan dağılım indirimsiz poliçeler ile %20 indirimli poliçeler arasında 0,05 anlamlılık düzeyinde fark var mıdır?

16- On dördüncü soruda %20 indirimli poliçelerde baz alınan yıla göre 0,05 anlamlılık düzeyinde diğer iki yıl farklı mıdır?

17- On dördüncü soruda %40 indirimli poliçelerde baz alınan yıla göre 0,05 anlamlılık düzeyinde diğer iki yıl farklı mıdır?

18- Seçim zamanı yapılan bir ön araştırmada örnekleme kurallarına uygun olarak seçilen 800 seçmenin tercih ettikleri siyasal parti ve meslekleri sorularak aşağıda verilen tablo hazırlanmıştır. 0,01 anlamlılık düzeyinde parti tercihi ile çalışma yeri arasında ilişki var mıdır?

Çalışma yerleri	A partisi	B partisi	C partisi	D partisi
Kamuda çalışanlar	40	30	80	150
Özel şirkette çalışan.	80	60	40	50
Serbest çalışan	85	80	35	40
İşsizler	15	9	5	1

19- TRT'nin yaptığı bir anket çalışmasında, yaş gruplarına göre TV izleyen bireylerin en çok izledikleri televizyon kanallarına ilişkin çapraz tablo aşağıda verilmiştir. 0,01 anlamlılık düzeyinde, yaş grupları ile en çok izlenen TV kanalları arasında ilişki var mıdır?

TV Kanalları	Yaş Grupları		
	14-24	25-30	31-40
TRT1	300	315	320
SHOW TV	375	262	265
ATV	323	262	255
KANAL D	356	227	218
NTV	98	89	103

20- TRT'nin yaptığı bir anket çalışmasında, çalışma durumuna göre Televizyon kanallarını izleyen bireylerin kanallarda yayınlanmasını istedikleri program türleri çapraz tablosu aşağıda verilmiştir. 0,05 anlamlılık düzeyinde, çalışma durumu ile izlenen programlar arasında ilişki var mıdır?

Program	Çalışma Durumu		
	Çalışıyor	Çalışmıyor	Toplam
Yerli Dizi	240	426	666
Yabancı Dizi	99	170	269
Belgesel Programlar	370	316	686
Müzik Eğlence Programları	162	207	369
Tartışma Programları	265	190	455
Toplam	1136	1309	2445

12. ONİKİNCİ BÖLÜM

12.1. GİRİŞ
12.2. KORELASYON KATSAYISI
12.3. DOĞRUSAL REGRESYON ANALİZİ
PROBLEMLER

12.1. GİRİŞ

Günlük yaşantımızda karşılaşılan sorunların çoğu, iki ya da daha çok değişken arasında bağıntının olup olmadığının araştırılması ile ilgilidir. Bu bölümde, yalnız iki değişkenin olması durumunda değişkenler arasındaki ilişkiler incelenecektir. Burada verilecek bilgilerin kullanılabilmesi için değişkenlerin normal dağılıma sahip olma varsayımının kabul edilmesi gerekmektedir. Belirtilen koşullar altında iki değişken arasındaki ilişki miktarı, korelasyon katsayısı ile değişkenler arasındaki bağıntının fonksiyonel yapısı regresyon ile ele alınacaktır.

12.2. KORELASYON KATSAYISI

Korelasyon, ilgilenilen iki ya da daha çok sayıda değişken arasındaki ilişkiyi gösterir. Bu ilişkinin miktarı da korelasyon katsayısı ile belirlenir. Bir şirketin reklamlar için harcadığı miktar ile o dönem yapılan satışlar arasındaki ilişki gibi. Korelasyon çözümlemesinin amacı, değişkenler arasındaki ilişkinin derecesini ve yönünü belirlemektir. Değişkenlerden biri X diğeri Y ise bu iki değişken arasındaki ilişkiye basit korelasyon denir. Değişken sayısı ikiden fazla ise, değişkenler arasında çoklu korelasyon ya da kısmi korelasyon katsayıları bulunur. Burada basit korelasyon incelenecektir.

Yukarıda da denildiği gibi, korelasyon ölçülecek olan ilişkinin miktarıdır. İlişki miktarının ölçülmesi iki değişken için verilen ortak varyans tanımına dayanır. Eğer X rastlantı değişkeni μ_x ortalama ve σ_x^2 varyansa, Y raslantı değişkeni μ_y ortalama ve σ_y^2 varyansa sahipse, ortak varyans (kovaryans)

$$\mathrm{Or}(V) = E\left[(X-\mu_x)(Y-\mu_y)\right] \qquad (12.1)$$

dır. Ancak ortak varyans incelenen değişkenlerin ölçü birimine bağlıdır. Bu durum farklı ölçeklerde olan değişkenler için ilişkinin ölçüsü olarak karşılaştırmada kullanılamaz. Ortak varyans değişkenlerin standart sapmalarına bölünerek, ölçü biriminden bağımsız bir ilişki katsayısı elde edilir. Bu ilişki katsayısı, ρ ile gösterilen Eşitlik (12.2) ile verilen korelasyon katsayısıdır. Korelasyon katsayısı, $-1 \le \rho \le 1$ arasındadır ve kitleye ulaşma güçlülüğünden ρ 'nun bilinmesi zordur. Ancak örneklem bilgilerinden yararlanarak bu parametre tahmin edilebilir.

$$\rho = \frac{Or(V)(X,Y)}{\sigma_x \sigma_y} = \frac{E[(X-\mu_x)(Y-\mu_y)]}{\sqrt{E(X-\mu_x)^2 E(Y-\mu_y)^2}} \tag{12.2}$$

Korelasyon Katsayısı için Tahmin Formülü

X ve Y değişkenleri arasındaki doğrusal ilişkinin basit korelasyon katsayısı tahmini r ile gösterilir ve formülü,

$$r = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\left[\sum(x-\bar{x})^2\right]\left[\sum(y-\bar{y})^2\right]}}$$

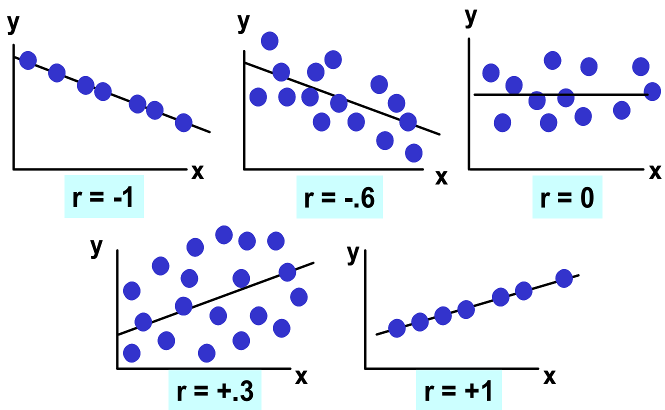
$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{\left[n\left(\sum x^2\right)-\left(\sum x\right)^2\right]\left[n\left(\sum y^2\right)-\left(\sum y\right)^2\right]}} \tag{12.3}$$

ya da

$$r = \frac{\sum x_i.y_i - \left(\sum x_i \sum y_i\right)/n}{\sqrt{\sum x_i^2 - \left(\sum x_i\right)^2/n}\ \sqrt{\sum y_i^2 - \left(\sum y_i\right)^2/n}} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sqrt{\left[\sum x_i^2 - n\bar{x}^2\right]\left[\sum y_i^2 - n\bar{y}^2\right]}} \tag{12.4}$$

dır. Kitle korelasyonu için verilen $-1 \leq \rho \leq 1$ kısıtı tahmin için de geçerlidir. Yani $-1 \leq r \leq 1$'dır. $r < 0$ ise negatif bir ilişki vardır, $r \cong -1$ ise kuvvetli negatif bir ilişki vardır, $r = 0$ ise doğrusal ilişki yoktur, $r > 0$ ise pozitif bir ilişki vardır, $r \cong 1$ ise kuvvetli doğrusal pozitif bir ilişki vardır denilir.

Değişkenler arasında olan ilişkinin, yönünü görebilmek için saçılım grafiği yapılır. Şekil 12.1'de, çeşitli korelasyon katsayıları için saçılım grafikleri görülmektedir.



Kaynak: Groebner vd., 2001.

Şekil 12.1 Gözlenen Noktaların Saçılımı ile Korelasyon Katsayısı Arasındaki Bağıntı

Şekil 12.1'de verilen saçılım grafiklerinden görüldüğü gibi, iki değişkenin her ikisi de aynı yönde değişim gösteriyorsa ilişki pozitiftir ve korelasyon katsayısının işareti (+) olur. Biri artarken diğeri azalıyor ise ilişki negatiftir ve korelasyon katsayısının işareti (–) olur. İki değişken birlikte değişim göstermiyor ise ilişki yoktur ($r = 0$) denilir.

Örnek 12.1: Bir iş yerinde çalışan 10 işçinin iş yetenekleri (X) ve ortak çalışmalarda gösterdikleri uyumları (Y) 10 üzerinden değerlendirilmiş ve Tablo 12.1'de verilmiştir. Verilere ilişkin korelasyon katsayısını hesaplayınız.

Tablo 12.1 İşçilerin İş Yetenek ve Ortak Çalışmalardaki Uyum Puanları

X	Y	XY	X^2	Y^2
9	9	81	81	81
5	8	40	25	64
7	8	56	49	64
7	4	28	49	16
6	6	36	36	36
5	5	25	25	25
3	6	18	9	36
3	3	9	9	9
1	1	1	1	1
4	5	20	16	25
$\sum x = 50$	$\sum y = 55$	$\sum xy = 314$	$\sum x^2 = 300$	$\sum y^2 = 357$

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{\left[n\left(\sum x^2\right) - \left(\sum x\right)^2\right]\left[n\left(\sum y^2\right) - \left(\sum y\right)^2\right]}} = \frac{10 \times 314 - 50 \times 55}{\sqrt{\left[10 \times 300 - 50^2\right]\left[10 \times 357 - 55^2\right]}} = \frac{3140 - 2750}{\sqrt{500 \times 545}}$$

$$= \frac{390}{\sqrt{272500}} = \frac{390}{522} = 0{,}747$$

Yorum: İşçilerin iş yetenek ve ortak çalışmalardaki uyum puanları arasında yaklaşık olarak 0,75'lik pozitif doğrusal bir ilişki vardır.

Örnek 12.2: 10 günlük dönemde A hisse senedinin kapanış fiyatı ile Merkez Bankası (MB) dolar alış kuru değerleri Tablo 12.2'de verilmiştir. Dolar alış fiyatı ile hisse senedi fiyatı arasındaki ilişki miktarını bulunuz.

Tablo 12.2 A Hisse Senedinin Kapanış Fiyatı ile MB Dolar Alış Kuru Değerleri

Hisse senedi kapanış fiyatı	1,450	1,455	1,375	1,300	1,280	1,350	1,375	1,450	1,400	1,425
MB dolar alış kuru	1,360	1,362	1,363	1,350	1,345	1,345	1,375	1,400	1,400	1,450

X: Hisse senedi kapanış fiyatı

Y: MB dolar alış kuru

$$n = 10 \quad \sum_{i=1}^{10} x_i = 13{,}86 \quad \sum_{i=1}^{10} y_i = 13{,}75 \quad \sum_{i=1}^{10} x_i^{\,2} = 19{,}2448 \quad \sum_{i=1}^{10} y_i^{\,2} = 18{,}91609 \quad \sum_{i=1}^{10} x_i y_i = 19{,}06706$$

$$r = \frac{\sum x_i . y_i - \left(\sum x_i \sum y_i\right)/n}{\sqrt{\sum x_i^2 - \left(\sum x_i\right)^2/n}\ \sqrt{\sum y_i^2 - \left(\sum y_i\right)^2/n}}$$

$$= \frac{19{,}06706 - \left[(13{,}86)(13{,}75)\right]/10}{\sqrt{19{,}2448 - (13{,}86)^2/10}\ \sqrt{18{,}91609 - (13{,}75)^2/10}} = \frac{0{,}00956}{0{,}01851} = 0{,}5164$$

Yorum: Dolar alış fiyatı ile hisse senedi fiyatı arasında 0,51'lik pozitif ilişki bulunur.

Örnek 12.3: A ve B hisse senetlerinin yıllar itibariyle getiri oranları Tablo 12.3'de verilmiştir. Bu iki hisse senedi arasında yıllar itibariyle ilişki miktarı ne kadardır?

Tablo 12.3 A ve B Hisse Senetlerinin Yıllar İtibariyle Getiri Oranları

Yıllar	1995	1996	1997	1998	1999	2000	2001
A hisse senedi getirisi	0,06	0,16	-0,10	0,10	0,18	-0,05	0,14
B hisse senedi getirisi	0,10	0,10	-0,10	0,18	0,21	0,10	0,25

X: A hisse senedi getirisi

Y: B hisse senedi getirisi

$$n = 7 \quad \sum_{i=1}^{7} x_i = 0{,}49 \quad \sum_{i=1}^{7} y_i = 0{,}84 \quad \sum_{i=1}^{7} x_i^{\,2} = 0{,}1037 \quad \sum_{i=1}^{7} y_i^{\,2} = 0{,}179 \quad \sum_{i=1}^{7} x_i y_i = 0{,}1178$$

$$r = \frac{0{,}1178 - \left[(0{,}49)(0{,}84)\right]/7}{\sqrt{0{,}1037 - (0{,}49)^2/7}\sqrt{0{,}1790 - (0{,}84)^2/7}} = \frac{0{,}0590}{\sqrt{0{,}0694}\sqrt{0{,}0782}} = \frac{0{,}0590}{0{,}0736} = 0{,}801$$

Yorum: A ve B hisse senetlerinin yıllar itibariyle ilişki miktarı, 0,80'lik pozitif bir ilişkidir.

Korelasyon Katsayısının Önem Kontrolü

Örneklemden tahmin edilen ilişki katsayısının çekildiği kitlenin, ilişki katsayısının önemli olup olmayacağı test edilmek istenebilir. Bu durumda hipotezlerin testi için kullanılacak test istatistiği;

$$t_H = \frac{r-\rho}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{(r-\rho)\sqrt{n-2}}{\sqrt{1-r^2}} \tag{12.5}$$

dir. Burada, r: Örneklemin korelasyon katsayısı, ρ: Çekildiği kitlenin hipotez ile belirlenen korelasyon katsayısı, $\sqrt{\frac{1-r^2}{n-2}}$: r'nin standart sapmasıdır. Burada hipotezler korelasyon katsayılarının sıfırdan farklı olup olmadığını sınamak için yapılacaktır. Bu nedenle hipotezlerde $\rho_0 = 0$ olarak alınacaktır.

Eğer hipotez,

$$\begin{array}{l} H_0: \rho \le \rho_0 \\ H_A: \rho > \rho_0 \end{array} \quad \text{ya da} \quad \begin{array}{l} H_0: \rho = \rho_0 \\ H_A: \rho > \rho_0 \end{array}$$

biçiminde ise, $t_H > t_{\alpha;\ n-2}$ olduğunda H_0 hipotezi reddedilir, $t_H \le t_{\alpha;\ n-2}$ olduğunda H_0 hipotezi reddedilemez.

Problem için uygun hipotez,

$$\begin{array}{l} H_0: \rho \ge \rho_0 \\ H_A: \rho < \rho_0 \end{array} \quad \text{ya da} \quad \begin{array}{l} H_0: \rho = \rho_0 \\ H_A: \rho < \rho_0 \end{array}$$

biçiminde ise, $t_H < -t_{\alpha;\ n-2}$ olduğunda H_0 hipotezi reddedilir, $t_H \ge -t_{\alpha;\ n-2}$ olduğunda H_0 hipotezi reddedilemez.

Eğer hipotez,

$$H_0: \rho = \rho_0$$
$$H_A: \rho \neq \rho_0$$

biçiminde ise, $t_H > t_{\alpha/2;\ n-2}$ ya da $t_H < -t_{\alpha/2;\ n-2}$ olduğunda H_0 hipotezi reddedilir, diğer durumda reddedilemez.

Örnek 12.4: Bir mağazaya 10 gün içinde her gün gelen müşterilerin sayısı ve günlük satış miktarları Tablo 12.4'de verilmiştir. 0,05 anlamlılık düzeyinde, gelen müşteri sayısı ile günlük satış miktarı arasında ilişki olmadığı yönündeki iddiayı test ediniz.

Tablo 12.4 Müşteri Sayısı ve Günlük Satış Miktarları

Gün	1	2	3	4	5	6	7	8	9	10
Müşteri sayısı	30	25	20	37	31	50	29	40	42	36
Satış miktarı (milyon TL)	2600	3000	2100	3200	2500	3900	2600	3700	4300	3100

X: Müşteri sayısı.

Y: Satış miktarı (milyon TL).

$$H_o: \rho = 0$$
$$H_A: \rho \neq 0$$

$$n = 10 \quad \sum_{i=1}^{10} x_i = 340 \quad \sum_{i=1}^{10} y_i = 31000 \quad \sum_{i=1}^{10} x_i^2 = 12256 \quad \sum_{i=1}^{10} y_i^2 = 100420000 \quad \sum_{i=1}^{10} x_i y_i = 1101500$$

$$r = \frac{1101500 - \left[(340)(31000)\right]/10}{\sqrt{12256 - (340)^2/10}\sqrt{100420000 - (31000)^2/10}} = \frac{47500}{\sqrt{696}\sqrt{4320000}} = \frac{47500}{(26,3818)(2078,461)} = 0,8663$$

$$t_H = r\sqrt{\frac{n-2}{1-r^2}} = 0,8663\sqrt{\frac{10-2}{1-(0,8663)^2}} = 4,9052$$

$$t_{\alpha/2;n-2} = t_{0,025;10-2} = 2,3060$$

4,90>2,3060 olduğundan H_o hipotezi reddedilir.

Yorum: 0,05 anlamlılık düzeyinde ρ sıfırdan farklıdır, yani değişkenler arasındaki ilişki önemlidir.

12.3. DOĞRUSAL REGRESYON ANALİZİ

Regresyonun değişkenler arasındaki fonksiyonel ilişkiyi gösterdiği daha önce belirtildi. Değişken sayısı iki olan regresyon modeline basit regresyon, ikiden çok bağımsız değişkeni olan regresyon modeline çoklu regresyon modeli denir. Basit regresyonda da çoklu regresyonda da bağlı olan değişken bir tanedir. Bu değişkene etki eden değişken bir ya da daha çok olur. Burada sadece basit regresyon modeli incelenecektir. Bir regresyon probleminde, değişkenler arasındaki ilişkiyi fonksiyonel olarak açıklamak ve bu ilişkiyi bir modelle tanımlayabilmek için çalışılır.

Bir kitlede gözlenen X ve Y değişkenleri arasındaki doğrusal ilişki aşağıdaki doğrusal regresyon modeli ile verilebilir.

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

$X \to$ bağımsız değişken

$Y \to$ bağımlı değişken

$\beta_0 \to X = 0$ olduğunda bağımlı değişkenin alacağı değer

$\beta_1 \to$ regresyon katsayısı

$\varepsilon \to$ ortalaması 0, varyansı σ^2 olan hata terimidir.

β_1 regresyon katsayısı, bağımsız değişkendeki bir birimlik değişimin, bağımlı değişkende yaratacağı değişikliği gösterir. ε hata terimi, önerilen modelin vereceği değer ile gerçekte gözlemlenen değer arasındaki farktır. $\varepsilon_i = (\beta_0 + \beta_1 X_i) - Y_i$ veya $Y_i - (\beta_0 + \beta_1 X_i)$.

Doğrusal regresyon modelinin aşağıda verilen varsayımları vardır:

1- x_i değerleri ya sabittir ya da ε_i hata teriminden bağımsız X raslantı değişkeninin gerçekleşmiş değerleridir.

2- ε, ortalaması 0 olan raslantı değişkenidir.

$E(\varepsilon_i)=0 \qquad (i=1,2,...,n)$

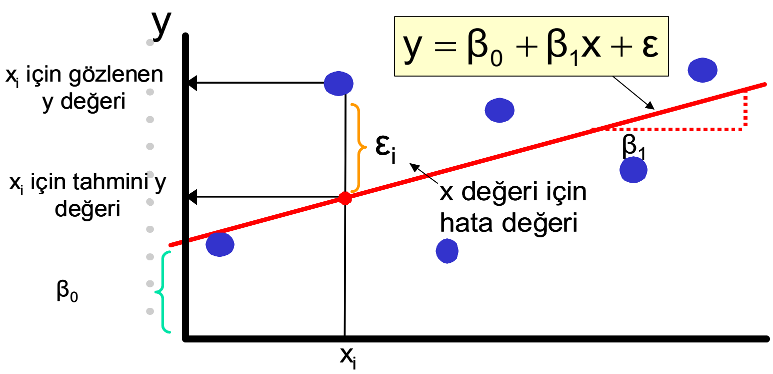
3- Tüm ε_i'lerin varyansı aynıdır.

$V(\varepsilon_i)=\sigma^2 \qquad (i=1,2,...,n)$

4- ε_i'ler birbiriyle ilişkili değildir.

$\rho(\varepsilon_i \varepsilon_j) = 0 \qquad i \neq j$

Şekil 12.2'de, gözlenen y değerleri ile tahmini y değerleri ve hata terimi arasındaki ilişkiyi veren grafik verilmiştir.



Kaynak: Groebner vd., 2001.

Şekil 12.2 Gözlenen y Değeri ile Tahmini y Değeri ve Hata Terimi Arasındaki İlişki

Tanımlanan doğrusal regresyon modeli, kitleden seçilen n gözlemli örneklem için;

$$\hat{Y} = b_0 + b_1 X \tag{12.6}$$

biçimindedir. Bu tahmini denklem bir bağımlı ve bir bağımsız değişken arasındaki bağıntının matematiksel gösterimidir. Bağımsız değişken değerleri denklemde yerine konularak tüm bağımsız değişkene ilişkin tahmini $\hat{Y}$ bulunabilir. Amaç, örneklem bilgilerinden yararlanarak, kitle için tutarlı ve güvenilir tahmin değerleri vermek olduğu için, gerçek değerler ile (Y_i) tahmini değerler $(\hat{Y}_i)$ arasında fark olmaması ya da çok az fark olması beklenir. Bu nedenle b_0 ve b_1 katsayıları $Y_i - \hat{Y}_i$ değerlerini her i için minimum yapacak şekilde bulunmalıdır. Bunun için çeşitli yöntemler geliştirilmiştir. Bu yöntemlerden biri en küçük kareler kriteridir. Bu kritere göre katsayılar $\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2 = \sum_{i=1}^{n}\left[Y_i - \left(b_0 + b_1 x\right)\right]^2$ eşitliğini en küçük yapacak b_0 ve b_1 bilinmeyen değerleri matematik kurallarına göre uygun olarak bulunur. Bu değerler:

$$b_1 = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(x-\bar{x})^2} = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sum x^2 - \frac{\left(\sum x\right)^2}{n}} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} \tag{12.7}$$

$$b_0 = \bar{y} - b_1\bar{x} \tag{12.8}$$

biçimindedir.

Değişkenler birlikte artıyor ya da azalıyorsa b_1 pozitif değerli, değişkenlerden biri artarken diğeri azalıyorsa b_1 negatif değerlidir.

Örnek 12.5: Örnek 12.1'deki verilere ilişkin regresyon modelini oluşturunuz.

Katsayılara sırasıyla,

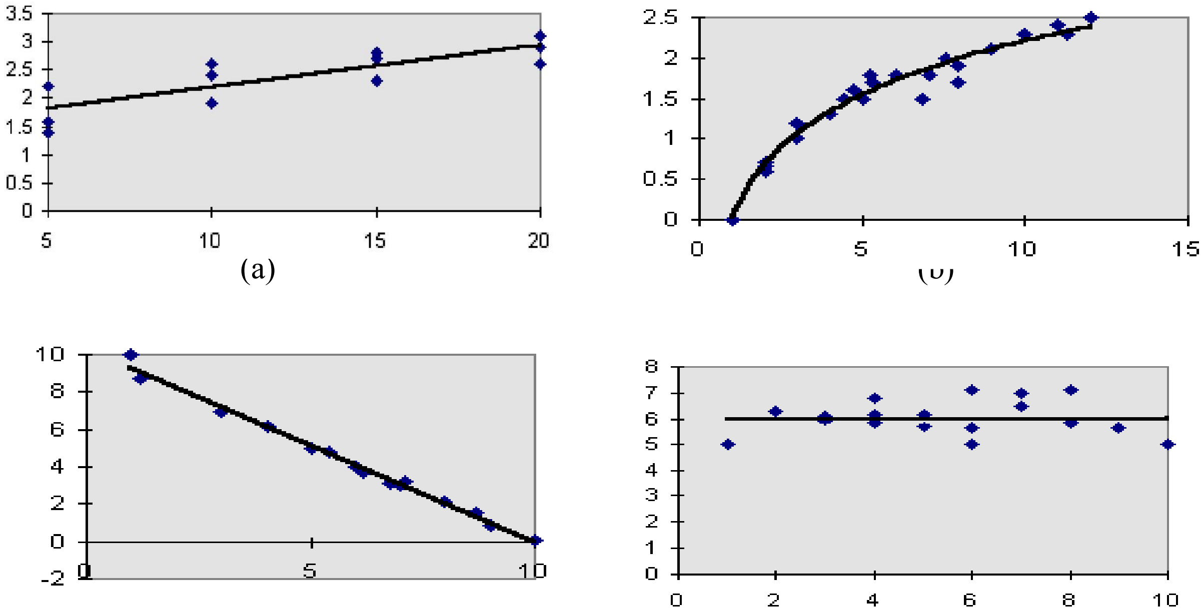
$$b_1 = \frac{314 - \frac{50 \times 55}{10}}{300 - \frac{50^2}{10}} = \frac{314 - 275}{50} = \frac{39}{50} = 0{,}78$$

$$b_0 = \bar{y} - b_1\bar{x} \Rightarrow 5{,}5 - 0{,}78 \times 5 = 1{,}6$$

olmak üzere regresyon modeli $\hat{Y} = 1{,}6 + 0{,}78X$ olur.

Regresyon Doğrusunun Çizimi

Doğrunun y eksenini kestiği nokta b_0'dır. Aşağıdaki grafik çeşitli regresyon eğrilerine ilişkindir.



(a) (b)

(c) (d)

Kaynak: Groebner vd., 2001.

Şekil 12.3 Çeşitli Regresyon Model Çizimleri; Pozitif Doğrusal İlişki (a), Doğrusal Olmayan İlişki (b), Negatif Doğrusal İlişki (c), Doğrusal İlişki Yok (d)

Regresyon Katsayısının Önem Kontrolü

X bağımsız değişkeni ile Y bağımlı değişkeni arasında doğrusal bir illişkinin varlığı, her bir deneğin x_i ve y_i değerlerinin, koordinat düzlemi üzerinde oluşturdukları noktaların dağılımına bakılarak tahmin edilebilir. Ancak, bu tahminin tutarlı olup olmadığının araştırılması gerekir. Bunun için regresyon katsayısının önem kontrolü, doğrusallıktan ayrılışın önem kontrolü yapılır. Bu kontroller yapılırken bazı eşitliklerin bilinmesi gerekir. Bu eşitlikler aşağıda verilmiştir.

Çeşitli Kareler Toplamı, Serbestlik Dereceleri ve Kareler Ortalaması

X ortalama ayrılış kareler toplamı (XOAKT),

$$XOAKT = \sum_{i=1}^{n} \left(x_i - \overline{X}\right)^2$$

$$= \sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n} = \sum x_i^2 - n\overline{x}^2 = S_{xx} \qquad (12.9)$$

olup, serbestlik derecesi= (n-1)'dir.

Y ortalama ayrılış kareler toplamı (YOAKT),

$$YOAKT = \sum_{i=1}^{n} \left(y_i - \overline{Y}\right)^2$$

$$= \sum_{i=1}^{n} y_i^2 - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n} = \sum y_i^2 - n\overline{y}^2 = S_{yy} \qquad (12.10)$$

olup, serbestlik derecesi=(n-1)'dir.

X Y çarpımlar toplamı (XYÇT) da aşağıda verilmiştir:

$$XYÇT = \sum_{i=1}^{n} \left(x_i - \overline{X}\right)\left(y_i - \overline{Y}\right) = \sum x_i y_i - n\overline{x}\,\overline{y}$$

$$= \sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n} = S_{xy} \qquad (12.11)$$

Eşitlik (12.9) ve Eşitlik (12.11)'den yararlanarak, regresyon katsayısı (b_1) aşağıdaki biçimde de yazılabilir:

$$b_1 = \frac{XYÇT}{XOAKT} = \frac{S_{xy}}{S_{xx}} \qquad (12.12)$$

Regresyon kareler toplamı (RKT)

Her bir x_i değerinin, $\hat{y}_i$ değeri ile $\overline{Y}$ arasında bir fark vardır. Bu farkların kareler toplamı, regresyon kareler toplamı olarak adlandırılır.

$$RKT = \sum_{i=1}^{n} \left(\hat{y}_i - \overline{Y}\right)^2$$

$$= \frac{\left(\sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n}\right)^2}{\sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n}} = b_1 S_{xy} \qquad (12.13)$$

Eşitlik (12.13), Eşitlik (12.14) biçiminde de verilebilir.

$$RKT = \frac{(XYÇT)^2}{XOAKT} \tag{12.14}$$

ya da

$$RKT = b_1 XYÇT \tag{12.15}$$

RKT'ye ilişkin serbestlik derecesi 1'dir.

Regresyondan ayrılışlar kareler toplamı (RAKT)

Her bir x_i değerinin, y_i değeri ile $\hat{y}_i$ değeri arasında bir fark vardır. Bu farkların kareler toplamı, regresyondan ayrılışlar kareler toplamı olarak adlandırılır.

$$RAKT = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

dır. Ayrıca RAKT,

$$RAKT = YOAKT - RKT \tag{12.16}$$

olarak da yazılabilir. Kareler toplamı arasındaki bağıntı serbestlik derecesi arasında da vardır. RAKT'ye ilişkin serbestlik derecesi, (n-1)-1 olmak üzere n-2 'ye eşittir.

Kareler Ortalaması

Kareler ortalaması, kareler toplamının serbestlik derecesine bölümünden elde edilir. RAKO,

$$RAKO = \frac{RAKT}{n-2} \tag{12.17}$$

ve RKO,

$$RKO = \frac{RKT}{1} \tag{12.18}$$

dır.

Doğrusal regresyon denkleminin varyansı;

$$S^2 = RAKO \tag{12.19}$$

dır.

Belirtme katsayısı (R^2)

Yüzde cinsinden ifade edilen belirtme katsayısı, regresyon çözümlemesinde çok önemlidir ve aşağıda verildiği biçimde bulunur:

$$R^2 = \frac{RKT}{YOAKT} = 1 - \frac{RAKT}{YOAKT} \tag{12.20}$$

Belirtme katsayısı, $0 \leq R^2 \leq 1$'dir. R^2 bire yakın bulunursa, bağımlı değişkendeki değişimim büyük bir kısmı bağımsız değişkenler tarafından açıklanabilmektedir. Belirtme katsayısı, Y'deki değişimin % ...sı, X tarafından açıklanmaktadır diye yorumlanabilir.

Regresyon modelinin yeterliliğinin belirlenmesi için, hipotez testi ve güven aralıklarının oluşturulması gerekir.

Regresyon katsayısının önem kontrolü için aşağıda verilen adımlar uygulanır.

1- Hipotez kurulur.

$H_{0:}$ $\beta_1=0$ (Regresyon katsayısı önemsizdir)

$H_{A:}$ $\beta_1 \neq 0$ (Regresyon katsayısı önemlidir)

Regresyon katsayısının önemsiz olması demek; örneklemin çekildiği kitlede, bağımsız değişkende bir birimlik değişimin, bağımlı değişkende değişiklik yaratmayacağı anlamına gelir.

2- t test istatistiği hesaplanır.

$$t_H = \frac{b_1 - \beta_1}{s_{b_1}} \tag{12.21}$$

Burada, $s_{b_1} = \sqrt{\dfrac{RAKO}{XOAKT}}$ dir.

3- Serbestlik derecesi (n-2) ve α anlamlılık düzeyinde, $t_{\alpha/2;\, n-2}$ tablo değeri bulunur. $t_H > t_{\alpha/2;\, n-2}$ ya da $t_H < -\, t_{\alpha/2;\, n-2}$ ise H_0 reddedilir, diğer durumda reddedilemez.

4- Regresyon katsayısının önemli olup olmadığına ilişkin yorum yapılır.

Örnek 12.6: Rasgele seçilen 20 personelin aldığı uyarı yazıları ile performans değerleri arasındaki ilişki araştırılmak isteniyor. 20 personel üzerinde yapılan bir araştırmada alınan sonuçlar aşağıda verilmiştir: Burada X alınan uyarıları, Y performans değerlendirmeleri göstermektedir.

$$n=20\ ; \quad \sum_{i=1}^{20} x_i = 96; \quad \sum_{i=1}^{20} y_i = 236\ ; \quad \sum_{i=1}^{20} x_i^2 = 616\ ; \quad \sum_{i=1}^{20} y_i^2 = 2824; \quad \sum_{i=1}^{20} x_i y_i = 1075$$

a- Yukarıdaki verilere göre regresyon denklemini oluşturunuz.

$$\text{XOAKT} = 616 - \frac{96^2}{20} = 155{,}2$$

$$\text{YOAKT} = 2824 - \frac{236^2}{20} = 39{,}2$$

$$\text{XYÇT} = 1075 - \frac{96 * 236}{20} = -57{,}8$$

$$b_1 = \frac{-57{,}8}{155{,}2} = -0{,}37$$

$$b_0 = \frac{236}{20} - (-0{,}37)\frac{96}{20} = 13{,}57$$

Regresyon modeli,

$$\hat{Y} = 13{,}57 - 0{,}37X$$

olacaktır. Uyarıda yapılacak bir birimlik değişiklik, performans değerleri üzerinde etkilidir.

b- Regresyon katsayısının önem kontrolünü 0,05 anlamlılık düzeyinde yapınız.

1- $H_0: \beta_1 = 0$ (Regresyon katsayısı önemsizdir)

$H_A: \beta_1 \neq 0$ (Regresyon katsayısı önemlidir)

2- $$\text{RKT} = \frac{(\text{XYÇT})^2}{\text{XOAKT}} = 21{,}5$$

RAKT=YOAKT-RKT=39,2-21,5=17,7

RAKO=17,7/18=0,98

$$s_{b_1} = \sqrt{\frac{\text{RAKO}}{\text{XOAKT}}} = \sqrt{\frac{0{,}98}{155{,}2}} = 0{,}079$$

$$t_H = \frac{b_1 - \beta_1}{s_{b_1}} = \frac{-0{,}37 - 0}{0{,}079} = -4{,}683$$

3- 18 serbestlik derecesinde ve 0,05 anlamlılık düzeyinde tablo değeri, $t_{0,025;18}$=2,10 t_H<- $t_{0,025;18}$ olduğundan H_0 reddedilir.

4- **Yorum:** β_1 regresyon katsayısı önemlidir. Personelin aldığı bir uyarı mektubunun performans değerinde -0,37 birim azalma yaptığını söyleyebiliriz.

Doğrusallıktan ayrılışın önem kontrolü

Gözlenen noktaları tahmin edecek doğrunun geçerli olması için, gözlenen noktaların doğru ile tahmin edilecek noktalardan ayrılışlarının test edilmesi gerekir. YOAKT=RAKT+RKT eşitliğinde, RAKT'nin RKT'ye göre, belirlenen olasılıklar sınırında küçük olması tahmin doğrusunun geçerli olabilmesi için gerekir. Bu amaçla test edilecek hipotez bu uzunluklar dikkate alınarak kurulur ve uygun test dağılımı kullanılır. Doğrusallıktan ayrılışın önem kontrolü için RAKO ve RKO varyanslarının oranı uygun test istatistiğidir. İki varyansın oranı F dağılımına yakınsayacağı için kullanılacak test dağılımı F'dir. Hipoteze ilişkin adımlar aşağıda verilmiştir.

1- Hipotez kurulur.

H_0: Model geçersizdir (Gözlenen noktaların regresyon doğrusuna uyumu önemsizdir).

H_A: Model geçerlidir (Gözlenen noktalar regresyon doğrusu ile tanımlanabilir).

2-Gerekli olan kareler toplamları, kareler ortalamaları ve serbestlik derecelerini içeren varyans analizi tablosu düzenlenir.

Varyasyon kaynağı	**sd**	**KT**	**KO**
Regresyon	1	RKT	RKT/1
Hata	n-2	RAKT	RAKT/n-2
Toplam	n-1	YOAKT	-

3- $F_H = \dfrac{RKO}{RAKO}$ değeri hesaplanır.

4- (n-2) serbestlik derecesinde ve belirlenen α anlamlılık düzeyinde, $F_{\alpha;1;n-2}$'deki tablo değeri bulunur. $F_H > F_{\alpha;1;n-2}$ ise H_0 reddedilir, diğer durumda reddedilemez.

5- Model geçerliliği hakkında yorum yapılır.

Örnek 12.7: Örnek 12.6'daki veriler için, a- R^2'yi bulunuz. b- 0,05 anlamlılık düzeyinde doğrusallıktan ayrılışlar önem kontrolünü yapınız.

a- $R^2 = 1 - \frac{RAKT}{YOAKT} = 1 - \frac{17,7}{39,2} = 0,55$

Yorum: Bağımlı değişkendeki değişmelerin %55'i modeldeki bağımsız değişken tarafından açıklanmaktadır.

b-

1- H_0: Model geçersizdir.

H_A: Model geçerlidir.

2-

Varyasyon kaynağı	sd	KT	KO
Regresyon	1	21,5	21,5
Hata	18	17,7	0,98
Toplam	19	39,2	-

3- $F_H = \frac{RKO}{RAKO} = \frac{21,5}{0,98} = 21,98$

4- $F_{0,05;1;18} = 4,41$ $F_H > F_{\alpha;1;n-2}$ olduğundan H_0 reddedilir.

5- **Yorum:** 0,05 anlamlılık düzeyinde modelin geçerli olduğu söylenebilir. Personelin aldığı uyarı yazıları ile performans değerleri arasında bulunan $\hat{Y} = 13,57 - 0,37X$ denklemi gözlenen noktalar için uygun bir modeldir.

Basit Doğrusal Regresyonda Aralık Tahmini

a- Regresyon katsayısı için güven aralığı;

$$P\left(b_1 - t_{\alpha/2;n-2}S_{b_1} < \beta_1 < b_1 + t_{\alpha/2;n-2}S_{b_1}\right) = 1 - \alpha \qquad (12.22)$$

dır.

Örnek 12.8: Örnek 12.6'daki verilere ilişkin regresyon katsayısının 0,95 güven düzeyinde güven aralığını bulunuz.

$$P(-0,37 - 2,10.0,079 < \beta_1 < -0,37 + 2,100,079) = 0,95$$

$$P(-0,53 < \beta_1 < -0,21) = 0,95$$

0,95 güven düzeyinde regresyon katsayının güven aralığı -0,53 ile -0,21 arasındadır.

b- Bilinen bir x_p değerine karşılık y değerinin ortalaması için güven aralığı verilebilir. x_p bilindiğinde, y değerinin ortalamasını tanımlayan ifade $E(Y/x_p)$ olup güven aralığı;

$$P\left(\hat{y} - t_{\alpha/2;n-2}\left[RAKO\left(\frac{1}{n}+\frac{(x_p-\overline{X})^2}{XOAKT}\right)\right]^{1/2} < E(Y/x_p) < \hat{y} + t_{\alpha/2;n-2}\left[RAKO\left(\frac{1}{n}+\frac{(x_p-\overline{X})^2}{XOAKT}\right)\right]^{1/2}\right) = 1-\alpha \tag{12.23}$$

dır.

Örnek 12.9: Örnek 12.6'daki veriler için, x_p=5 değerine karşılık gelen y değerinin ortalaması için 0,95 güven düzeyinde güven aralığını bulunuz.

Verilere ilişkin regresyon modeli: $\hat{Y} = 13{,}57 - 0{,}37X$.

x_p=5 ise y tahmin değeri 11,72'dir.

$$P\left(11{,}72 - 2{,}10\left[0{,}98\left(\frac{1}{20}+\frac{(5-4{,}8)^2}{155{,}2}\right)\right]^{1/2} < E(Y/x_p) < 11{,}72 + 2{,}10\left[0{,}98\left(\frac{1}{20}+\frac{(5-4{,}8)^2}{155{,}2}\right)\right]^{1/2}\right) = 0{,}95$$

$$P(11{,}72 - 0{,}4661 < E(Y/x_p) < 11{,}72 + 0{,}4661) = 0{,}95$$

$$P(11{,}25 < E(Y/x_p) < 12{,}18) = 0{,}95$$

0,95 güven düzeyinde x_p=5 değerine karşılık gelen y değerinin ortalamasının güven aralığı 11,25 ile 12,18 arasındadır.

c- Bilinen bir x_p değerine karşı gelen herhangi bir y'nin değerini tahmin etmek regresyon çözümlemesinde önemlidir. X raslantı değişkeninin x_p gibi bir değeri verildiğinde, Y raslantı değişkeninin herhangi bir değeri için y/x_p olup güven aralığı;

$$P\left(\hat{y} - t_{\alpha/2}\left[RAKO\left(1+\frac{1}{n}+\frac{(x_p-\overline{X})^2}{XOAKT}\right)\right]^{1/2} < Y/x_p < \hat{y} + t_{\alpha/2}\left[RAKO\left(1+\frac{1}{n}+\frac{(x_p-\overline{X})^2}{XOAKT}\right)\right]^{1/2}\right) = 1-\alpha \tag{12.24}$$

dır.

Örnek 12.10: Örnek 12.6'daki veriler için, x_p=5 değerine karşılık gelen y değerinin 0,95 güven düzeyinde güven aralığını bulunuz.

$$P\left(11{,}72-2{,}10\left[0{,}98\left(1+\frac{1}{20}+\frac{(5-4{,}8)^2}{155{,}2}\right)\right]^{1/2}<Y/x_p<11{,}72+2{,}10\left[0{,}98\left(1+\frac{1}{20}+\frac{(5-4{,}8)^2}{155{,}2}\right)\right]^{1/2}\right)=0{,}95$$

$$P(11{,}72-2{,}13<Y/x_p<11{,}72+2{,}13)=0{,}95$$

$$P(9{,}59<Y/x_p<13{,}85)=0{,}95$$

0,95 güven düzeyinde, x_p=5 değerine karşılık gelen y değerinin güven aralığı 9,59 ile 13,85 arasındadır.

PROBLEMLER

1- 10 kişinin yaşları ve kandaki kolesterol miktarları aşağıda verilmiştir.

Yaş	:	44	46	48	55	60	62	57	47	50	45
Kolesterol	:	172	179	185	190	209	211	200	173	184	176

a- Kolesterolün yaş ile olan ilişkisini hesaplayınız. Aralarındaki doğrusal regresyon modelini kurunuz.

b- Kurduğunuz denklemden nasıl yararlanabilirsiniz?

2- X ve Y değişkenleri için aşağıdaki bilgiler verilmiştir.

$$\sum_{i=1}^{20} x_i = 2015 \qquad \sum_{i=1}^{20} y_i = 990 \qquad \sum_{i=1}^{20} x_i^2 = 204053 \qquad \sum_{i=1}^{20} y_i^2 = 49088 \qquad \sum_{i=1}^{20} x_i y_i = 99884$$

a- Korelasyon katsayısını bulunuz ve yorumlayınız

b- Korelasyon katsayısının önem kontrolünü yapınız. α=0,05 alınız.

c- Modelin geçerliliğini (Doğrusallıktan ayrılış) α=0,05 için test ediniz.

3- 11 öğrenci bulunan bir sınıfta sınav yapılıyor. Sınav kağıtları ayrı ayrı öğretmen tarafından değerlendiriliyor. Birinci öğretmenin değerlendirmesi X, ikinci öğretmenin değerlendirmesi Y olmak üzere, aşağıdaki bilgilere ulaşılmıştır.

X:	9	5	6	5	3	4	4	5	6	8	8
Y:	9	8	4	6	5	3	5	5	5	7	5

a- Korelasyon katsayısını bulunuz ve yorumlayınız.

b- Korelasyon katsayısının önem kontrolünü yapınız. α=0,05 alınız.

c- Regresyon modelini oluşturunuz.

d- Birinci öğretmenin 7 olarak değerlendirdiği bir notu, ikinci öğretmenin kaç olarak değerlendireceğini bulunuz.

4- Aşağıdaki bilgilerden yararlanarak,

a- İlişki katsayısını bulunuz.

b- Regresyon denklemini oluşturunuz.

c- b_1 katsayısının önem kontrolünü 0,05 anlamlılık düzeyinde yapınız.

$$n = 6 \quad \sum_{i=1}^{6} x_i = 40{,}71 \quad \sum_{i=1}^{6} y_i = 31{,}59 \quad \sum_{i=1}^{6} x_i^2 = 276{,}23 \quad \sum_{i=1}^{6} y_i^2 = 169{,}64 \quad \sum_{i=1}^{6} x_i y_i = 214{,}52$$

5- Bir işyerindeki 6 çalışan, iki farklı amir tarafından ayrı ayrı değerlendiriliyor. Birinci amirin değerlendirmesi X, ikinci amirin değerlendirmesi Y olmak üzere, aşağıdaki bilgilere bulunuyor.

X:	7	5	7	5	3
Y:	6	5	5	4	4

a- Korelasyon katsayısını bulunuz ve yorumlayınız.

b- 0,05 anlamlılık düzeyinde korelasyon katsayısının önem kontrolünü yapınız.

c- Regresyon modelini oluşturunuz ve 0,05 anlamlılık düzeyinde geçerliliğini test ediniz.

d- Birinci amirin 8 olarak değerlendirdiği bir kişi, ikinci amirin kaç olarak değerlendirebileceğini bulunuz.

6- Bir iş yerinde çalışan işçi sayısı ve bir yıl süresince üretilen mal miktarları 6 yıl için gözlenmiş aşağıdaki değerler bulunmuştur.

İşçi Sayısı	(1000 Kişi):	1	2	5	5,5	6	5,8
Mal Miktarları	(1000 Ton):	4	7	10	11	12	11,7

a- Doğrusal regresyon denklemini oluşturunuz.

b- Regresyon katsayısının önem kontrolünü 0,05 anlamlılık düzeyinde yapınız.

c- Korelasyon katsayısını hesaplayınız.

7- 9 yıl için bir malın fiyatı ile talep miktarları verilmiştir. Aşağıdaki verilere göre talep miktarı ile fiyat arasındaki regresyon denklemini oluşturunuz. Korelasyon katsayısını bulunuz.

Malın Fiyatı (100 YTL):	2	2	4	6	7	8	8,5	9	8,8
Talep Miktarları (1000 Kg):	8	9	7,5	7	6,8	6	6,2	5	5,4

8- 8 tür şarap, iki uzman tarafından 10 üzerinden değerlendirilmiştir. Uzmanların kararları arasındaki ilişki miktarı nedir?

Şaraplar	A	B	C	D	E	F	G	H
I Uzman1	10	8	7	6	7	4	4	9
II Uzman	8	7	10	6	5	5	6	8

9- Yapılan reklam kampanyası ile gelir arasında bir ilişki varsa firma yönetimi reklam harcamalarını artırmayı düşünmektedir. Bu amaçla son bir yıl içinde aylık olarak yaptığı reklam harcaması ve gelirler aşağıda özetlenmiştir. Harcamalar ile gelirler arasında pozitif doğrusal bir ilişki olup olmadığını 0,10 anlamlılık düzeyi ile test ediniz.

Aylar	1	2	3	4	5	6	7	8	9	10	11	12
Harcama (1000YTL)	1,2	1,3	1,0	2,4	1,8	2,0	1,8	1,8	2,2	2,6	2,7	2,6
Gelir (1000 YTL)	7,0	7,3	6,5	6,5	7,0	7,2	7,4	7,7	7,6	7,5	8,1	8,0

10- Bir sigorta şirketi 9 aylık dönemde yeni yapılan poliçelerin gelirler üzerindeki etkilerini incelemek istemektedir. Aşağıdaki bilgilere göre, poliçe sayısı ile gelir arasındaki ilişki katsayısını bulunuz ve 0,05 anlamlılık düzeyinde önem kontrolünü yapınız.

Poliçe sayısı	: 10	20	22	24	25	25	26	27	27
Gelir (milyon)	:10	35	24	30	34	34	35	30	28

11- 6 aileye ait bir aylık gelir ve sağlık harcamalarına ilişkin veriler aşağıdadır.

Harcama (YTL)	20	30	40	50	55	60
Gelir (YTL)	500	600	612	400	645	625

a- Verilere ilişkin regresyon modelini bulunuz ve modelin geçerliliğini 0,01 anlamlılık düzeyinde test ediniz.

b- Geliri 650 olan bir ailenin sağlık harcaması ne olabilir?

12- Bir araştırmacı, çalışma süresi (X) ile başarı puanı (Y) arasındaki ilişki miktarını belirlemek amacıyla, rasgele seçtiği 10 kişinin değerlerini aşağıdaki biçimde kaydetmiştir.

X (saat):	1	1	2	2	3	3	4	4	5	5
Y (puan):	38	64	86	51	71	84	75	78	60	80

Verilere ilişkin ilişki miktarını ve regresyon modelini oluşturunuz. Önem kontrollerini 0,05 anlamlılık düzeyinde yapınız.

13- Bebeklerin boy uzunluğunda, annenin karın çevresinin etkili olduğu sanılmaktadır. Bu iki değişken arasındaki ilişki miktarını belirlemek amacıyla 11 bebek rasgele seçiliyor ve aşağıdaki bulgulara ulaşılıyor.

X:	99	96	100	97	106	92	109	100	99	105	104
Y:	52	48	47	45	47	50	53	50	52	53	54

a- Verilere ilişkin regresyon modelini kurunuz ve 0,01 anlamlılık düzeyinde geçerliliğini test ediniz.

b- Boy uzunluğu 95 olan bir bebeğin annesinin karın çevresinin ne olabileceğini hesaplayınız.

14- Örneklem büyüklüğü 20 olan bir örneklem için regresyon katsayısı –0,74'dür. 0,95 güven düzeyinde korelasyon katsayısının güven aralığını bulunuz.

15- Bir şirket sahibi, çalışanlarının mali durumunu araştırmak istiyor. Bu amaçla şirketinde çalışan 10 kişiyi rasgele seçiyor ve ailedeki birey sayısı ile bir ayda harcadıkları para miktarını (YTL) aşağıdaki biçimde kaydediyor.

Birey sayısı (X):	4	2	3	1	2	3	4	5	5	4
Harcanan para (Y) :	1480	960	1950	880	650	1100	1200	2700	1750	1250

a- Ailedeki birey sayısı ile harcanan para arasındaki ilişki miktarını bulunuz ve 0,05 anlamlılık düzeyinde önem kontrolünü yapınız.

b- Regresyon modelini oluşturunuz ve modelin geçerliliğini 0,05 anlamlılık düzeyinde test ediniz.

16- Üniversiteye giriş puanlarının, üniversiteden mezun puanlarını etkilediği düşünülmektedir. Bir üniversitenin istatistik bölümüne giren 120 öğrencinin giriş ve mezun olduğu puanlar aşağıda verilmiştir.

Giriş puanı :	2,5	3,2	3,5	2,8	3,0	2,4	3,4	2,9	2,8	3,8
Mezun ol.puan :	650	700	710	550	625	490	700	655	589	699

a- Üniversiteye giriş puanı ile mezun olduğu puan arasındaki ilişki miktarını bulunuz ve 0,05 anlamlılık düzeyinde önem kontrolünü yapınız.

b- Regresyon modelini oluşturunuz ve modelin geçerliliğini 0,05 anlamlılık düzeyinde test ediniz.

17- 8 hastanın yaşları ve tedavi sonrası iyileşme süreleri aşağıda verilmiştir.

X (yıl):	20	25	26	25	26	23	32	35	38	39
Y(gün):	9	9	10	8	20	25	30	25	26	35

Verilere ilişkin regresyon modelini oluşturunuz. Yaşı 24 olan hastanın, tedavi sonrasında kaç günde iyileşebileceğini bulunuz.

18- Soru 17'deki veriler için, $x_p=24$ değerine karşılık gelen y değerinin ortalaması için 0,95 güven düzeyinde güven aralığını bulunuz.

19- Soru 15'deki verilere ilişkin,

a- Regresyon katsayısının 0,95 güven düzeyinde güven aralığını bulunuz.

b- $x_p=6$ değerine karşılık gelen y değerinin 0,95 güven düzeyinde güven aralığını bulunuz.

20- Soru 16'daki verilere ilişkin,

a- Regresyon katsayısının 0,95 güven düzeyinde güven aralığını bulunuz.

b- $x_p=2,6$ değerine karşılık gelen y değerinin 0,95 güven düzeyinde güven aralığını bulunuz.

SEÇİLMİŞ BAZI PROBLEMLERİN YANITLARI

BİRİNCİ BÖLÜM

2. **a.** 2000 sayfalık 3 ciltlik bir ansiklopedide sayfaların tümü **b.** Rasgele seçilen 100 sayfa **c**. Basit rasgele örnekleme yöntemi kullanılmıştır. **d.** Hatalı sözcük sayısı nicel değişkendir.

4. **a.** Çevre kirliliğinin yoğun olduğu bir bölgede yaşayan toplam 20000 kişi
b. Tabakalı örnekleme **c.** Bu bölgede yaşayan kişiler **d.** Yaşadığı bölge, eğitimi, evinizde TV var mı? Çevrenizi kirli buluyor musunuz? Aylık geliriniz nedir?
e. Yüz yüze görüşme anket uygulayarak **f**. Sıralanmış (ordinal)

6. **a.** Araştırmaya konu olan, nüfusu 4 milyon olan bir yerleşim bölgesinde yaşayan 12 yaş ve üstündeki kişilerin tümü **b.** Basit rasgele örnekleme **c.** Orantılı **d.** Kesikli
e. Sürekli **f.** Kesikli

8. Verilerin doğru zaman ve yerden toplanmaması bilgilerde bir yanlılık ortaya çıkarır. Örneğin çalışan bayanlarla ilgili bir çalışmada doğru bilgi toplama zamanının seçilmemesi gibi

10. **a.** Orantılı ölçüm türüdür.
b. Hesaplayamayız çünkü 4 ve daha fazla gibi bir seçenek olup, açık uçlu bir cevap seçeneği vardır.

12. **a.** Eğer örneklemeyi gün içinde markete gelen müşteriler içinden oluşturacaksak olasılığa bağlı olmayan örnekleme türlerinden keyfi örnekleme kullanırız. Eğer belli günlerde ya da saatlerde müşteri niteliği ve çokluğu değişiyor ve bu durum devamlılık gösteriyor ise olasılığa bağlı olmayan örnekleme türlerinden kota örneklemesi kullanılır.
b. Saha araştırmalarından yazılı anketlerin doldurulması ile bilgi elde ederiz.

14. **a.** Sürekli nicel **b.** Sıralanabilir nitel **c.** Sınıflanabilir nitel **d.** Sürekli nicel

16. Bu konuda öğrencinin kendi yaratıcılığı ile farklı formlar hazırlanabilir. Burada basit bir form örneği aşağıda verilmiştir:

İade edilen ürünün adı	Seri No	Üretim Tarihi	İadenin alındığı tarih	İade nedeni

18. Bu konuda öğrencinin kendi yaratıcılığı ile farklı formlar hazırlanabilir. Burada basit bir form örneği aşağıda verilmiştir:

Tarih:

Yemek Öğlen, …… akşam……..

Yemek Çeşitleri							
1.çeşit	2.çeşit	3.çeşit	4.çeşit	5.çeşit	6.çeşit	7.çeşit	8.çeşit

.

1.çeşit içki	2. çeşit içki	3. çeşit içki	4. çeşit içki	5. çeşit içki

Çizelgeler her müşteri için işaretlenerek oluşturulur.

20. **a.** Bir işyeri sınavına başvuran kişiler **b.** Basit rasgele örnekleme yöntemi

c. Aldığı puan nicel, veri yapısı sürekli; cinsiyet nitel, veri yapısı kesikli **d.** Aralıklı

İKİNCİ BÖLÜM

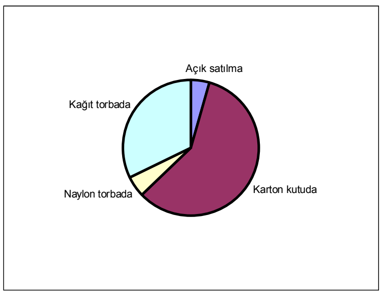
2. **a.**

				Birikimli Sıklıklar	
Alt Sınır-Üst Sınır	**S_i**	**f_i**	**p_i**	**...Ve Az**	**...Ve Çok**
2-5	3.5	4	0.12	4	30
6-9	7.5	6	0.20	10	26
10-13	11.5	12	0.40	22	20
14-17	15.5	2	0.07	24	8
18-21	19.5	2	0.07	26	6
22-25	23.5	2	0.07	28	4
26-29	27.5	2	0.07	30	2

b. Öğrencilerin % 76'sı 15 puandan daha az puan almıştır

4.

Sınıflar	**f_i**	**p_i**	**P_i*360**
Açık satılma	30	0.03	10.8
Karton kutuda	400	0.40	144
Naylon torbada	350	0.35	126
Kağıt torbada	220	0.22	79.2



6.

Sınıflar	f_i	p_i
Sigara içmiyor	100	0.50
Az sigara içiyor	60	0.30
Çok sigara içiyor	40	0.20

8.

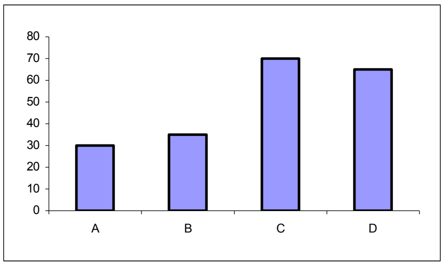
Notlar	**Sıklık**	**Göreli sıklık**	**Ve daha az**	**Ve daha çok**
F	40	0,40	40	100
D	20	0,20	60	60
C	12	0,12	72	40
B	24	0,24	96	28
A	4	0,04	100	4

10. **a.** Sınıflanabilen nitel

b.

Sınıflar	f_i	p_i
A	30	0.150
B	35	0.175
C	70	0.350
D	65	0.325

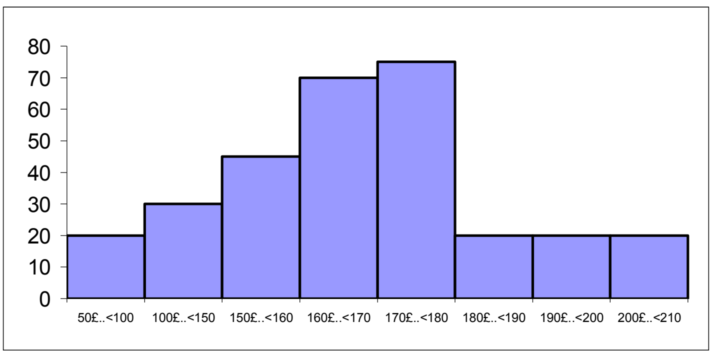
c.



12.

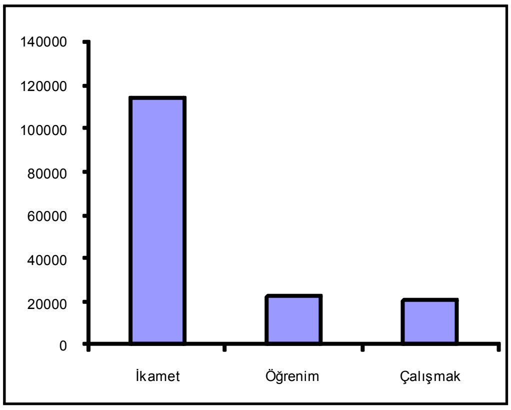
Alt Sınır-Üst Sınır	S_i	f_i	**...Ve Az**	**...Ve Çok**
2-4	3	6	6	33
5-7	6	4	10	27
8-10	9	4	14	23
11-13	12	2	16	19
14-16	15	6	22	17
17-19	18	2	24	11
20-22	21	9	33	9

14.



16.

Sınıflar	f_i	p_i
İkamet	114315	0.73
Öğrenim	22107	0.14
Çalışmak	21140	0.13



18.

Alt Sınır-Üst Sınır	S_i	f_i	...Ve Az	...Ve Çok
18-26	22	7	7	26
27-35	31	2	9	19
36-44	40	5	14	17
45-53	49	6	20	12
54-62	58	4	24	6
63-71	67	2	26	2

20. Bu amaçla ve daha çok ya da ve daha az göreli sıklıkları bulunur. Bu değerlerde %50'nin üstünde olan değerler için sıklık dağılımında olan o değere karşıt gelen yaşlar bize bir fikir verir.

ÜÇÜNCÜ BÖLÜM

1. **a.** $\overline{X} = 7,125$; Bu verilerde tepe değerinden söz edilemez **;** $\overline{X}' = 7$

b. Dağılımın tepe değeri yoktur. Ortalama ve ortanca eşit oldukları için dağılım simetriktir. **c.** Q_1=4+0,5=4,5

3. **a.** $\overline{X} = 1,125$; $\overline{X}' = 1$; $\hat{X}_1 = 1$; $\hat{X}_1 = 3$

b. <u>5 sınıf için</u>

A_s	$Ü_s$	S_i	f_i	$S_i * f_i$	S_i^2	$S_i^2 * f_i$
-6,0	-3,7	-4,85	4	-19,4	23,5225	94,09
-3,6	-1,3	-2,45	4	-9,8	6,0025	24,01
-1,2	1,1	-0,05	15	-0,75	0,0025	0,0375
1,2	3,5	2,35	9	21,15	5,5225	49,7025
3,6	6	4,80	8	38,4	23,04	184,32
Toplam				29,6		352,16

<u>8 sınıf için</u>

A_s	$Ü_s$	S_i	f_i	$S_i * f_i$	S_i^2	$S_i^2 * f_i$
-6	-4,6	-5,3	2	-10,6	28,09	56,18
-4,5	-3,1	-3,8	2	-7,6	14,44	28,88
-3	-1,6	-2,3	4	-9,2	5,29	21,16
-1,5	-0,9	-1,2	2	-2,4	0,81	2,88
0	1,4	0,7	13	9,1	0,49	6,37
1,5	2,9	2,2	0	0	0	0
3	4,4	3,7	12	44,4	13,69	164,28
4,5	6	5,2	5	26	27,04	135,2
Toplam			$\sum_{i=1}^{8} f_i * S_i = 49,7$			414,24

c. 8 sınıf için $\overline{X} = 1,24$; $S^2 = 9,04$

5 sınıf için $\overline{X} = 0,74$; $S^2 = 8,46$

d. Sınıflandırılmamış verilerde bulunan ortalama=1,125 ve varyans=9,95 dir.

8 sınıf yapıldığında ortalama=1,24 varyans=9,04

5 sınıf yapıldığında ortalama=0,74 varyans=8,46

Yukarıda verilen çözümde görüldüğü gibi verileri yerleştireceğimiz sınıf sayısı azaldıkça ortalama ve varyansda sınıflandırılmamış verilerden uzaklaşma artmıştır.

5. **a.**D **b.**Y **c.**D **d.**Y

7. **a.** $\overline{X} = 4,42$ $S^2 = 4,401$ S= $\sqrt{4,401}$ =2,098 $\hat{X} = 5,1$

b. $\overline{X}'$ =4 dak.

9. **a.**

A_s	$Ü_s$	S_i	f_i	$S_{i}*f_i$	S_i^2	$S_i^2*f_i$...	**…ve daha az**
2	6	4	4	16	16	64	4
6	10	8	6	48	64	384	10
10	14	12	12	144	144	1728	22
14	18	16	2	32	256	512	24
18	22	20	2	40	400	800	26
22	26	24	2	48	576	1152	28
26	30	28	2	56	784	1568	30
Toplam			30	384		6208	

b. $\overline{X} = 12{,}8$ $S^2 = 1292{,}8$ S=35,95 **c.** $\overline{X}'$=11,67 dak. $\hat{X} = 11{,}43$ litre

11. **a.**

A_s	$Ü_s$	S_i	f_i	$S_{i}*f_i$	S_i^2	$S_i^2*f_i$...
35	40	37,5	4	150	1406,25	5625
40	45	42,5	3	127,5	1806,25	5418,75
45	50	47,5	4	190	2256,25	9025
50	55	52,5	5	262,5	2756,25	13781,25
55	60	57,5	1	57,5	3306,25	3306,25
60	65	62,5	3	187,5	3906,25	11718,75

b $\overline{X} = 48{,}75$ $S^2 = 70{,}72$ S=8,41

c Sapan değer yoktur

13. **a**. DG= 34-10=24

b. $\overline{X} = (10+11+11+\ldots+34)/20 = 21{,}4$ $MS = 7{,}1$ $S^2 = 65{,}72$

c. Her ikisi de değişim ölçüsüdür. Ancak varyans ortalamadan olan uzaklıkların kareleri toplamını verdiği için sayısal olarak daha büyük bir rakamdır. Varyansın kare kökü MS'ye yakın bir değer vermektedir.

15. I. örneklem daha dar alanda dağılım gösteriyor simetrisi biraz bozulmuş ortalaması II.örneklemden küçük, II. örneklem simetrik ve daha geniş alanda dağılım gösteriyor.

17. $\sum_{i=1}^{10} |X_i - \overline{X}| = 8$ MS=8/10=0,8 $S^2 = (-1^2+4^2+\ldots+3^2)/9 = 5{,}5$

19. $CV = V = \frac{S}{\overline{X}} \times 100$

I.bölge için; 4/4200*100=0,0952

II.bölge için; 0,7/3,9*100=17,94

olup, ikinci bölgede kız çocuklarının ağırlık dağılımı daha heterojendir.

DÖRDÜNCÜ BÖLÜM

1. **a.** 5×4×3=60 **b.** 1×4×3=12 **c.** 3×4×2=24 **d.** 3×4×3=36 **e.** 3×4×1=12

3. **a.** 5!=120 **b.** 2×2! ×3!=24 **c.** 2! ×4!=48

5. **a.** $\binom{30}{20} = \frac{30!}{20!\,10!} = 30045015$ **b.** $\binom{2}{2}\binom{28}{18} = \frac{28!}{18!\,10!} = 13123110$

c. $\binom{10}{3}\binom{20}{17} = 136800$ **d.** $\binom{10}{3}\binom{20}{17} + \binom{10}{4}\binom{20}{16} + \ldots + \binom{10}{10}\binom{20}{10}$ **e.** $2\binom{28}{19}$

7. **a.** $P(KKK) = \frac{6}{15} \times \frac{5}{14} \times \frac{4}{13} = 0{,}044$ **b.** $P(BBB) = \frac{9}{15} \times \frac{8}{14} \times \frac{7}{13} = 0{,}184$

P(BBB)+ P(KKK)=0.184+0.044=0.228

9. **a.** $P(A) \times P(A) \times P(A) = \frac{6}{12} \times \frac{6}{12} \times \frac{6}{12} = \frac{1}{8}$ **b.** $3 \times \left(\frac{6}{12} + \frac{4}{12}\right) \times \frac{2}{12} \times \frac{2}{12} = \frac{15}{216} = 0{,}069$

c. $P(A) \times P(B) \times P(A) + P(B) \times P(A) \times P(B) = \frac{6}{12} \times \frac{4}{12} \times \frac{6}{12} + \frac{4}{12} \times \frac{6}{12} \times \frac{4}{12} = 0{,}138$

11. **a.** $P(A_1 \cap B_1) = \frac{105}{260} = 0{,}404$

b. $P(A_1 \cup B_1) = P(A_1) + P(B_1) - P(A_1 \cap B_1) = \frac{120}{260} + \frac{225}{260} - \frac{105}{260} = 0{,}923$

c. $P(B2) = \frac{35}{260} = 0{,}13$

13. $P(E/H) = \frac{P(E \cap H)}{P(H)} = \frac{0{,}15}{0{,}30} = 0{,}5$

15. **a.** $E(X) = 200 \times 0{,}40 + 250 \times 0{,}35 + 260 \times 0{,}25 = 232{,}5$ **b.** σ=26.8

17. **a.** $P(A) \times P(A) \times P(A) \times P(B) = \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} = \frac{1}{81} = 0{,}0123$

b. $P(A) \times P(A) \times P(A) = \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} = \frac{1}{27} = 0{,}037$

19. **a.** $P(\dot{I}) \times P(\overline{H}) = 0{,}85 \times (1 - 0{,}40) = 0{,}51$ **b.** $P(\dot{I} \cap T) = P(\dot{I}/T) \times P(T) = 0{,}85 \times 0{,}55 = 0{,}4675$

BEŞİNCİ BÖLÜM

2. **a.** P(X=3)=0.1536 **b.** P(X≥2)=0.3456+0.1536+0.0256=0.5248

4. **a.** $\mu=8$ $\sigma^2=1{,}6$ $\sigma=1{,}2649$

b. P(X>8)=P(X=9)+ P(X=10)= 0,2684+0,1074=0,3758

P(X<5)=P(X=0)+P(X=1)+P(X=2)+P(X=3)+P(X=4)=0,0+0,0+0,0001+0,0008+0,0055
=0,0064

P(2≤X<6)=P(X=2)+P(X=3)+P(X=4)+P(X=5)=0,0001+0,0008+0,0055+0,0264=0,0328

6. **a**. $P(\ 1\le X)=1-\binom{4}{0}0{,}10^0(0{,}90)^4=1-0{,}6561=0{,}3439$

b. $P(\ 1\le X)=1-\binom{4}{0}0{,}05^0(0{,}95)^4=1-0{,}8145=0{,}1855$

8. **a.** $\mu=1{,}5$ **b.** $\sigma^2=1{,}05$ $\sigma=\sqrt{npq}=\sqrt{1{,}05}=1{,}025$

10. **a.** $P(X=4)=0{,}1755$

b. P(X>3)=1-(P(X=0)+ P(X=1)+ P(X=2)+ P(X=3))=1-(0,0067+0,0337+0,0842+0,1404)=0,735

c. P(X≤2)= P(X=0)+ P(X=1)+ P(X=2)= 0,0067+0,0337+0,0842=0,1246

12. **a.** P(X=5)=0,0916

b. P(4≤X) =1-(P(X=0)+ P(X=1)+ P(X=2)+ P(X=3)+ P(X=4))=1-0,0996=0,9004

c. P(X≤6)=P(X=0)+P(X=1)+P(X=2)+P(X=3)+ P(X=4)+ P(X=5)+ P(X=6)=0,3133

14. $\mu=11{,}997$

16. $(5300<X)=P(\frac{5300-5000}{300}<Z)=P(1<Z)=0{,}5-0{,}3413=0{,}1587$

18. X=91,6

20. **a.** $P(1{,}68<Z<1{,}8)=P\left(\frac{1{,}68-1{,}75}{1{,}2}<Z<\frac{1{,}8-1{,}75}{1{,}2}\right)=P(-0{,}05<Z<0{,}04)=0{,}0199+0{,}016=0{,}0359$

b. $P(Z<1{,}82)=P\left(Z<\frac{1{,}82-1{,}75}{1{,}2}\right)=P(Z<0{,}05)=0{,}5+0{,}0199=0{,}5199$

ALTINCI BÖLÜM

2. **a.** $P(\overline{X}<3)=P\left(Z<\frac{3-4}{1/\sqrt{96}}\right)=P\left(Z<\frac{-1}{0{,}1020}\right)=P(Z<-3{,}16)=0{,}5-0{,}4990=0{,}001$

b. $P(3{,}5<\overline{X}<4{,}5)=P\left(\frac{3{,}5-4}{0{,}1020}<Z<\frac{4{,}5-4}{0{,}1020}\right)=P(-1{,}58<Z<1{,}58)=0{,}4429+0{,}4429=0{,}8858$

4. **a.** $P(66<\overline{X}<68)=P\left(\frac{66-68}{\sqrt{1{,}175}}<Z<\frac{4{,}5-4}{\sqrt{1{,}175}}\right)=P(-1{,}84<Z<0)=0{,}4671$

b. $P(\overline{X}>69)=P\left(Z>\frac{69-68}{\sqrt{1{,}175}}\right)=P(Z>0{,}92)=0{,}5-0{,}3212=0{,}1788$

6. **a.** $P(66{,}8 < \overline{X} < 68{,}3) = P\left(\frac{66{,}8-68}{\sqrt{0{,}357}} < Z < \frac{68{,}3-68}{\sqrt{0{,}357}}\right) = P(-2{,}03 < Z < 0{,}50) = 0{,}4788 + 0{,}1915 = 0{,}6703$

$80 \times 0.6703 \cong 54$

b. $P(\overline{X} < 66{,}4) = P\left(Z < \frac{66{,}4-68}{\sqrt{0{,}357}}\right) = P(Z < -2{,}71) = 0{,}5 - 0{,}4966 = 0{,}0034$

8. **a.** $E(\hat{p}) = p = 0{,}20 \quad Var(\hat{p}) = \sigma^2_{\hat{p}} = \frac{p(1-p)}{n} = \frac{0{,}20 \times 0{,}80}{130} = 0{,}0012$

b. $P(0{,}18 < \hat{p} < 0{,}22) = P\left(\frac{0{,}18-0{,}20}{\sqrt{0{,}0012}} < Z < \frac{0{,}22-0{,}20}{\sqrt{0{,}0012}}\right) = P(-0{,}57 < Z < 0{,}57) = 0{,}4314$

10. **a.** $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}.\sqrt{\frac{N-n}{N-1}} = \sqrt{\frac{0{,}45 \times 0{,}55}{144}}\sqrt{\frac{200-144}{200-1}} = 0{,}0219$

b. $P(\hat{p} > 0{,}50) = P\left(Z > \frac{0{,}50-0{,}45}{0{,}0219}\right) = P(Z > 2{,}28) = 0{,}5 - 0{,}4887 = 0{,}0113$

c. $P(\hat{p} < 0{,}36) = P\left(Z < \frac{0{,}36-0{,}45}{0{,}0219}\right) = P(Z < -4{,}11) = 0{,}5 - 0{,}4990 = 0{,}001$

12. **a.** $E(\hat{p}) = p = 0{,}12 \quad \sigma^2_{\hat{p}} = \frac{p(1-p)}{n}\frac{N-n}{N-1} = \frac{0{,}12 \times 0{,}88}{450} \times \frac{1000-450}{1000-1} = 0{,}000129$

b. $P(0{,}10 < \hat{p} < 0{,}15) = P\left(\frac{0{,}10-0{,}12}{\sqrt{0{,}000129}} < Z < \frac{0{,}15-0{,}12}{\sqrt{0{,}000129}}\right) = P(-1{,}76 < Z < 2{,}65) = 0{,}4608 + 0{,}4960 = 0{,}9568$

14. **a.** $P(S^2 < 154{,}72) = P\left(\frac{(n-1)S^2}{\sigma^2} < \frac{(4-1)154{,}72}{100}\right) = P(\chi^2_3 < 4{,}6416) = 0{,}80$

b. $P(78{,}87 < S^2 < 208{,}38) = P\left(\frac{3 \times 78{,}87}{100} < \chi^2_3 < \frac{3 \times 208{,}38}{100}\right) = P(2{,}3661 < \chi^2_3 < 6{,}2514) = 0{,}50 - 0{,}10 = 0{,}40$

16. $P(\overline{X} < 41{,}84683) = P\left(t_8 < \frac{41{,}84683-40}{5/3}\right) = P(t_8 < 1{,}1081) = 0{,}85$

18. $P(S^2 > ?) = 0{,}01 \Rightarrow P\left(\chi^2_{19} < \frac{? \times 9}{22}\right) = 0{,}01 \Rightarrow \frac{? \times 9}{22} = 27{,}2036 \Rightarrow ? = 31{,}49$

20. Müşteri isteği sağlanmıştır.

YEDİNCİ BÖLÜM

1. **a.** $\mu = \frac{\sum X}{N} = 24{,}41$ **b.** $\sigma^2 = \frac{\sum (X-\mu)^2}{N} = 78{,}57$ **c.** $P(X > 30) = \frac{3}{12}$

3. **a.** $E(\overline{X}_a) = E\left(\frac{X_1 + 2X_2 + 3X_3 + 4X_4}{10}\right) = \frac{1}{10}\mu + \frac{2}{10}\mu + \frac{3}{10}\mu + \frac{4}{10}\mu = \mu$

$E(\overline{X}_b) = E\left(\frac{X_1 + 4X_2 + 4X_3 + X_4}{10}\right) = \frac{1}{10}\mu + \frac{4}{10}\mu + \frac{4}{10}\mu + \frac{1}{10}\mu = \mu$

b. $V(\overline{X}_a) = V\left(\frac{X_1 + 2X_2 + 3X_3 + 4X_4}{10}\right) = \frac{1}{100}\sigma^2 + \frac{4}{100}\sigma^2 + \frac{9}{100}\sigma^2 + \frac{16}{100}\sigma^2 = 0{,}30\sigma^2$

$$V(\overline{X}_b)=V\left(\frac{X_1+4X_2+4X_3+X_4}{10}\right)=\frac{1}{100}\sigma^2+\frac{16}{100}\sigma^2+\frac{16}{100}\sigma^2+\frac{1}{100}\sigma^2=0{,}34\sigma^2$$

c. $GE=\dfrac{Var(\hat{\theta}_2)}{Var(\hat{\theta}_1)}=\dfrac{0{,}30\sigma^2}{0{,}34\sigma^2}=0{,}88$

5. $P(2{,}02<\mu<3{,}98)=0{,}95$

7. **a.** $P(5{,}714<\mu<6{,}886)=0{,}80$

$P(5{,}544<\mu<7{,}056)=0{,}90$

$P(5{,}402<\mu<7{,}198)=0{,}95$

b. 5.3 ile 7.3 arasında bulunma olasılığı 0.2434'dür.

9. $P(284{,}93<\mu<285{,}6)=0{,}95$

11. $P\left(40{,}23-2{,}0930\times\dfrac{1{,}2}{\sqrt{20}}<\mu<40{,}23+2{,}0930\times\dfrac{1{,}2}{\sqrt{20}}\right)=0{,}95$

$P(39{,}66<\mu<40{,}79)=0{,}95$

13. $P(20{,}29<\mu<27{,}70)=0{,}95$

15. $P(0{,}3<p<0{,}37)=0{,}90$

17. $P(345{,}6<\sigma^2<1195{,}8)=0{,}80$

19. $n\geq Z_{\alpha/2}^2\dfrac{p(1-p)}{e^2}\Rightarrow n\geq 1{,}96^2\dfrac{0{,}5\times 0{,}5}{0{,}03^2}\Rightarrow n\geq 1067$

SEKİZİNCİ BÖLÜM

1. **a.** $\overline{X}=\dfrac{\sum X}{n}=3{,}6$ $\quad S^2=\dfrac{\sum(X-\overline{X})^2}{n-1}=2{,}4631$ **b.** $P(2{,}597<\mu<4{,}603)=0{,}99$ **c.** H_0:μ=7 hipotezi reddedilir.

3. **a.** $P(0{,}598<p<0{,}661)=0{,}90$ **b.** H_0:p=0.50 hipotezi reddedilir.

5. **a.** $H_0:\mu\leq 1600$ hipotezi reddedilemez. **b.** $H_0:\mu\geq 1600$ hipotezi reddedilir. **c.** H_0: μ=1600 hipotezi reddedilemez.

7. $H_0:\mu\leq 600$ hipotezi reddedilemez.

9. $H_0:\sigma^2\geq 0{,}81$ hipotezi reddedilemez.

11. $H_0:p\leq 0{,}20$ hipotezi reddedilemez.

13. $H_0:\sigma^2\leq 1{,}69$ hipotezi reddedilemez.

15. $H_0:\mu=1800$ hipotezi reddedilir.

17. $H_0:\sigma^2\leq 9$ hipotezi reddedilemez.

19. $H_0:\mu\geq 30$ hipotezi reddedilemez.

DOKUZUNCU BÖLÜM

2. H_0: $\sigma_1^2=\sigma_2^2$ hipotezi reddedilemez.

4. H_0:$\mu_1=\mu_2$ hipotezi reddedilir.

6. H_0: $p_1-p_2=0$ hipotezi reddedilir.

8. **a.** $P[-0,0092 < p_1 - p_2 < 0,1492] = 0,99$ **b.** H_0: p_1-p_2=0 hipotezi reddedilemez.

10. **a.** H_0: p_1-p_2=0 hipotezi reddedilir. **b.** $P[-0,26 < p_1 - p_2 < -0,1876] = 0,95$

12. H_0:μ_d=0 hipotezi reddedilir.

14. H_0:μ_d=0 hipotezi reddedilir.

16. H_0:μ_1=μ_2 hipotezi reddedilir.

18. H_0: μ_1=μ_2 hipotezi reddedilemez.

20. **a.** $P[-17,77 < \mu_1 - \mu_2 < -14,23] = 0,90$ **b.** H_0: μ_1=μ_2 hipotezi reddedilir.

ONUNCU BÖLÜM

2. A, B, C, D gibi belli markalar karşılaştırılıyor ise, bu grupların çekildikleri kitlelerin ortalamaları sırasıyla $\mu_A, \mu_B,...$ gibi kesin değerlerdir. Bu ortalamaların eşitliği;

H_o: $\mu_A = \mu_B = ... = \mu_D$ (Ortalamalar arasında fark yoktur.)

$H_{A:}$ En az bir ortalama diğerinden farklıdır

hipotezleri ile test edilir. Sonuç genellenemez yalnız incelenen ortalamalar için geçerlidir. H_o hipotezi reddedilmiş ise, hangi ortalamanın farklı olduğunu bulmak için ortalamalar arası farkı belirleyen yöntemlerden biri kullanılır.

4. H_o: $\mu_1 = \mu_2 = \mu_3$ (Ortalamalar arasında fark yoktur.)

$H_{A:}$ En az bir ortalama diğerinden farklıdır

H_0 reddedilemez. 0.05 anlamlılık düzeyinde ortalamalar arasında fark yoktur.

6. H_o: $\mu_1 = \mu_2 = \mu_3$ (Ortalamalar arasında fark yoktur.)

$H_{A:}$ En az bir ortalama diğerinden farklıdır

H_0 reddedilir. 0.05 anlamlılık düzeyinde en az bir ortalama diğerlerinden farklıdır.

8. H_o: $\mu_1 = \mu_2 = \mu_3$ (Ortalamalar arasında fark yoktur.)

$H_{A:}$ En az bir ortalama diğerinden farklıdır

H_0 reddedilir. 0.05 anlamlılık düzeyinde en az bir ortalama diğerinden farklıdır.

10. H_o: $\mu_1 = \mu_2 = \mu_3$ (Ortalamalar arasında fark yoktur.)

$H_{A:}$ En az bir ortalama diğerinden farklıdır

H_0 reddedilemez. 0.05 anlamlılık düzeyinde ortalamalar arasında fark yoktur.

12. H_o: $\mu_1 = \mu_2 = \mu_3$ (Ortalamalar arasında fark yoktur.)

$H_{A:}$ En az bir ortalama diğerinden farklıdır

H_0 reddedilir. 0.05 anlamlılık düzeyinde en az bir ortalama diğerlerinden farklıdır.

14. H_o: $\mu_1 = \mu_2 = \mu_3$ (Ortalamalar arasında fark yoktur.)

H_A: En az bir ortalama diğerinden farklıdır

H_0 reddedilir. 0.05 anlamlılık düzeyinde en az bir ortalama diğerlerinden farklıdır.

16. **a.** Birinci varsayım; X rastlantı değişkenleri yani bitkilerdeki hastalık yüzdeleri normal dağılıma sahiptir. İkinci varsayım; Gruplara ait varyanslar eşittir.

b. H_o: $\mu_0 = \mu_1 = \mu_2 = \mu_3 = \mu$ (Ortalamalar arasında fark yoktur.)

H_A: En az bir ortalama diğerinden farklıdır

H_0 reddedilir. 0.05 anlamlılık düzeyinde ortalamalar farklıdır.

c. Tüm gruplar birbirinden farklıdır.

18. H_o: $\mu_1 = \mu_2 = \mu_3$ (Ortalamalar arasında fark yoktur.)

H_A: En az bir ortalama diğerinden farklıdır

H_0 reddedilemez. 0.05 anlamlılık düzeyinde ortalamalar farklı değildir.

20. H_o: $\mu_1 = \mu_2 = \mu_3$ (Ortalamalar arasında fark yoktur.)

H_A: En az bir ortalama diğerinden farklıdır

H_0 reddedilir. 0.05 anlamlılık düzeyinde en az bir ortalama diğerlerinden farklıdır.

ONBİRİNCİ BÖLÜM

1. H_o: Değişkenler arasında ilişki yoktur.

H_A: Değişkenler arasında ilişki vardır.

H_o hipotezi reddedilir. 0,05 anlamlılık düzeyinde, değişkenler arasında ilişki vardır.

3. H_o: Değişkenler arasında ilişki yoktur.

H_A: Değişkenler arasında ilişki vardır.

H_o hipotezi reddedilemez. 0,01 anlamlılık düzeyinde, değişkenler arasında ilişki yoktur.

5. H_o: Gruplar arasında fark yoktur (Gruplar homojendir).

H_A: Gruplar arasında fark vardır (Gruplar homojen değildir).

H_o hipotezi reddedilemez. 0,01 anlamlılık düzeyinde, gruplar arasında fark yoktur.

7. H_o: Değişkenler arasında ilişki yoktur.

H_A: Değişkenler arasında ilişki vardır.

H_o hipotezi reddedilir. 0,05 anlamlılık düzeyinde, değişkenler arasında ilişki vardır.

9. H_o: Değişkenler arasında ilişki yoktur.

H_A: Değişkenler arasında ilişki vardır.

H_o hipotezi reddedilir. 0,05 anlamlılık düzeyinde, değişkenler arasında ilişki vardır.

Pearson C katsayısına göre beslenme ile başarılı olma arasında %50'lik bir ilişki vardır.

11. Pearson C katsayısına göre %25'lik bir ilişki vardır.

13. H_o: Gruplar arasında fark yoktur (Gruplar homojendir).

H_A: Gruplar arasında fark vardır (Gruplar homojen değildir).

H_o hipotezi reddedilir. 0,01 anlamlılık düzeyinde, gruplar arasında fark vardır.

15. H_o: Gruplar arasında fark yoktur (Gruplar homojendir).

H_A: Gruplar arasında fark vardır (Gruplar homojen değildir).

H_o hipotezi reddedilir. 0,05 anlamlılık düzeyinde, gruplar arasında fark vardır.

17. H_o: Değişkenler arasında ilişki yoktur.

H_A: Değişkenler arasında ilişki vardır.

H_o hipotezi reddedilir. 0,05 anlamlılık düzeyinde, gruplar arasında fark vardır.

19. H_o: Değişkenler arasında ilişki yoktur.

H_A: Değişkenler arasında ilişki vardır.

H_o hipotezi reddedilir. 0,01 anlamlılık düzeyinde, değişkenler arasında ilişki vardır.

ONİKİNCİ BÖLÜM

1. **a.** $$r = \frac{\sum x_i.y_i - \left(\sum x_i \sum y_i\right)/n}{\sqrt{\sum x_i^2 - \left(\sum x_i\right)^2/n}\ \sqrt{\sum y_i^2 - \left(\sum y_i\right)^2/n}} = 0{,}973$$ **b.** $\hat{Y} = 78{,}801 + 2{,}123X$

3. **a.** $$r = \frac{\sum x_i.y_i - \left(\sum x_i \sum y_i\right)/n}{\sqrt{\sum x_i^2 - \left(\sum x_i\right)^2/n}\ \sqrt{\sum y_i^2 - \left(\sum y_i\right)^2/n}} = 0{,}539$$

b. 0.05 anlamlılık düzeyinde korelasyon katsayısı önemsizdir.

c. $\hat{Y} = 2{,}802 + 0{,}495X$ **d.** $\hat{Y} = 2{,}802 + 0{,}495 \times 7 = 6{,}267$

5. **a.** $$r = \frac{\sum x_i.y_i - \left(\sum x_i \sum y_i\right)/n}{\sqrt{\sum x_i^2 - \left(\sum x_i\right)^2/n}\ \sqrt{\sum y_i^2 - \left(\sum y_i\right)^2/n}} = 0{,}786$$

b. H_o: $\rho = 0$

H_A: $\rho \neq 0$

H_o hipotezi reddedilemez. 0.05 anlamlılık düzeyinde korelasyon katsayısı önemsizdir.

c. $\hat{Y} = 2{,}679 + 0{,}393X$

H_0 : Model geçersizdir.

H_1 : Model geçerlidir. 0.05 anlamlılık düzeyinde model geçersizdir.

7. $$r=\frac{\sum x_i.y_i-\left(\sum x_i\sum y_i\right)/n}{\sqrt{\sum x_i^2-\left(\sum x_i\right)^2/n}\sqrt{\sum y_i^2-\left(\sum y_i\right)^2/n}}=-0{,}95$$ $\hat{Y}=9{,}399-0{,}428X$

9. $$r=\frac{\sum x_i.y_i-\left(\sum x_i\sum y_i\right)/n}{\sqrt{\sum x_i^2-\left(\sum x_i\right)^2/n}\sqrt{\sum y_i^2-\left(\sum y_i\right)^2/n}}=0{,}559$$

H_o: $\rho=0$

H_A: $\rho\neq 0$

H_o hipotezi reddedilir. 0.10 anlamlılık düzeyinde korelasyon katsayısı önemlidir.

11. **a.** $\hat{Y}=20{,}895+0{,}038X$

H_0 : Model geçersizdir.

H_1 : Model geçerlidir.

0.01 anlamlılık düzeyinde model geçersizdir.

b. Yaklaşık olarak 46 YTL

13. **a.** $\hat{Y}=26{,}073+0{,}239X$

H_0 : Model geçersizdir.

H_1 : Model geçerlidir.

0.01 anlamlılık düzeyinde model geçersizdir.

b. Yaklaşık olarak 49

15. **a.** $$r=\frac{\sum x_i.y_i-\left(\sum x_i\sum y_i\right)/n}{\sqrt{\sum x_i^2-\left(\sum x_i\right)^2/n}\sqrt{\sum y_i^2-\left(\sum y_i\right)^2/n}}=0{,}725$$

H_o: $\rho=0$

H_A: $\rho\neq 0$

H_o hipotezi reddedilir. 0.05 anlamlılık düzeyinde korelasyon katsayısı önemlidir.

b. $\hat{Y}=308{,}944+328{,}199X$

H_0 : Model geçersizdir.

H_1 : Model geçerlidir.

0.05 anlamlılık düzeyinde model geçerlidir.

17. $\hat{Y}=-13{,}746+1{,}157X$

H_0 : Model geçersizdir.

H_1 : Model geçerlidir.

0.05 anlamlılık düzeyinde model geçerlidir.

Yaklaşık olarak 14

19. **a.** $P\left(b_1 - t_{\alpha/2;n-2}S_{b_1} < \beta_1 < b_1 + t_{\alpha/2;n-2}S_{b_1}\right) = 1-\alpha$

$P(328{,}199 - 2{,}3060 \times 110{,}43 < \beta_1 < 328{,}199 + 2{,}3060 \times 110{,}43) = 0{,}95$

$P(73{,}55 < \beta_1 < 582{,}85) = 0{,}95$

b. x_p=6 ise y tahmin değeri 2278.13'dür.

$$P\left(\hat{y} - t_{\alpha/2;n-2}\left[RAKO\left(\frac{1}{n} + \frac{(x_p - \overline{X})^2}{XOAKT}\right)\right]^{1/2} < E(Y/x_p) < \hat{y} + t_{\alpha/2;n-2}\left[RAKO\left(\frac{1}{n} + \frac{(x_p - \overline{X})^2}{XOAKT}\right)\right]^{1/2}\right) = 1-\alpha$$

$$P\left(2278{,}13 - 2{,}306\left[195510{,}78\left(\frac{1}{10} + \frac{(6-3{,}3)^2}{16{,}1}\right)\right]^{1/2} < E(Y/x_p) < 2278{,}13 + 2{,}306\left[195510{,}78\left(\frac{1}{10} + \frac{(6-3{,}3)^2}{16{,}1}\right)\right]^{1/2}\right) = 0{,}95$$

$P(1520 < E(Y/x_p) < 3036{,}23) = 0{,}95$

KAYNAKLAR

Akdeniz, F.: Olasılık ve İstatistik, Baki Kitapevi, Adana, 2000.

Apaydın, A., Kutsal, A., Atakan, C.: Uygulamalı İstatistik, Kültür Kitap ve Yayınevi, Ankara, 1997.

Aytaç, M.: Matematiksel İstatistik, Ezgi Kitapevi Yayınları, 1991.

Başkan, Ş.: Uygulamalı İstatistik, Ege Üniversitesi Basımevi, İzmir, 1993.

Çıngı, H.: Örnekleme Kuramı, H.Ü. Fen Fakültesi Yayınları Ders Kitapları Dizisi 20, H.Ü. Fen Fakültesi Basımevi, Ankara, 1990.

Draper, N.R., Smith, H.: Applied Regression Analysis, John wileyand Sons, New York, 1966.

Edwards, A.L.: An Introduction to Linear Regression and Correlation, W.H. Freeman and Company, New York, 1984.

Erar, A.: Bağlanım (Regresyon) Çözümlemesi, H.Ü. Ders Notları, Ankara, 1985.

Esensoy, Ö.: Olasılık, Cilt I Temel Tanımlar ve Kavramlar, DIE Matbaası, Ankara, 1994.

Esin, A., Ekni, M., Gamgam, H.: Sağlık Bilimlerinde İstatistik, Gazi Üniversitesi Yayını, 171, Gazi Üniversitesi Basım-Yayın Yüksekokulu Matbaası, Ankara, 1991.

Ersoy, N., Erbas (Oral),S.: Olasılık ve İstatistiğe Giriş, Gazi Üniversitesi Yayını, 175, Ankara, 1992.

Groebner, D. F., Shannon, P.W., Fry, P.C., Smith,K.D.: Business Statistics, Sixty Edition, Pearson Education, New Jersey, 2001.

Hossack, I.B., Pollard, J.H., Zehnwirth, B.: Introductory Statistics with Applications in General Insurance, Cambridge Universty ,Press, New York, 1983.

İnal, H.C., Günay, S.: Olasılık ve Matematiksel İstatistik, H.Ü. Fen Fakültesi Basımevi, Ankara, 1982.

Kohler, H.: Statistics for Business and Economics, Scott, Foresman and Company, England, 1985.

Kutsal, A., Muluk, F.Z.: Uygulamalı Temel İstatistik, H.Ü. Fen Fakültesi Yayınları, Ders Kitapları Dizisi 8, H.Ü. Fen Fakültesi Basımevi, Ankara, 1990.

Özdamar, K.: Paket Programlar İle İstatistiksel Veri Analizi, Kaan Kitapevi, Eskişehir, 2002.

Özden, H.: İstatistik Kuramı ve Uygulamaları, H.Ü. Fen Fakültesi Basımevi, Ankara, 1993.

Öztürk, A.: Tarım Biyoloji ve Sağlık Bilimlerinde Uygulamalı İstatistik, Ege Üniversitesi Matbaası, İzmir, 1978.

Saraçbaşı, T., Kutsal, A.: Betimsel İstatistik, H.Ü. Fen Fakültesi Yayınları Ders Kitapları Dizisi 17, H.Ü. Fen Fakültesi Basımevi, Ankara, 1987.

Sıegel, A. F. and Morgan, C. J.: Statistics And Data Analysıs, John Wıley&Sons Inc, New York, 1996.

Şenesen, Ü.: İşletme ve İktisat için İstatistik, 4. Basım Çevirisi, Paul Newbold, 2000.

Türkiye İstatistik Yıllığı : DIE, 2002.

Ünver, Ö.: Uygulamalı İstatistik Yöntemler Giriş, Genişletilmiş Baskı, Siyasal Kitapevi, Ankara, 1995.

Ünver, Ö.,Gamgam, H.: Uygulamalı İstatistik Yöntemler, Siyasal Kitapevi, Ankara, 1996.

TABLOLAR

EK 1: Binom Dağılım Tablosu

$n = 1$

x	$p = 0.01$	$p = 0.02$	$p = 0.03$	$p = 0.04$	$p = 0.05$	$p = 0.06$	$p = 0.07$	$p = 0.08$	$p = 0.09$	$p = 0.10$	$n - x$
0	0.9900	0.9800	0.9700	0.9600	0.9500	0.9400	0.9300	0.9200	0.9100	0.9000	1
1	0.0100	0.0200	0.0300	0.0400	0.0500	0.0600	0.0700	0.0800	0.0900	0.1000	0
	$q = 0.99$	$q = 0.98$	$q = 0.97$	$q = 0.96$	$q = 0.95$	$q = 0.94$	$q = 0.93$	$q = 0.92$	$q = 0.91$	$q = 0.90$	
x	$p = 0.11$	$p = 0.12$	$p = 0.13$	$p = 0.14$	$p = 0.15$	$p = 0.16$	$p = 0.17$	$p = 0.18$	$p = 0.19$	$p = 0.20$	$n - x$
0	0.8900	0.8800	0.8700	0.8600	0.8500	0.8400	0.8300	0.8200	0.8100	0.8000	1
1	0.1100	0.1200	0.1300	0.1400	0.1500	0.1600	0.1700	0.1800	0.1900	0.2000	0
	$q = 0.89$	$q = 0.88$	$q = 0.87$	$q = 0.86$	$q = 0.85$	$q = 0.84$	$q = 0.83$	$q = 0.82$	$q = 0.81$	$q = 0.80$	
x	$p = 0.21$	$p = 0.22$	$p = 0.23$	$p = 0.24$	$p = 0.25$	$p = 0.26$	$p = 0.27$	$p = 0.28$	$p = 0.29$	$p = 0.30$	$n - x$
0	0.7900	0.7800	0.7700	0.7600	0.7500	0.7400	0.7300	0.7200	0.7100	0.7000	1
1	0.2100	0.2200	0.2300	0.2400	0.2500	0.2600	0.2700	0.2800	0.2900	0.3000	0
	$q = 0.79$	$q = 0.78$	$q = 0.77$	$q = 0.76$	$q = 0.75$	$q = 0.74$	$q = 0.73$	$q = 0.72$	$q = 0.71$	$q = 0.70$	
x	$p = 0.31$	$p = 0.32$	$p = 0.33$	$p = 0.34$	$p = 0.35$	$p = 0.36$	$p = 0.37$	$p = 0.38$	$p = 0.39$	$p = 0.40$	$n - x$
0	0.6900	0.6800	0.6700	0.6600	0.6500	0.6400	0.6300	0.6200	0.6100	0.6000	1
1	0.3100	0.3200	0.3300	0.3400	0.3500	0.3600	0.3700	0.3800	0.3900	0.4000	0
	$q = 0.69$	$q = 0.68$	$q = 0.67$	$q = 0.66$	$q = 0.65$	$q = 0.64$	$q = 0.63$	$q = 0.62$	$q = 0.61$	$q = 0.60$	
x	$p = 0.41$	$p = 0.42$	$p = 0.43$	$p = 0.44$	$p = 0.45$	$p = 0.46$	$p = 0.47$	$p = 0.48$	$p = 0.49$	$p = 0.50$	$n - x$
0	0.5900	0.5800	0.5700	0.5600	0.5500	0.5400	0.5300	0.5200	0.5100	0.5000	1
1	0.4100	0.4200	0.4300	0.4400	0.4500	0.4600	0.4700	0.4800	0.4900	0.5000	0
	$q = 0.59$	$q = 0.58$	$q = 0.57$	$q = 0.56$	$q = 0.55$	$q = 0.54$	$q = 0.53$	$q = 0.52$	$q = 0.51$	$q = 0.50$	

$n = 2$

x	$p = 0.01$	$p = 0.02$	$p = 0.03$	$p = 0.04$	$p = 0.05$	$p = 0.06$	$p = 0.07$	$p = 0.08$	$p = 0.09$	$p = 0.10$	$n - x$
0	0.9801	0.9604	0.9409	0.9216	0.9025	0.8836	0.8649	0.8464	0.8281	0.8100	2
1	0.0198	0.0392	0.0582	0.0768	0.0950	0.1128	0.1302	0.1472	0.1638	0.1800	1
2	0.0001	0.0004	0.0009	0.0016	0.0025	0.0036	0.0049	0.0064	0.0081	0.0100	0
	$q = 0.99$	$q = 0.98$	$q = 0.97$	$q = 0.96$	$q = 0.95$	$q = 0.94$	$q = 0.93$	$q = 0.92$	$q = 0.91$	$q = 0.90$	
x	$p = 0.11$	$p = 0.12$	$p = 0.13$	$p = 0.14$	$p = 0.15$	$p = 0.16$	$p = 0.17$	$p = 0.18$	$p = 0.19$	$p = 0.20$	$n - x$
0	0.7921	0.7744	0.7569	0.7396	0.7225	0.7056	0.6889	0.6724	0.6561	0.6400	2
1	0.1958	0.2112	0.2262	0.2408	0.2550	0.2688	0.2822	0.2952	0.3078	0.3200	1
2	0.0121	0.0144	0.0169	0.0196	0.0225	0.0256	0.0289	0.0324	0.0361	0.0400	0
	$q = 0.89$	$q = 0.88$	$q = 0.87$	$q = 0.86$	$q = 0.85$	$q = 0.84$	$q = 0.83$	$q = 0.82$	$q = 0.81$	$q = 0.80$	
x	$p = 0.21$	$p = 0.22$	$p = 0.23$	$p = 0.24$	$p = 0.25$	$p = 0.26$	$p = 0.27$	$p = 0.28$	$p = 0.29$	$p = 0.30$	$n - x$
0	0.6241	0.6084	0.5929	0.5776	0.5625	0.5476	0.5329	0.5184	0.5041	0.4900	2
1	0.3318	0.3432	0.3542	0.3648	0.3750	0.3848	0.3942	0.4032	0.4118	0.4200	1
2	0.0441	0.0484	0.0529	0.0576	0.0625	0.0676	0.0729	0.0784	0.0841	0.0900	0
	$q = 0.79$	$q = 0.78$	$q = 0.77$	$q = 0.76$	$q = 0.75$	$q = 0.74$	$q = 0.73$	$q = 0.72$	$q = 0.71$	$q = 0.70$	
x	$p = 0.31$	$p = 0.32$	$p = 0.33$	$p = 0.34$	$p = 0.35$	$p = 0.36$	$p = 0.37$	$p = 0.38$	$p = 0.39$	$p = 0.40$	$n - x$
0	0.4761	0.4624	0.4489	0.4356	0.4225	0.4096	0.3969	0.3844	0.3721	0.3600	2
1	0.4278	0.4352	0.4422	0.4488	0.4550	0.4608	0.4662	0.4712	0.4758	0.4800	1
2	0.0961	0.1024	0.1089	0.1156	0.1225	0.1296	0.1369	0.1444	0.1521	0.1600	0
	$q = 0.69$	$q = 0.68$	$q = 0.67$	$q = 0.66$	$q = 0.65$	$q = 0.64$	$q = 0.63$	$q = 0.62$	$q = 0.61$	$q = 0.60$	
x	$p = 0.41$	$p = 0.42$	$p = 0.43$	$p = 0.44$	$p = 0.45$	$p = 0.46$	$p = 0.47$	$p = 0.48$	$p = 0.49$	$p = 0.50$	$n - x$
0	0.3481	0.3364	0.3249	0.3136	0.3025	0.2916	0.2809	0.2704	0.2601	0.2500	2
1	0.4838	0.4872	0.4902	0.4928	0.4950	0.4968	0.4982	0.4992	0.4998	0.5000	1
2	0.1681	0.1764	0.1849	0.1936	0.2025	0.2116	0.2209	0.2304	0.2401	0.2500	0
	$q = 0.59$	$q = 0.58$	$q = 0.57$	$q = 0.56$	$q = 0.55$	$q = 0.54$	$q = 0.53$	$q = 0.52$	$q = 0.51$	$q = 0.50$	

Kaynak: Groebner vd., Business Statistics, 2005.

$n = 3$

x	$p = 0.01$	$p = 0.02$	$p = 0.03$	$p = 0.04$	$p = 0.05$	$p = 0.06$	$p = 0.07$	$p = 0.08$	$p = 0.09$	$p = 0.10$	$n - x$
0	0.9703	0.9412	0.9127	0.8847	0.8574	0.8306	0.8044	0.7787	0.7536	0.7290	3
1	0.0294	0.0576	0.0847	0.1106	0.1354	0.1590	0.1816	0.2031	0.2236	0.2430	2
2	0.0003	0.0012	0.0026	0.0046	0.0071	0.0102	0.0137	0.0177	0.0221	0.0270	1
3	0.0000	0.0000	0.0000	0.0001	0.0001	0.0002	0.0003	0.0005	0.0007	0.0010	0
	$q = 0.99$	$q = 0.98$	$q = 0.97$	$q = 0.96$	$q = 0.95$	$q = 0.94$	$q = 0.93$	$q = 0.92$	$q = 0.91$	$q = 0.90$	
x	$p = 0.11$	$p = 0.12$	$p = 0.13$	$p = 0.14$	$p = 0.15$	$p = 0.16$	$p = 0.17$	$p = 0.18$	$p = 0.19$	$p = 0.20$	$n - x$
0	0.7050	0.6815	0.6585	0.6361	0.6141	0.5927	0.5718	0.5514	0.5314	0.5120	3
1	0.2614	0.2788	0.2952	0.3106	0.3251	0.3387	0.3513	0.3631	0.3740	0.3840	2
2	0.0323	0.0380	0.0441	0.0506	0.0574	0.0645	0.0720	0.0797	0.0877	0.0960	1
3	0.0013	0.0017	0.0022	0.0027	0.0034	0.0041	0.0049	0.0058	0.0069	0.0080	0
	$q = 0.89$	$q = 0.88$	$q = 0.87$	$q = 0.86$	$q = 0.85$	$q = 0.84$	$q = 0.83$	$q = 0.82$	$q = 0.81$	$q = 0.80$	
x	$p = 0.21$	$p = 0.22$	$p = 0.23$	$p = 0.24$	$p = 0.25$	$p = 0.26$	$p = 0.27$	$p = 0.28$	$p = 0.29$	$p = 0.30$	$n - x$
0	0.4930	0.4746	0.4565	0.4390	0.4219	0.4052	0.3890	0.3732	0.3579	0.3430	3
1	0.3932	0.4015	0.4091	0.4159	0.4219	0.4271	0.4316	0.4355	0.4386	0.4410	2
2	0.1045	0.1133	0.1222	0.1313	0.1406	0.1501	0.1597	0.1693	0.1791	0.1890	1
3	0.0093	0.0106	0.0122	0.0138	0.0156	0.0176	0.0197	0.0220	0.0244	0.0270	0
	$q = 0.79$	$q = 0.78$	$q = 0.77$	$q = 0.76$	$q = 0.75$	$q = 0.74$	$q = 0.73$	$q = 0.72$	$q = 0.71$	$q = 0.70$	
x	$p = 0.31$	$p = 0.32$	$p = 0.33$	$p = 0.34$	$p = 0.35$	$p = 0.36$	$p = 0.37$	$p = 0.38$	$p = 0.39$	$p = 0.40$	$n - x$
0	0.3285	0.3144	0.3008	0.2875	0.2746	0.2621	0.2500	0.2383	0.2270	0.2160	3
1	0.4428	0.4439	0.4444	0.4443	0.4436	0.4424	0.4406	0.4382	0.4354	0.4320	2
2	0.1989	0.2089	0.2189	0.2289	0.2389	0.2488	0.2587	0.2686	0.2783	0.2880	1
3	0.0298	0.0328	0.0359	0.0393	0.0429	0.0467	0.0507	0.0549	0.0593	0.0640	0
	$q = 0.69$	$q = 0.68$	$q = 0.67$	$q = 0.66$	$q = 0.65$	$q = 0.64$	$q = 0.63$	$q = 0.62$	$q = 0.61$	$q = 0.60$	
x	$p = 0.41$	$p = 0.42$	$p = 0.43$	$p = 0.44$	$p = 0.45$	$p = 0.46$	$p = 0.47$	$p = 0.48$	$p = 0.49$	$p = 0.50$	$n - x$
0	0.2054	0.1951	0.1852	0.1756	0.1664	0.1575	0.1489	0.1406	0.1327	0.1250	3
1	0.4282	0.4239	0.4191	0.4140	0.4084	0.4024	0.3961	0.3894	0.3823	0.3750	2
2	0.2975	0.3069	0.3162	0.3252	0.3341	0.3428	0.3512	0.3594	0.3674	0.3750	1
3	0.0689	0.0741	0.0795	0.0852	0.0911	0.0973	0.1038	0.1106	0.1176	0.1250	0
	$q = 0.59$	$q = 0.58$	$q = 0.57$	$q = 0.56$	$q = 0.55$	$q = 0.54$	$q = 0.53$	$q = 0.52$	$q = 0.51$	$q = 0.50$	

$n = 4$

x	$p = 0.01$	$p = 0.02$	$p = 0.03$	$p = 0.04$	$p = 0.05$	$p = 0.06$	$p = 0.07$	$p = 0.08$	$p = 0.09$	$p = 0.10$	$n - x$
0	0.9606	0.9224	0.8853	0.8493	0.8145	0.7807	0.7481	0.7164	0.6857	0.6561	4
1	0.0388	0.0753	0.1095	0.1416	0.1715	0.1993	0.2252	0.2492	0.2713	0.2916	3
2	0.0006	0.0023	0.0051	0.0088	0.0135	0.0191	0.0254	0.0325	0.0402	0.0486	2
3	0.0000	0.0000	0.0001	0.0002	0.0005	0.0008	0.0013	0.0019	0.0027	0.0036	1
4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0
	$q = 0.99$	$q = 0.98$	$q = 0.97$	$q = 0.96$	$q = 0.95$	$q = 0.94$	$q = 0.93$	$q = 0.92$	$q = 0.91$	$q = 0.90$	
x	$p = 0.11$	$p = 0.12$	$p = 0.13$	$p = 0.14$	$p = 0.15$	$p = 0.16$	$p = 0.17$	$p = 0.18$	$p = 0.19$	$p = 0.20$	$n - x$
0	0.6274	0.5997	0.5729	0.5470	0.5220	0.4979	0.4746	0.4521	0.4305	0.4096	4
1	0.3102	0.3271	0.3424	0.3562	0.3685	0.3793	0.3888	0.3970	0.4039	0.4096	3
2	0.0575	0.0669	0.0767	0.0870	0.0975	0.1084	0.1195	0.1307	0.1421	0.1536	2
3	0.0047	0.0061	0.0076	0.0094	0.0115	0.0138	0.0163	0.0191	0.0222	0.0256	1
4	0.0001	0.0002	0.0003	0.0004	0.0005	0.0007	0.0008	0.0010	0.0013	0.0016	0
	$q = 0.89$	$q = 0.88$	$q = 0.87$	$q = 0.86$	$q = 0.85$	$q = 0.84$	$q = 0.83$	$q = 0.82$	$q = 0.81$	$q = 0.80$	

x	$p = 0.21$	$p = 0.22$	$p = 0.23$	$p = 0.24$	$p = 0.25$	$p = 0.26$	$p = 0.27$	$p = 0.28$	$p = 0.29$	$p = 0.30$	$n - x$
0	0.3895	0.3702	0.3515	0.3336	0.3164	0.2999	0.2840	0.2687	0.2541	0.2401	4
1	0.4142	0.4176	0.4200	0.4214	0.4219	0.4214	0.4201	0.4180	0.4152	0.4116	3
2	0.1651	0.1767	0.1882	0.1996	0.2109	0.2221	0.2331	0.2439	0.2544	0.2646	2
3	0.0293	0.0332	0.0375	0.0420	0.0469	0.0520	0.0575	0.0632	0.0693	0.0756	1
4	0.0019	0.0023	0.0028	0.0033	0.0039	0.0046	0.0053	0.0061	0.0071	0.0081	0
	$q = 0.79$	$q = 0.78$	$q = 0.77$	$q = 0.76$	$q = 0.75$	$q = 0.74$	$q = 0.73$	$q = 0.72$	$q = 0.71$	$q = 0.70$	

x	$p = 0.31$	$p = 0.32$	$p = 0.33$	$p = 0.34$	$p = 0.35$	$p = 0.36$	$p = 0.37$	$p = 0.38$	$p = 0.39$	$p = 0.40$	$n - x$
0	0.2267	0.2138	0.2015	0.1897	0.1785	0.1678	0.1575	0.1478	0.1385	0.1296	4
1	0.4074	0.4025	0.3970	0.3910	0.3845	0.3775	0.3701	0.3623	0.3541	0.3456	3
2	0.2745	0.2841	0.2933	0.3021	0.3105	0.3185	0.3260	0.3330	0.3396	0.3456	2
3	0.0822	0.0891	0.0963	0.1038	0.1115	0.1194	0.1276	0.1361	0.1447	0.1536	1
4	0.0092	0.0105	0.0119	0.0134	0.0150	0.0168	0.0187	0.0209	0.0231	0.0256	0
	$q = 0.69$	$q = 0.68$	$q = 0.67$	$q = 0.66$	$q = 0.65$	$q = 0.64$	$q = 0.63$	$q = 0.62$	$q = 0.61$	$q = 0.60$	

x	$p = 0.41$	$p = 0.42$	$p = 0.43$	$p = 0.44$	$p = 0.45$	$p = 0.46$	$p = 0.47$	$p = 0.48$	$p = 0.49$	$p = 0.50$	$n - x$
0	0.1212	0.1132	0.1056	0.0983	0.0915	0.0850	0.0789	0.0731	0.0677	0.0625	4
1	0.3368	0.3278	0.3185	0.3091	0.2995	0.2897	0.2799	0.2700	0.2600	0.2500	3
2	0.3511	0.3560	0.3604	0.3643	0.3675	0.3702	0.3723	0.3738	0.3747	0.3750	2
3	0.1627	0.1719	0.1813	0.1908	0.2005	0.2102	0.2201	0.2300	0.2400	0.2500	1
4	0.0283	0.0311	0.0342	0.0375	0.0410	0.0448	0.0488	0.0531	0.0576	0.0625	0
	$q = 0.59$	$q = 0.58$	$q = 0.57$	$q = 0.56$	$q = 0.55$	$q = 0.54$	$q = 0.53$	$q = 0.52$	$q = 0.51$	$q = 0.50$	

$n = 5$

x	$p = 0.01$	$p = 0.02$	$p = 0.03$	$p = 0.04$	$p = 0.05$	$p = 0.06$	$p = 0.07$	$p = 0.08$	$p = 0.09$	$p = 0.10$	$n - x$
0	0.9510	0.9039	0.8587	0.8154	0.7738	0.7339	0.6957	0.6591	0.6240	0.5905	5
1	0.0480	0.0922	0.1328	0.1699	0.2036	0.2342	0.2618	0.2866	0.3086	0.3281	4
2	0.0010	0.0038	0.0082	0.0142	0.0214	0.0299	0.0394	0.0498	0.0610	0.0729	3
3	0.0000	0.0001	0.0003	0.0006	0.0011	0.0019	0.0030	0.0043	0.0060	0.0081	2
4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0002	0.0003	0.0005	1
5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0
	$q = 0.99$	$q = 0.98$	$q = 0.97$	$q = 0.96$	$q = 0.95$	$q = 0.94$	$q = 0.93$	$q = 0.92$	$q = 0.91$	$q = 0.90$	

x	$p = 0.11$	$p = 0.12$	$p = 0.13$	$p = 0.14$	$p = 0.15$	$p = 0.16$	$p = 0.17$	$p = 0.18$	$p = 0.19$	$p = 0.20$	$n - x$
0	0.5584	0.5277	0.4984	0.4704	0.4437	0.4182	0.3939	0.3707	0.3487	0.3277	5
1	0.3451	0.3598	0.3724	0.3829	0.3915	0.3983	0.4034	0.4069	0.4089	0.4096	4
2	0.0853	0.0981	0.1113	0.1247	0.1382	0.1517	0.1652	0.1786	0.1919	0.2048	3
3	0.0105	0.0134	0.0166	0.0203	0.0244	0.0289	0.0338	0.0392	0.0450	0.0512	2
4	0.0007	0.0009	0.0012	0.0017	0.0022	0.0028	0.0035	0.0043	0.0053	0.0064	1
5	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002	0.0003	0
	$q = 0.89$	$q = 0.88$	$q = 0.87$	$q = 0.86$	$q = 0.85$	$q = 0.84$	$q = 0.83$	$q = 0.82$	$q = 0.81$	$q = 0.80$	

x	$p = 0.21$	$p = 0.22$	$p = 0.23$	$p = 0.24$	$p = 0.25$	$p = 0.26$	$p = 0.27$	$p = 0.28$	$p = 0.29$	$p = 0.30$	$n - x$
0	0.3077	0.2887	0.2707	0.2536	0.2373	0.2219	0.2073	0.1935	0.1804	0.1681	5
1	0.4090	0.4072	0.4043	0.4003	0.3955	0.3898	0.3834	0.3762	0.3685	0.3602	4
2	0.2174	0.2297	0.2415	0.2529	0.2637	0.2739	0.2836	0.2926	0.3010	0.3087	3
3	0.0578	0.0648	0.0721	0.0798	0.0879	0.0962	0.1049	0.1138	0.1229	0.1323	2
4	0.0077	0.0091	0.0108	0.0126	0.0146	0.0169	0.0194	0.0221	0.0251	0.0284	1
5	0.0004	0.0005	0.0006	0.0008	0.0010	0.0012	0.0014	0.0017	0.0021	0.0024	0
	$q = 0.79$	$q = 0.78$	$q = 0.77$	$q = 0.76$	$q = 0.75$	$q = 0.74$	$q = 0.73$	$q = 0.72$	$q = 0.71$	$q = 0.70$	

x	$p = 0.31$	$p = 0.32$	$p = 0.33$	$p = 0.34$	$p = 0.35$	$p = 0.36$	$p = 0.37$	$p = 0.38$	$p = 0.39$	$p = 0.40$	$n - x$
0	0.1564	0.1454	0.1350	0.1252	0.1160	0.1074	0.0992	0.0916	0.0845	0.0778	5
1	0.3513	0.3421	0.3325	0.3226	0.3124	0.3020	0.2914	0.2808	0.2700	0.2592	4
2	0.3157	0.3220	0.3275	0.3323	0.3364	0.3397	0.3423	0.3441	0.3452	0.3456	3
3	0.1418	0.1515	0.1613	0.1712	0.1811	0.1911	0.2010	0.2109	0.2207	0.2304	2
4	0.0319	0.0357	0.0397	0.0441	0.0488	0.0537	0.0590	0.0646	0.0706	0.0768	1
5	0.0029	0.0034	0.0039	0.0045	0.0053	0.0060	0.0069	0.0079	0.0090	0.0102	0
	$q = 0.69$	$q = 0.68$	$q = 0.67$	$q = 0.66$	$q = 0.65$	$q = 0.64$	$q = 0.63$	$q = 0.62$	$q = 0.61$	$q = 0.60$	

x	$p = 0.41$	$p = 0.42$	$p = 0.43$	$p = 0.44$	$p = 0.45$	$p = 0.46$	$p = 0.47$	$p = 0.48$	$p = 0.49$	$p = 0.50$	$n - x$
0	0.0715	0.0656	0.0602	0.0551	0.0503	0.0459	0.0418	0.0380	0.0345	0.0313	5
1	0.2484	0.2376	0.2270	0.2164	0.2059	0.1956	0.1854	0.1755	0.1657	0.1563	4
2	0.3452	0.3442	0.3424	0.3400	0.3369	0.3332	0.3289	0.3240	0.3185	0.3125	3
3	0.2399	0.2492	0.2583	0.2671	0.2757	0.2838	0.2916	0.2990	0.3060	0.3125	2
4	0.0834	0.0902	0.0974	0.1049	0.1128	0.1209	0.1293	0.1380	0.1470	0.1563	1
5	0.0116	0.0131	0.0147	0.0165	0.0185	0.0206	0.0229	0.0255	0.0282	0.0313	0
	$q = 0.59$	$q = 0.58$	$q = 0.57$	$q = 0.56$	$q = 0.55$	$q = 0.54$	$q = 0.53$	$q = 0.52$	$q = 0.51$	$q = 0.50$	

$n = 6$

x	$p = 0.01$	$p = 0.02$	$p = 0.03$	$p = 0.04$	$p = 0.05$	$p = 0.06$	$p = 0.07$	$p = 0.08$	$p = 0.09$	$p = 0.10$	$n - x$
0	0.9415	0.8858	0.8330	0.7828	0.7351	0.6899	0.6470	0.6064	0.5679	0.5314	6
1	0.0571	0.1085	0.1546	0.1957	0.2321	0.2642	0.2922	0.3164	0.3370	0.3543	5
2	0.0014	0.0055	0.0120	0.0204	0.0305	0.0422	0.0550	0.0688	0.0833	0.0984	4
3	0.0000	0.0002	0.0005	0.0011	0.0021	0.0036	0.0055	0.0080	0.0110	0.0146	3
4	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0003	0.0005	0.0008	0.0012	2
5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	1
6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0
	$q = 0.99$	$q = 0.98$	$q = 0.97$	$q = 0.96$	$q = 0.95$	$q = 0.94$	$q = 0.93$	$q = 0.92$	$q = 0.91$	$q = 0.90$	

x	$p = 0.11$	$p = 0.12$	$p = 0.13$	$p = 0.14$	$p = 0.15$	$p = 0.16$	$p = 0.17$	$p = 0.18$	$p = 0.19$	$p = 0.20$	$n - x$
0	0.4970	0.4644	0.4336	0.4046	0.3771	0.3513	0.3269	0.3040	0.2824	0.2621	6
1	0.3685	0.3800	0.3888	0.3952	0.3993	0.4015	0.4018	0.4004	0.3975	0.3932	5
2	0.1139	0.1295	0.1452	0.1608	0.1762	0.1912	0.2057	0.2197	0.2331	0.2458	4
3	0.0188	0.0236	0.0289	0.0349	0.0415	0.0486	0.0562	0.0643	0.0729	0.0819	3
4	0.0017	0.0024	0.0032	0.0043	0.0055	0.0069	0.0086	0.0106	0.0128	0.0154	2
5	0.0001	0.0001	0.0002	0.0003	0.0004	0.0005	0.0007	0.0009	0.0012	0.0015	1
6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0
	$q = 0.89$	$q = 0.88$	$q = 0.87$	$q = 0.86$	$q = 0.85$	$q = 0.84$	$q = 0.83$	$q = 0.82$	$q = 0.81$	$q = 0.80$	

x	$p = 0.21$	$p = 0.22$	$p = 0.23$	$p = 0.24$	$p = 0.25$	$p = 0.26$	$p = 0.27$	$p = 0.28$	$p = 0.29$	$p = 0.30$	$n - x$
0	0.2431	0.2252	0.2084	0.1927	0.1780	0.1642	0.1513	0.1393	0.1281	0.1176	6
1	0.3877	0.3811	0.3735	0.3651	0.3560	0.3462	0.3358	0.3251	0.3139	0.3025	5
2	0.2577	0.2687	0.2789	0.2882	0.2966	0.3041	0.3105	0.3160	0.3206	0.3241	4
3	0.0913	0.1011	0.1111	0.1214	0.1318	0.1424	0.1531	0.1639	0.1746	0.1852	3
4	0.0182	0.0214	0.0249	0.0287	0.0330	0.0375	0.0425	0.0478	0.0535	0.0595	2
5	0.0019	0.0024	0.0030	0.0036	0.0044	0.0053	0.0063	0.0074	0.0087	0.0102	1
6	0.0001	0.0001	0.0001	0.0002	0.0002	0.0003	0.0004	0.0005	0.0006	0.0007	0
	$q = 0.79$	$q = 0.78$	$q = 0.77$	$q = 0.76$	$q = 0.75$	$q = 0.74$	$q = 0.73$	$q = 0.72$	$q = 0.71$	$q = 0.70$	

x	$p = 0.31$	$p = 0.32$	$p = 0.33$	$p = 0.34$	$p = 0.35$	$p = 0.36$	$p = 0.37$	$p = 0.38$	$p = 0.39$	$p = 0.40$	$n - x$
0	0.1079	0.0989	0.0905	0.0827	0.0754	0.0687	0.0625	0.0568	0.0515	0.0467	6
1	0.2909	0.2792	0.2673	0.2555	0.2437	0.2319	0.2203	0.2089	0.1976	0.1866	5
2	0.3267	0.3284	0.3292	0.3290	0.3280	0.3261	0.3235	0.3201	0.3159	0.3110	4
3	0.1957	0.2061	0.2162	0.2260	0.2355	0.2446	0.2533	0.2616	0.2693	0.2765	3
4	0.0660	0.0727	0.0799	0.0873	0.0951	0.1032	0.1116	0.1202	0.1291	0.1382	2
5	0.0119	0.0137	0.0157	0.0180	0.0205	0.0232	0.0262	0.0295	0.0330	0.0369	1
6	0.0009	0.0011	0.0013	0.0015	0.0018	0.0022	0.0026	0.0030	0.0035	0.0041	0
	$q = 0.69$	$q = 0.68$	$q = 0.67$	$q = 0.66$	$q = 0.65$	$q = 0.64$	$q = 0.63$	$q = 0.62$	$q = 0.61$	$q = 0.60$	

x	$p = 0.41$	$p = 0.42$	$p = 0.43$	$p = 0.44$	$p = 0.45$	$p = 0.46$	$p = 0.47$	$p = 0.48$	$p = 0.49$	$p = 0.50$	$n - x$
0	0.0422	0.0381	0.0343	0.0308	0.0277	0.0248	0.0222	0.0198	0.0176	0.0156	6
1	0.1759	0.1654	0.1552	0.1454	0.1359	0.1267	0.1179	0.1095	0.1014	0.0938	5
2	0.3055	0.2994	0.2928	0.2856	0.2780	0.2699	0.2615	0.2527	0.2436	0.2344	4
3	0.2831	0.2891	0.2945	0.2992	0.3032	0.3065	0.3091	0.3110	0.3121	0.3125	3
4	0.1475	0.1570	0.1666	0.1763	0.1861	0.1958	0.2056	0.2153	0.2249	0.2344	2
5	0.0410	0.0455	0.0503	0.0554	0.0609	0.0667	0.0729	0.0795	0.0864	0.0938	1
6	0.0048	0.0055	0.0063	0.0073	0.0083	0.0095	0.0108	0.0122	0.0138	0.0156	0
	$q = 0.59$	$q = 0.58$	$q = 0.57$	$q = 0.56$	$q = 0.55$	$q = 0.54$	$q = 0.53$	$q = 0.52$	$q = 0.51$	$q = 0.50$	

$n = 7$

x	$p = 0.01$	$p = 0.02$	$p = 0.03$	$p = 0.04$	$p = 0.05$	$p = 0.06$	$p = 0.07$	$p = 0.08$	$p = 0.09$	$p = 0.10$	$n - x$
0	0.9321	0.8681	0.8080	0.7514	0.6983	0.6485	0.6017	0.5578	0.5168	0.4783	7
1	0.0659	0.1240	0.1749	0.2192	0.2573	0.2897	0.3170	0.3396	0.3578	0.3720	6
2	0.0020	0.0076	0.0162	0.0274	0.0406	0.0555	0.0716	0.0886	0.1061	0.1240	5
3	0.0000	0.0003	0.0008	0.0019	0.0036	0.0059	0.0090	0.0128	0.0175	0.0230	4
4	0.0000	0.0000	0.0000	0.0001	0.0002	0.0004	0.0007	0.0011	0.0017	0.0026	3
5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0002	2
6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1
7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0
	$q = 0.99$	$q = 0.98$	$q = 0.97$	$q = 0.96$	$q = 0.95$	$q = 0.94$	$q = 0.93$	$q = 0.92$	$q = 0.91$	$q = 0.90$	

x	$p = 0.11$	$p = 0.12$	$p = 0.13$	$p = 0.14$	$p = 0.15$	$p = 0.16$	$p = 0.17$	$p = 0.18$	$p = 0.19$	$p = 0.20$	$n - x$
0	0.4423	0.4087	0.3773	0.3479	0.3206	0.2951	0.2714	0.2493	0.2288	0.2097	7
1	0.3827	0.3901	0.3946	0.3965	0.3960	0.3935	0.3891	0.3830	0.3756	0.3670	6
2	0.1419	0.1596	0.1769	0.1936	0.2097	0.2248	0.2391	0.2523	0.2643	0.2753	5
3	0.0292	0.0363	0.0441	0.0525	0.0617	0.0714	0.0816	0.0923	0.1033	0.1147	4
4	0.0036	0.0049	0.0066	0.0086	0.0109	0.0136	0.0167	0.0203	0.0242	0.0287	3
5	0.0003	0.0004	0.0006	0.0008	0.0012	0.0016	0.0021	0.0027	0.0034	0.0043	2
6	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001	0.0002	0.0003	0.0004	1
7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0
	$q = 0.89$	$q = 0.88$	$q = 0.87$	$q = 0.86$	$q = 0.85$	$q = 0.84$	$q = 0.83$	$q = 0.82$	$q = 0.81$	$q = 0.80$	

x	$p = 0.21$	$p = 0.22$	$p = 0.23$	$p = 0.24$	$p = 0.25$	$p = 0.26$	$p = 0.27$	$p = 0.28$	$p = 0.29$	$p = 0.30$	$n - x$
0	0.1920	0.1757	0.1605	0.1465	0.1335	0.1215	0.1105	0.1003	0.0910	0.0824	7
1	0.3573	0.3468	0.3356	0.3237	0.3115	0.2989	0.2860	0.2731	0.2600	0.2471	6
2	0.2850	0.2935	0.3007	0.3067	0.3115	0.3150	0.3174	0.3186	0.3186	0.3177	5
3	0.1263	0.1379	0.1497	0.1614	0.1730	0.1845	0.1956	0.2065	0.2169	0.2269	4
4	0.0336	0.0389	0.0447	0.0510	0.0577	0.0648	0.0724	0.0803	0.0886	0.0972	3
5	0.0054	0.0066	0.0080	0.0097	0.0115	0.0137	0.0161	0.0187	0.0217	0.0250	2
6	0.0005	0.0006	0.0008	0.0010	0.0013	0.0016	0.0020	0.0024	0.0030	0.0036	1
7	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002	0
	$q = 0.79$	$q = 0.78$	$q = 0.77$	$q = 0.76$	$q = 0.75$	$q = 0.74$	$q = 0.73$	$q = 0.72$	$q = 0.71$	$q = 0.70$	

x	$p = 0.31$	$p = 0.32$	$p = 0.33$	$p = 0.34$	$p = 0.35$	$p = 0.36$	$p = 0.37$	$p = 0.38$	$p = 0.39$	$p = 0.40$	$n - x$
0	0.0745	0.0672	0.0606	0.0546	0.0490	0.0440	0.0394	0.0352	0.0314	0.0280	7
1	0.2342	0.2215	0.2090	0.1967	0.1848	0.1732	0.1619	0.1511	0.1407	0.1306	6
2	0.3156	0.3127	0.3088	0.3040	0.2985	0.2922	0.2853	0.2778	0.2698	0.2613	5
3	0.2363	0.2452	0.2535	0.2610	0.2679	0.2740	0.2793	0.2838	0.2875	0.2903	4
4	0.1062	0.1154	0.1248	0.1345	0.1442	0.1541	0.1640	0.1739	0.1838	0.1935	3
5	0.0286	0.0326	0.0369	0.0416	0.0466	0.0520	0.0578	0.0640	0.0705	0.0774	2
6	0.0043	0.0051	0.0061	0.0071	0.0084	0.0098	0.0113	0.0131	0.0150	0.0172	1
7	0.0003	0.0003	0.0004	0.0005	0.0006	0.0008	0.0009	0.0011	0.0014	0.0016	0
	$q = 0.69$	$q = 0.68$	$q = 0.67$	$q = 0.66$	$q = 0.65$	$q = 0.64$	$q = 0.63$	$q = 0.62$	$q = 0.61$	$q = 0.60$	

x	$p = 0.41$	$p = 0.42$	$p = 0.43$	$p = 0.44$	$p = 0.45$	$p = 0.46$	$p = 0.47$	$p = 0.48$	$p = 0.49$	$p = 0.50$	$n - x$
0	0.0249	0.0221	0.0195	0.0173	0.0152	0.0134	0.0117	0.0103	0.0090	0.0078	7
1	0.1211	0.1119	0.1032	0.0950	0.0872	0.0798	0.0729	0.0664	0.0604	0.0547	6
2	0.2524	0.2431	0.2336	0.2239	0.2140	0.2040	0.1940	0.1840	0.1740	0.1641	5
3	0.2923	0.2934	0.2937	0.2932	0.2918	0.2897	0.2867	0.2830	0.2786	0.2734	4
4	0.2031	0.2125	0.2216	0.2304	0.2388	0.2468	0.2543	0.2612	0.2676	0.2734	3
5	0.0847	0.0923	0.1003	0.1086	0.1172	0.1261	0.1353	0.1447	0.1543	0.1641	2
6	0.0196	0.0223	0.0252	0.0284	0.0320	0.0358	0.0400	0.0445	0.0494	0.0547	1
7	0.0019	0.0023	0.0027	0.0032	0.0037	0.0044	0.0051	0.0059	0.0068	0.0078	0
	$q = 0.59$	$q = 0.58$	$q = 0.57$	$q = 0.56$	$q = 0.55$	$q = 0.54$	$q = 0.53$	$q = 0.52$	$q = 0.51$	$q = 0.50$	

$n = 8$

x	$p = 0.01$	$p = 0.02$	$p = 0.03$	$p = 0.04$	$p = 0.05$	$p = 0.06$	$p = 0.07$	$p = 0.08$	$p = 0.09$	$p = 0.10$	$n - x$
0	0.9227	0.8508	0.7837	0.7214	0.6634	0.6096	0.5596	0.5132	0.4703	0.4305	8
1	0.0746	0.1389	0.1939	0.2405	0.2793	0.3113	0.3370	0.3570	0.3721	0.3826	7
2	0.0026	0.0099	0.0210	0.0351	0.0515	0.0695	0.0888	0.1087	0.1288	0.1488	6
3	0.0001	0.0004	0.0013	0.0029	0.0054	0.0089	0.0134	0.0189	0.0255	0.0331	5
4	0.0000	0.0000	0.0001	0.0002	0.0004	0.0007	0.0013	0.0021	0.0031	0.0046	4
5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0002	0.0004	3
6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	2
7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1
8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0
	$q = 0.99$	$q = 0.98$	$q = 0.97$	$q = 0.96$	$q = 0.95$	$q = 0.94$	$q = 0.93$	$q = 0.92$	$q = 0.91$	$q = 0.90$	

x	$p = 0.11$	$p = 0.12$	$p = 0.13$	$p = 0.14$	$p = 0.15$	$p = 0.16$	$p = 0.17$	$p = 0.18$	$p = 0.19$	$p = 0.20$	$n - x$
0	0.3937	0.3596	0.3282	0.2992	0.2725	0.2479	0.2252	0.2044	0.1853	0.1678	8
1	0.3892	0.3923	0.3923	0.3897	0.3847	0.3777	0.3691	0.3590	0.3477	0.3355	7
2	0.1684	0.1872	0.2052	0.2220	0.2376	0.2518	0.2646	0.2758	0.2855	0.2936	6
3	0.0416	0.0511	0.0613	0.0723	0.0839	0.0959	0.1084	0.1211	0.1339	0.1468	5
4	0.0064	0.0087	0.0115	0.0147	0.0185	0.0228	0.0277	0.0332	0.0393	0.0459	4
5	0.0006	0.0009	0.0014	0.0019	0.0026	0.0035	0.0045	0.0058	0.0074	0.0092	3
6	0.0000	0.0001	0.0001	0.0002	0.0002	0.0003	0.0005	0.0006	0.0009	0.0011	2
7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	1
8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0
	$q = 0.89$	$q = 0.88$	$q = 0.87$	$q = 0.86$	$q = 0.85$	$q = 0.84$	$q = 0.83$	$q = 0.82$	$q = 0.81$	$q = 0.80$	

x	$p = 0.21$	$p = 0.22$	$p = 0.23$	$p = 0.24$	$p = 0.25$	$p = 0.26$	$p = 0.27$	$p = 0.28$	$p = 0.29$	$p = 0.30$	$n - x$
0	0.1517	0.1370	0.1236	0.1113	0.1001	0.0899	0.0806	0.0722	0.0646	0.0576	8
1	0.3226	0.3092	0.2953	0.2812	0.2670	0.2527	0.2386	0.2247	0.2110	0.1977	7
2	0.3002	0.3052	0.3087	0.3108	0.3115	0.3108	0.3089	0.3058	0.3017	0.2965	6
3	0.1596	0.1722	0.1844	0.1963	0.2076	0.2184	0.2285	0.2379	0.2464	0.2541	5
4	0.0530	0.0607	0.0689	0.0775	0.0865	0.0959	0.1056	0.1156	0.1258	0.1361	4
5	0.0113	0.0137	0.0165	0.0196	0.0231	0.0270	0.0313	0.0360	0.0411	0.0467	3
6	0.0015	0.0019	0.0025	0.0031	0.0038	0.0047	0.0058	0.0070	0.0084	0.0100	2
7	0.0001	0.0002	0.0002	0.0003	0.0004	0.0005	0.0006	0.0008	0.0010	0.0012	1
8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0
	$q = 0.79$	$q = 0.78$	$q = 0.77$	$q = 0.76$	$q = 0.75$	$q = 0.74$	$q = 0.73$	$q = 0.72$	$q = 0.71$	$q = 0.70$	

x	$p = 0.31$	$p = 0.32$	$p = 0.33$	$p = 0.34$	$p = 0.35$	$p = 0.36$	$p = 0.37$	$p = 0.38$	$p = 0.39$	$p = 0.40$	$n - x$
0	0.0514	0.0457	0.0406	0.0360	0.0319	0.0281	0.0248	0.0218	0.0192	0.0168	8
1	0.1847	0.1721	0.1600	0.1484	0.1373	0.1267	0.1166	0.1071	0.0981	0.0896	7
2	0.2904	0.2835	0.2758	0.2675	0.2587	0.2494	0.2397	0.2297	0.2194	0.2090	6
3	0.2609	0.2668	0.2717	0.2756	0.2786	0.2805	0.2815	0.2815	0.2806	0.2787	5
4	0.1465	0.1569	0.1673	0.1775	0.1875	0.1973	0.2067	0.2157	0.2242	0.2322	4
5	0.0527	0.0591	0.0659	0.0732	0.0808	0.0888	0.0971	0.1058	0.1147	0.1239	3
6	0.0118	0.0139	0.0162	0.0188	0.0217	0.0250	0.0285	0.0324	0.0367	0.0413	2
7	0.0015	0.0019	0.0023	0.0028	0.0033	0.0040	0.0048	0.0057	0.0067	0.0079	1
8	0.0001	0.0001	0.0001	0.0002	0.0002	0.0003	0.0004	0.0004	0.0005	0.0007	0
	$q = 0.69$	$q = 0.68$	$q = 0.67$	$q = 0.66$	$q = 0.65$	$q = 0.64$	$q = 0.63$	$q = 0.62$	$q = 0.61$	$q = 0.60$	

x	p = 0.41	p = 0.42	p = 0.43	p = 0.44	p = 0.45	p = 0.46	p = 0.47	p = 0.48	p = 0.49	p = 0.50	n − x
0	0.0147	0.0128	0.0111	0.0097	0.0084	0.0072	0.0062	0.0053	0.0046	0.0039	8
1	0.0816	0.0742	0.0672	0.0608	0.0548	0.0493	0.0442	0.0395	0.0352	0.0313	7
2	0.1985	0.1880	0.1776	0.1672	0.1569	0.1469	0.1371	0.1275	0.1183	0.1094	6
3	0.2759	0.2723	0.2679	0.2627	0.2568	0.2503	0.2431	0.2355	0.2273	0.2188	5
4	0.2397	0.2465	0.2526	0.2580	0.2627	0.2665	0.2695	0.2717	0.2730	0.2734	4
5	0.1332	0.1428	0.1525	0.1622	0.1719	0.1816	0.1912	0.2006	0.2098	0.2188	3
6	0.0463	0.0517	0.0575	0.0637	0.0703	0.0774	0.0848	0.0926	0.1008	0.1094	2
7	0.0092	0.0107	0.0124	0.0143	0.0164	0.0188	0.0215	0.0244	0.0277	0.0313	1
8	0.0008	0.0010	0.0012	0.0014	0.0017	0.0020	0.0024	0.0028	0.0033	0.0039	0
	q = 0.59	q = 0.58	q = 0.57	q = 0.56	q = 0.55	q = 0.54	q = 0.53	q = 0.52	q = 0.51	q = 0.50	

$n = 9$

x	p = 0.01	p = 0.02	p = 0.03	p = 0.04	p = 0.05	p = 0.06	p = 0.07	p = 0.08	p = 0.09	p = 0.10	n − x
0	0.9135	0.8337	0.7602	0.6925	0.6302	0.5730	0.5204	0.4722	0.4279	0.3874	9
1	0.0830	0.1531	0.2116	0.2597	0.2985	0.3292	0.3525	0.3695	0.3809	0.3874	8
2	0.0034	0.0125	0.0262	0.0433	0.0629	0.0840	0.1061	0.1285	0.1507	0.1722	7
3	0.0001	0.0006	0.0019	0.0042	0.0077	0.0125	0.0186	0.0261	0.0348	0.0446	6
4	0.0000	0.0000	0.0001	0.0003	0.0006	0.0012	0.0021	0.0034	0.0052	0.0074	5
5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0003	0.0005	0.0008	4
6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	3
7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	2
8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1
9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0
	q = 0.99	q = 0.98	q = 0.97	q = 0.96	q = 0.95	q = 0.94	q = 0.93	q = 0.92	q = 0.91	q = 0.90	

x	p = 0.11	p = 0.12	p = 0.13	p = 0.14	p = 0.15	p = 0.16	p = 0.17	p = 0.18	p = 0.19	p = 0.20	n − x
0	0.3504	0.3165	0.2855	0.2573	0.2316	0.2082	0.1869	0.1676	0.1501	0.1342	9
1	0.3897	0.3884	0.3840	0.3770	0.3679	0.3569	0.3446	0.3312	0.3169	0.3020	8
2	0.1927	0.2119	0.2295	0.2455	0.2597	0.2720	0.2823	0.2908	0.2973	0.3020	7
3	0.0556	0.0674	0.0800	0.0933	0.1069	0.1209	0.1349	0.1489	0.1627	0.1762	6
4	0.0103	0.0138	0.0179	0.0228	0.0283	0.0345	0.0415	0.0490	0.0573	0.0661	5
5	0.0013	0.0019	0.0027	0.0037	0.0050	0.0066	0.0085	0.0108	0.0134	0.0165	4
6	0.0001	0.0002	0.0003	0.0004	0.0006	0.0008	0.0012	0.0016	0.0021	0.0028	3
7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001	0.0002	0.0003	2
8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1
9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0
	q = 0.89	q = 0.88	q = 0.87	q = 0.86	q = 0.85	q = 0.84	q = 0.83	q = 0.82	q = 0.81	q = 0.80	

x	p = 0.21	p = 0.22	p = 0.23	p = 0.24	p = 0.25	p = 0.26	p = 0.27	p = 0.28	p = 0.29	p = 0.30	n − x
0	0.1199	0.1069	0.0952	0.0846	0.0751	0.0665	0.0589	0.0520	0.0458	0.0404	9
1	0.2867	0.2713	0.2558	0.2404	0.2253	0.2104	0.1960	0.1820	0.1685	0.1556	8
2	0.3049	0.3061	0.3056	0.3037	0.3003	0.2957	0.2899	0.2831	0.2754	0.2668	7
3	0.1891	0.2014	0.2130	0.2238	0.2336	0.2424	0.2502	0.2569	0.2624	0.2668	6
4	0.0754	0.0852	0.0954	0.1060	0.1168	0.1278	0.1388	0.1499	0.1608	0.1715	5
5	0.0200	0.0240	0.0285	0.0335	0.0389	0.0449	0.0513	0.0583	0.0657	0.0735	4
6	0.0036	0.0045	0.0057	0.0070	0.0087	0.0105	0.0127	0.0151	0.0179	0.0210	3
7	0.0004	0.0005	0.0007	0.0010	0.0012	0.0016	0.0020	0.0025	0.0031	0.0039	2
8	0.0000	0.0000	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002	0.0003	0.0004	1
9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0
	q = 0.79	q = 0.78	q = 0.77	q = 0.76	q = 0.75	q = 0.74	q = 0.73	q = 0.72	q = 0.71	q = 0.70	

x	p = 0.31	p = 0.32	p = 0.33	p = 0.34	p = 0.35	p = 0.36	p = 0.37	p = 0.38	p = 0.39	p = 0.40	n − x
0	0.0355	0.0311	0.0272	0.0238	0.0207	0.0180	0.0156	0.0135	0.0117	0.0101	9
1	0.1433	0.1317	0.1206	0.1102	0.1004	0.0912	0.0826	0.0747	0.0673	0.0605	8
2	0.2576	0.2478	0.2376	0.2270	0.2162	0.2052	0.1941	0.1831	0.1721	0.1612	7
3	0.2701	0.2721	0.2731	0.2729	0.2716	0.2693	0.2660	0.2618	0.2567	0.2508	6
4	0.1820	0.1921	0.2017	0.2109	0.2194	0.2272	0.2344	0.2407	0.2462	0.2508	5
5	0.0818	0.0904	0.0994	0.1086	0.1181	0.1278	0.1376	0.1475	0.1574	0.1672	4
6	0.0245	0.0284	0.0326	0.0373	0.0424	0.0479	0.0539	0.0603	0.0671	0.0743	3
7	0.0047	0.0057	0.0069	0.0082	0.0098	0.0116	0.0136	0.0158	0.0184	0.0212	2
8	0.0005	0.0007	0.0008	0.0011	0.0013	0.0016	0.0020	0.0024	0.0029	0.0035	1
9	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002	0.0003	0
	q = 0.69	q = 0.68	q = 0.67	q = 0.66	q = 0.65	q = 0.64	q = 0.63	q = 0.62	q = 0.61	q = 0.60	

x	p = 0.41	p = 0.42	p = 0.43	p = 0.44	p = 0.45	p = 0.46	p = 0.47	p = 0.48	p = 0.49	p = 0.50	n − x
0	0.0087	0.0074	0.0064	0.0054	0.0046	0.0039	0.0033	0.0028	0.0023	0.0020	9
1	0.0542	0.0484	0.0431	0.0383	0.0339	0.0299	0.0263	0.0231	0.0202	0.0176	8
2	0.1506	0.1402	0.1301	0.1204	0.1110	0.1020	0.0934	0.0853	0.0776	0.0703	7
3	0.2442	0.2369	0.2291	0.2207	0.2119	0.2027	0.1933	0.1837	0.1739	0.1641	6
4	0.2545	0.2573	0.2592	0.2601	0.2600	0.2590	0.2571	0.2543	0.2506	0.2461	5
5	0.1769	0.1863	0.1955	0.2044	0.2128	0.2207	0.2280	0.2347	0.2408	0.2461	4
6	0.0819	0.0900	0.0983	0.1070	0.1160	0.1253	0.1348	0.1445	0.1542	0.1641	3
7	0.0244	0.0279	0.0318	0.0360	0.0407	0.0458	0.0512	0.0571	0.0635	0.0703	2
8	0.0042	0.0051	0.0060	0.0071	0.0083	0.0097	0.0114	0.0132	0.0153	0.0176	1
9	0.0003	0.0004	0.0005	0.0006	0.0008	0.0009	0.0011	0.0014	0.0016	0.0020	0
	q = 0.59	q = 0.58	q = 0.57	q = 0.56	q = 0.55	q = 0.54	q = 0.53	q = 0.52	q = 0.51	q = 0.50	

$n = 10$

x	p = 0.01	p = 0.02	p = 0.03	p = 0.04	p = 0.05	p = 0.06	p = 0.07	p = 0.08	p = 0.09	p = 0.10	n − x
0	0.9044	0.8171	0.7374	0.6648	0.5987	0.5386	0.4840	0.4344	0.3894	0.3487	10
1	0.0914	0.1667	0.2281	0.2770	0.3151	0.3438	0.3643	0.3777	0.3851	0.3874	9
2	0.0042	0.0153	0.0317	0.0519	0.0746	0.0988	0.1234	0.1478	0.1714	0.1937	8
3	0.0001	0.0008	0.0026	0.0058	0.0105	0.0168	0.0248	0.0343	0.0452	0.0574	7
4	0.0000	0.0000	0.0001	0.0004	0.0010	0.0019	0.0033	0.0052	0.0078	0.0112	6
5	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0003	0.0005	0.0009	0.0015	5
6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	4
7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	3
8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	2
9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1
10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0
	q = 0.99	q = 0.98	q = 0.97	q = 0.96	q = 0.95	q = 0.94	q = 0.93	q = 0.92	q = 0.91	q = 0.90	

x	p = 0.11	p = 0.12	p = 0.13	p = 0.14	p = 0.15	p = 0.16	p = 0.17	p = 0.18	p = 0.19	p = 0.20	n − x
0	0.3118	0.2785	0.2484	0.2213	0.1969	0.1749	0.1552	0.1374	0.1216	0.1074	10
1	0.3854	0.3798	0.3712	0.3603	0.3474	0.3331	0.3178	0.3017	0.2852	0.2684	9
2	0.2143	0.2330	0.2496	0.2639	0.2759	0.2856	0.2929	0.2980	0.3010	0.3020	8
3	0.0706	0.0847	0.0995	0.1146	0.1298	0.1450	0.1600	0.1745	0.1883	0.2013	7
4	0.0153	0.0202	0.0260	0.0326	0.0401	0.0483	0.0573	0.0670	0.0773	0.0881	6
5	0.0023	0.0033	0.0047	0.0064	0.0085	0.0111	0.0141	0.0177	0.0218	0.0264	5
6	0.0002	0.0004	0.0006	0.0009	0.0012	0.0018	0.0024	0.0032	0.0043	0.0055	4
7	0.0000	0.0000	0.0000	0.0001	0.0001	0.0002	0.0003	0.0004	0.0006	0.0008	3
8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	2
9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1
10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0
	q = 0.89	q = 0.88	q = 0.87	q = 0.86	q = 0.85	q = 0.84	q = 0.83	q = 0.82	q = 0.81	q = 0.80	

EK 2: Poisson Olasılıkları Tablosu

λt

x	0.005	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	0.9950	0.9900	0.9802	0.9704	0.9608	0.9512	0.9418	0.9324	0.9231	0.9139
1	0.0050	0.0099	0.0196	0.0291	0.0384	0.0476	0.0565	0.0653	0.0738	0.0823
2	0.0000	0.0000	0.0002	0.0004	0.0008	0.0012	0.0017	0.0023	0.0030	0.0037
3	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001

λt

x	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
0	0.9048	0.8187	0.7408	0.6703	0.6065	0.5488	0.4966	0.4493	0.4066	0.3679
1	0.0905	0.1637	0.2222	0.2681	0.3033	0.3293	0.3476	0.3595	0.3659	0.3679
2	0.0045	0.0164	0.0333	0.0536	0.0758	0.0988	0.1217	0.1438	0.1647	0.1839
3	0.0002	0.0011	0.0033	0.0072	0.0126	0.0198	0.0284	0.0383	0.0494	0.0613
4	0.0000	0.0001	0.0003	0.0007	0.0016	0.0030	0.0050	0.0077	0.0111	0.0153
5	0.0000	0.0000	0.0000	0.0001	0.0002	0.0004	0.0007	0.0012	0.0020	0.0031
6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0003	0.0005
7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001

λt

x	1.10	1.20	1.30	1.40	1.50	1.60	1.70	1.80	1.90	2.00
0	0.3329	0.3012	0.2725	0.2466	0.2231	0.2019	0.1827	0.1653	0.1496	0.1353
1	0.3662	0.3614	0.3543	0.3452	0.3347	0.3230	0.3106	0.2975	0.2842	0.2707
2	0.2014	0.2169	0.2303	0.2417	0.2510	0.2584	0.2640	0.2678	0.2700	0.2707
3	0.0738	0.0867	0.0998	0.1128	0.1255	0.1378	0.1496	0.1607	0.1710	0.1804
4	0.0203	0.0260	0.0324	0.0395	0.0471	0.0551	0.0636	0.0723	0.0812	0.0902
5	0.0045	0.0062	0.0084	0.0111	0.0141	0.0176	0.0216	0.0260	0.0309	0.0361
6	0.0008	0.0012	0.0018	0.0026	0.0035	0.0047	0.0061	0.0078	0.0098	0.0120
7	0.0001	0.0002	0.0003	0.0005	0.0008	0.0011	0.0015	0.0020	0.0027	0.0034
8	0.0000	0.0000	0.0001	0.0001	0.0001	0.0002	0.0003	0.0005	0.0006	0.0009
9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001	0.0002

λt

x	2.10	2.20	2.30	2.40	2.50	2.60	2.70	2.80	2.90	3.00
0	0.1225	0.1108	0.1003	0.0907	0.0821	0.0743	0.0672	0.0608	0.0550	0.0498
1	0.2572	0.2438	0.2306	0.2177	0.2052	0.1931	0.1815	0.1703	0.1596	0.1494
2	0.2700	0.2681	0.2652	0.2613	0.2565	0.2510	0.2450	0.2384	0.2314	0.2240
3	0.1890	0.1966	0.2033	0.2090	0.2138	0.2176	0.2205	0.2225	0.2237	0.2240
4	0.0992	0.1082	0.1169	0.1254	0.1336	0.1414	0.1488	0.1557	0.1622	0.1680
5	0.0417	0.0476	0.0538	0.0602	0.0668	0.0735	0.0804	0.0872	0.0940	0.1008
6	0.0146	0.0174	0.0206	0.0241	0.0278	0.0319	0.0362	0.0407	0.0455	0.0504
7	0.0044	0.0055	0.0068	0.0083	0.0099	0.0118	0.0139	0.0163	0.0188	0.0216
8	0.0011	0.0015	0.0019	0.0025	0.0031	0.0038	0.0047	0.0057	0.0068	0.0081
9	0.0003	0.0004	0.0005	0.0007	0.0009	0.0011	0.0014	0.0018	0.0022	0.0027
10	0.0001	0.0001	0.0001	0.0002	0.0002	0.0003	0.0004	0.0005	0.0006	0.0008
11	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001	0.0002	0.0002
12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001

λt

x	3.10	3.20	3.30	3.40	3.50	3.60	3.70	3.80	3.90	4.00
0	0.0450	0.0408	0.0369	0.0334	0.0302	0.0273	0.0247	0.0224	0.0202	0.0183
1	0.1397	0.1304	0.1217	0.1135	0.1057	0.0984	0.0915	0.0850	0.0789	0.0733
2	0.2165	0.2087	0.2008	0.1929	0.1850	0.1771	0.1692	0.1615	0.1539	0.1465
3	0.2237	0.2226	0.2209	0.2186	0.2158	0.2125	0.2087	0.2046	0.2001	0.1954
4	0.1733	0.1781	0.1823	0.1858	0.1888	0.1912	0.1931	0.1944	0.1951	0.1954
5	0.1075	0.1140	0.1203	0.1264	0.1322	0.1377	0.1429	0.1477	0.1522	0.1563
6	0.0555	0.0608	0.0662	0.0716	0.0771	0.0826	0.0881	0.0936	0.0989	0.1042
7	0.0246	0.0278	0.0312	0.0348	0.0385	0.0425	0.0466	0.0508	0.0551	0.0595
8	0.0095	0.0111	0.0129	0.0148	0.0169	0.0191	0.0215	0.0241	0.0269	0.0298
9	0.0033	0.0040	0.0047	0.0056	0.0066	0.0076	0.0089	0.0102	0.0116	0.0132
10	0.0010	0.0013	0.0016	0.0019	0.0023	0.0028	0.0033	0.0039	0.0045	0.0053
11	0.0003	0.0004	0.0005	0.0006	0.0007	0.0009	0.0011	0.0013	0.0016	0.0019
12	0.0001	0.0001	0.0001	0.0002	0.0002	0.0003	0.0003	0.0004	0.0005	0.0006
13	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002
14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001

λt

x	4.10	4.20	4.30	4.40	4.50	4.60	4.70	4.80	4.90	5.00
0	0.0166	0.0150	0.0136	0.0123	0.0111	0.0101	0.0091	0.0082	0.0074	0.0067
1	0.0679	0.0630	0.0583	0.0540	0.0500	0.0462	0.0427	0.0395	0.0365	0.0337
2	0.1393	0.1323	0.1254	0.1188	0.1125	0.1063	0.1005	0.0948	0.0894	0.0842
3	0.1904	0.1852	0.1798	0.1743	0.1687	0.1631	0.1574	0.1517	0.1460	0.1404
4	0.1951	0.1944	0.1933	0.1917	0.1898	0.1875	0.1849	0.1820	0.1789	0.1755
5	0.1600	0.1633	0.1662	0.1687	0.1708	0.1725	0.1738	0.1747	0.1753	0.1755
6	0.1093	0.1143	0.1191	0.1237	0.1281	0.1323	0.1362	0.1398	0.1432	0.1462
7	0.0640	0.0686	0.0732	0.0778	0.0824	0.0869	0.0914	0.0959	0.1002	0.1044
8	0.0328	0.0360	0.0393	0.0428	0.0463	0.0500	0.0537	0.0575	0.0614	0.0653
9	0.0150	0.0168	0.0188	0.0209	0.0232	0.0255	0.0281	0.0307	0.0334	0.0363
10	0.0061	0.0071	0.0081	0.0092	0.0104	0.0118	0.0132	0.0147	0.0164	0.0181
11	0.0023	0.0027	0.0032	0.0037	0.0043	0.0049	0.0056	0.0064	0.0073	0.0082
12	0.0008	0.0009	0.0011	0.0013	0.0016	0.0019	0.0022	0.0026	0.0030	0.0034
13	0.0002	0.0003	0.0004	0.0005	0.0006	0.0007	0.0008	0.0009	0.0011	0.0013
14	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002	0.0003	0.0003	0.0004	0.0005
15	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002

λt

x	5.10	5.20	5.30	5.40	5.50	5.60	5.70	5.80	5.90	6.00
0	0.0061	0.0055	0.0050	0.0045	0.0041	0.0037	0.0033	0.0030	0.0027	0.0025
1	0.0311	0.0287	0.0265	0.0244	0.0225	0.0207	0.0191	0.0176	0.0162	0.0149
2	0.0793	0.0746	0.0701	0.0659	0.0618	0.0580	0.0544	0.0509	0.0477	0.0446
3	0.1348	0.1293	0.1239	0.1185	0.1133	0.1082	0.1033	0.0985	0.0938	0.0892
4	0.1719	0.1681	0.1641	0.1600	0.1558	0.1515	0.1472	0.1428	0.1383	0.1339
5	0.1753	0.1748	0.1740	0.1728	0.1714	0.1697	0.1678	0.1656	0.1632	0.1606
6	0.1490	0.1515	0.1537	0.1555	0.1571	0.1584	0.1594	0.1601	0.1605	0.1606
7	0.1086	0.1125	0.1163	0.1200	0.1234	0.1267	0.1298	0.1326	0.1353	0.1377
8	0.0692	0.0731	0.0771	0.0810	0.0849	0.0887	0.0925	0.0962	0.0998	0.1033
9	0.0392	0.0423	0.0454	0.0486	0.0519	0.0552	0.0586	0.0620	0.0654	0.0688
10	0.0200	0.0220	0.0241	0.0262	0.0285	0.0309	0.0334	0.0359	0.0386	0.0413
11	0.0093	0.0104	0.0116	0.0129	0.0143	0.0157	0.0173	0.0190	0.0207	0.0225
12	0.0039	0.0045	0.0051	0.0058	0.0065	0.0073	0.0082	0.0092	0.0102	0.0113
13	0.0015	0.0018	0.0021	0.0024	0.0028	0.0032	0.0036	0.0041	0.0046	0.0052
14	0.0006	0.0007	0.0008	0.0009	0.0011	0.0013	0.0015	0.0017	0.0019	0.0022
15	0.0002	0.0002	0.0003	0.0003	0.0004	0.0005	0.0006	0.0007	0.0008	0.0009
16	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002	0.0002	0.0003	0.0003
17	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001	0.0001	0.0001

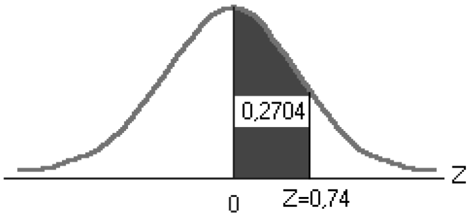
	λt									
x	6.10	6.20	6.30	6.40	6.50	6.60	6.70	6.80	6.90	7.00
0	0.0022	0.0020	0.0018	0.0017	0.0015	0.0014	0.0012	0.0011	0.0010	0.0009
1	0.0137	0.0126	0.0116	0.0106	0.0098	0.0090	0.0082	0.0076	0.0070	0.0064
2	0.0417	0.0390	0.0364	0.0340	0.0318	0.0296	0.0276	0.0258	0.0240	0.0223
3	0.0848	0.0806	0.0765	0.0726	0.0688	0.0652	0.0617	0.0584	0.0552	0.0521
4	0.1294	0.1249	0.1205	0.1162	0.1118	0.1076	0.1034	0.0992	0.0952	0.0912
5	0.1579	0.1549	0.1519	0.1487	0.1454	0.1420	0.1385	0.1349	0.1314	0.1277
6	0.1605	0.1601	0.1595	0.1586	0.1575	0.1562	0.1546	0.1529	0.1511	0.1490
7	0.1399	0.1418	0.1435	0.1450	0.1462	0.1472	0.1480	0.1486	0.1489	0.1490
8	0.1066	0.1099	0.1130	0.1160	0.1188	0.1215	0.1240	0.1263	0.1284	0.1304
9	0.0723	0.0757	0.0791	0.0825	0.0858	0.0891	0.0923	0.0954	0.0985	0.1014
10	0.0441	0.0469	0.0498	0.0528	0.0558	0.0588	0.0618	0.0649	0.0679	0.0710
11	0.0244	0.0265	0.0285	0.0307	0.0330	0.0353	0.0377	0.0401	0.0426	0.0452
12	0.0124	0.0137	0.0150	0.0164	0.0179	0.0194	0.0210	0.0227	0.0245	0.0263
13	0.0058	0.0065	0.0073	0.0081	0.0089	0.0099	0.0108	0.0119	0.0130	0.0142
14	0.0025	0.0029	0.0033	0.0037	0.0041	0.0046	0.0052	0.0058	0.0064	0.0071
15	0.0010	0.0012	0.0014	0.0016	0.0018	0.0020	0.0023	0.0026	0.0029	0.0033
16	0.0004	0.0005	0.0005	0.0006	0.0007	0.0008	0.0010	0.0011	0.0013	0.0014
17	0.0001	0.0002	0.0002	0.0002	0.0003	0.0003	0.0004	0.0004	0.0005	0.0006
18	0.0000	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002	0.0002
19	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001	0.0001

	λt									
x	7.10	7.20	7.30	7.40	7.50	7.60	7.70	7.80	7.90	8.00
0	0.0008	0.0007	0.0007	0.0006	0.0006	0.0005	0.0005	0.0004	0.0004	0.0003
1	0.0059	0.0054	0.0049	0.0045	0.0041	0.0038	0.0035	0.0032	0.0029	0.0027
2	0.0208	0.0194	0.0180	0.0167	0.0156	0.0145	0.0134	0.0125	0.0116	0.0107
3	0.0492	0.0464	0.0438	0.0413	0.0389	0.0366	0.0345	0.0324	0.0305	0.0286
4	0.0874	0.0836	0.0799	0.0764	0.0729	0.0696	0.0663	0.0632	0.0602	0.0573
5	0.1241	0.1204	0.1167	0.1130	0.1094	0.1057	0.1021	0.0986	0.0951	0.0916
6	0.1468	0.1445	0.1420	0.1394	0.1367	0.1339	0.1311	0.1282	0.1252	0.1221
7	0.1489	0.1486	0.1481	0.1474	0.1465	0.1454	0.1442	0.1428	0.1413	0.1396
8	0.1321	0.1337	0.1351	0.1363	0.1373	0.1381	0.1388	0.1392	0.1395	0.1396
9	0.1042	0.1070	0.1096	0.1121	0.1144	0.1167	0.1187	0.1207	0.1224	0.1241
10	0.0740	0.0770	0.0800	0.0829	0.0858	0.0887	0.0914	0.0941	0.0967	0.0993
11	0.0478	0.0504	0.0531	0.0558	0.0585	0.0613	0.0640	0.0667	0.0695	0.0722
12	0.0283	0.0303	0.0323	0.0344	0.0366	0.0388	0.0411	0.0434	0.0457	0.0481
13	0.0154	0.0168	0.0181	0.0196	0.0211	0.0227	0.0243	0.0260	0.0278	0.0296
14	0.0078	0.0086	0.0095	0.0104	0.0113	0.0123	0.0134	0.0145	0.0157	0.0169
15	0.0037	0.0041	0.0046	0.0051	0.0057	0.0062	0.0069	0.0075	0.0083	0.0090
16	0.0016	0.0019	0.0021	0.0024	0.0026	0.0030	0.0033	0.0037	0.0041	0.0045
17	0.0007	0.0008	0.0009	0.0010	0.0012	0.0013	0.0015	0.0017	0.0019	0.0021
18	0.0003	0.0003	0.0004	0.0004	0.0005	0.0006	0.0006	0.0007	0.0008	0.0009
19	0.0001	0.0001	0.0001	0.0002	0.0002	0.0002	0.0003	0.0003	0.0003	0.0004
20	0.0000	0.0000	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002
21	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001

					λt					
x	8.10	8.20	8.30	8.40	8.50	8.60	8.70	8.80	8.90	9.00
0	0.0003	0.0003	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0001	0.0001
1	0.0025	0.0023	0.0021	0.0019	0.0017	0.0016	0.0014	0.0013	0.0012	0.0011
2	0.0100	0.0092	0.0086	0.0079	0.0074	0.0068	0.0063	0.0058	0.0054	0.0050
3	0.0269	0.0252	0.0237	0.0222	0.0208	0.0195	0.0183	0.0171	0.0160	0.0150
4	0.0544	0.0517	0.0491	0.0466	0.0443	0.0420	0.0398	0.0377	0.0357	0.0337
5	0.0882	0.0849	0.0816	0.0784	0.0752	0.0722	0.0692	0.0663	0.0635	0.0607
6	0.1191	0.1160	0.1128	0.1097	0.1066	0.1034	0.1003	0.0972	0.0941	0.0911
7	0.1378	0.1358	0.1338	0.1317	0.1294	0.1271	0.1247	0.1222	0.1197	0.1171
8	0.1395	0.1392	0.1388	0.1382	0.1375	0.1366	0.1356	0.1344	0.1332	0.1318
9	0.1256	0.1269	0.1280	0.1290	0.1299	0.1306	0.1311	0.1315	0.1317	0.1318
10	0.1017	0.1040	0.1063	0.1084	0.1104	0.1123	0.1140	0.1157	0.1172	0.1186
11	0.0749	0.0776	0.0802	0.0828	0.0853	0.0878	0.0902	0.0925	0.0948	0.0970
12	0.0505	0.0530	0.0555	0.0579	0.0604	0.0629	0.0654	0.0679	0.0703	0.0728
13	0.0315	0.0334	0.0354	0.0374	0.0395	0.0416	0.0438	0.0459	0.0481	0.0504
14	0.0182	0.0196	0.0210	0.0225	0.0240	0.0256	0.0272	0.0289	0.0306	0.0324
15	0.0098	0.0107	0.0116	0.0126	0.0136	0.0147	0.0158	0.0169	0.0182	0.0194
16	0.0050	0.0055	0.0060	0.0066	0.0072	0.0079	0.0086	0.0093	0.0101	0.0109
17	0.0024	0.0026	0.0029	0.0033	0.0036	0.0040	0.0044	0.0048	0.0053	0.0058
18	0.0011	0.0012	0.0014	0.0015	0.0017	0.0019	0.0021	0.0024	0.0026	0.0029
19	0.0005	0.0005	0.0006	0.0007	0.0008	0.0009	0.0010	0.0011	0.0012	0.0014
20	0.0002	0.0002	0.0002	0.0003	0.0003	0.0004	0.0004	0.0005	0.0005	0.0006
21	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002	0.0002	0.0002	0.0003
22	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001

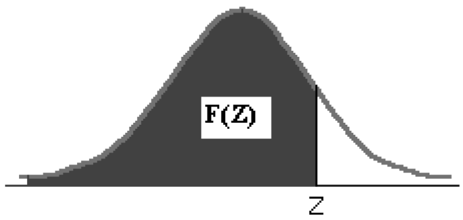
					λt					
x	9.10	9.20	9.30	9.40	9.50	9.60	9.70	9.80	9.90	10.00
0	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0000
1	0.0010	0.0009	0.0009	0.0008	0.0007	0.0007	0.0006	0.0005	0.0005	0.0005
2	0.0046	0.0043	0.0040	0.0037	0.0034	0.0031	0.0029	0.0027	0.0025	0.0023
3	0.0140	0.0131	0.0123	0.0115	0.0107	0.0100	0.0093	0.0087	0.0081	0.0076
4	0.0319	0.0302	0.0285	0.0269	0.0254	0.0240	0.0226	0.0213	0.0201	0.0189
5	0.0581	0.0555	0.0530	0.0506	0.0483	0.0460	0.0439	0.0418	0.0398	0.0378
6	0.0881	0.0851	0.0822	0.0793	0.0764	0.0736	0.0709	0.0682	0.0656	0.0631
7	0.1145	0.1118	0.1091	0.1064	0.1037	0.1010	0.0982	0.0955	0.0928	0.0901
8	0.1302	0.1286	0.1269	0.1251	0.1232	0.1212	0.1191	0.1170	0.1148	0.1126
9	0.1317	0.1315	0.1311	0.1306	0.1300	0.1293	0.1284	0.1274	0.1263	0.1251
10	0.1198	0.1210	0.1219	0.1228	0.1235	0.1241	0.1245	0.1249	0.1250	0.1251
11	0.0991	0.1012	0.1031	0.1049	0.1067	0.1083	0.1098	0.1112	0.1125	0.1137
12	0.0752	0.0776	0.0799	0.0822	0.0844	0.0866	0.0888	0.0908	0.0928	0.0948
13	0.0526	0.0549	0.0572	0.0594	0.0617	0.0640	0.0662	0.0685	0.0707	0.0729
14	0.0342	0.0361	0.0380	0.0399	0.0419	0.0439	0.0459	0.0479	0.0500	0.0521
15	0.0208	0.0221	0.0235	0.0250	0.0265	0.0281	0.0297	0.0313	0.0330	0.0347
16	0.0118	0.0127	0.0137	0.0147	0.0157	0.0168	0.0180	0.0192	0.0204	0.0217
17	0.0063	0.0069	0.0075	0.0081	0.0088	0.0095	0.0103	0.0111	0.0119	0.0128
18	0.0032	0.0035	0.0039	0.0042	0.0046	0.0051	0.0055	0.0060	0.0065	0.0071
19	0.0015	0.0017	0.0019	0.0021	0.0023	0.0026	0.0028	0.0031	0.0034	0.0037
20	0.0007	0.0008	0.0009	0.0010	0.0011	0.0012	0.0014	0.0015	0.0017	0.0019
21	0.0003	0.0003	0.0004	0.0004	0.0005	0.0006	0.0006	0.0007	0.0008	0.0009
22	0.0001	0.0001	0.0002	0.0002	0.0002	0.0002	0.0003	0.0003	0.0004	0.0004
23	0.0000	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002
24	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001

EK 3: Standart Normal Dağılım Tablosu



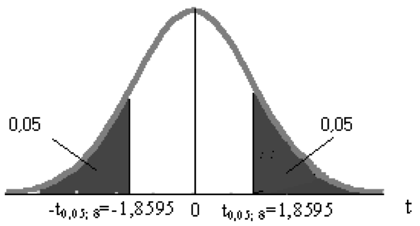
Z	**0**	**0,01**	**0,02**	**0,03**	**0,04**	**0,05**	**0,06**	**0,07**	**0,08**	**0,09**
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	**0,2704**	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4661	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990

EK 4: Standart Normal Dağılımın Birikimli Dağılım Tablosu



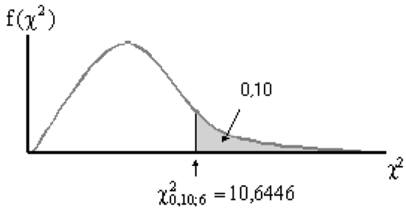
Z	F(Z)	Z	F(Z)	Z	F(Z)	Z	F(Z)	Z	F(Z)	Z	F(Z)	Z	F(Z)
0,00	0,5000												
0,01	0,5040	**0,51**	0,6950	**1,01**	0,8438	**1,51**	0,9345	**2,01**	0,9778	**2,51**	0,9940	**3,01**	0,9987
0,02	0,5040	**0,52**	0,6985	**1,02**	0,8461	**1,52**	0,9357	**2,02**	0,9783	**2,52**	0,9941	**3,02**	0,9987
0,03	0,5120	**0,53**	0,7019	**1,03**	0,8485	**1,53**	0,9370	**2,03**	0,9788	**2,53**	0,9943	**3,03**	0,9988
0,04	0,5160	**0,54**	0,7054	**1,04**	0,8508	**1,54**	0,9382	**2,04**	0,9793	**2,54**	0,9945	**3,04**	0,9988
0,05	0,5199	**0,55**	0,7088	**1,05**	0,8531	**1,55**	0,9394	**2,05**	0,9798	**2,55**	0,9946	**3,05**	0,9989
0,06	0,5239	**0,56**	0,7123	**1,06**	0,8554	**1,56**	0,9406	**2,06**	0,9803	**2,56**	0,9948	**3,06**	0,9989
0,07	0,5279	**0,57**	0,7157	**1,07**	0,8577	**1,57**	0,9418	**2,07**	0,9808	**2,57**	0,9949	**3,07**	0,9989
0,08	0,5319	**0,58**	0,7190	**1,08**	0,8599	**1,58**	0,9429	**2,08**	0,9812	**2,58**	0,9951	**3,08**	0,9990
0,09	0,5359	**0,59**	0,7224	**1,09**	0,8621	**1,59**	0,9441	**2,09**	0,9817	**2,59**	0,9952	**3,09**	0,9990
0,10	0,5398	**0,60**	0,7257	**1,10**	0,8643	**1,60**	0,9542	**2,10**	0,9821	**2,60**	0,9953	**3,10**	0,9990
0,11	0,5438	**0,61**	0,7291	**1,11**	0,8665	**1,61**	0,9463	**2,11**	0,9826	**2,61**	0,9955	**3,11**	0,9991
0,12	0,5478	**0,62**	0,7324	**1,12**	0,8686	**1,62**	0,9474	**2,12**	0,9830	**2,62**	0,9956	**3,12**	0,9991
0,13	0,5517	**0,63**	0,7357	**1,13**	0,8708	**1,63**	0,9484	**2,13**	0,9834	**2,63**	0,9957	**3,13**	0,9991
0,14	0,5557	**0,64**	0,7389	**1,14**	0,8729	**1,64**	0,9495	**2,14**	0,9838	**2,64**	0,9959	**3,14**	0,9992
0,15	0,5596	**0,65**	0,7422	**1,15**	0,8749	**1,65**	0,9505	**2,15**	0,9842	**2,65**	0,9960	**3,15**	0,9992
0,16	0,5636	**0,66**	0,7454	**1,16**	0,8770	**1,66**	0,9515	**2,16**	0,9846	**2,66**	0,9961	**3,16**	0,9992
0,17	0,5675	**0,67**	0,7486	**1,17**	0,8790	**1,67**	0,9525	**2,17**	0,9850	**2,67**	0,9962	**3,17**	0,9992
0,18	0,5714	**0,68**	0,7517	**1,18**	0,8810	**1,68**	0,9535	**2,18**	0,9854	**2,68**	0,9963	**3,18**	0,9993
0,19	0,5753	**0,69**	0,7549	**1,19**	0,8830	**1,69**	0,9545	**2,19**	0,9857	**2,69**	0,9964	**3,19**	0,9993
0,20	0,5793	**0,70**	0,7580	**1,20**	0,8849	**1,70**	0,9554	**2,20**	0,9861	**2,70**	0,9965	**3,20**	0,9993
0,21	0,5832	**0,71**	0,7611	**1,21**	0,8869	**1,71**	0,9564	**2,21**	0,9864	**2,71**	0,9966	**3,21**	0,9993
0,22	0,5871	**0,72**	0,7642	**1,22**	0,8888	**1,72**	0,9573	**2,22**	0,9868	**2,72**	0,9967	**3,22**	0,9994
0,23	0,5910	**0,73**	0,7643	**1,23**	0,8907	**1,73**	0,9582	**2,23**	0,9871	**2,73**	0,9968	**3,23**	0,9994
0,24	0,5948	**0,74**	0,7704	**1,24**	0,8925	**1,74**	0,9591	**2,24**	0,9875	**2,74**	0,9969	**3,24**	0,9994
0,25	0,5987	**0,75**	0,7734	**1,25**	0,8944	**1,75**	0,9599	**2,25**	0,9878	**2,75**	0,9970	**3,25**	0,9994
0,26	0,6026	**0,76**	0,7764	**1,26**	0,8962	**1,76**	0,9608	**2,26**	0,9881	**2,76**	0,9971	**3,26**	0,9994
0,27	0,6064	**0,77**	0,7794	**1,27**	0,8980	**1,77**	0,9616	**2,27**	0,9884	**2,77**	0,9972	**3,27**	0,9995
0,28	0,6103	**0,78**	0,7823	**1,28**	0,8997	**1,78**	0,9625	**2,28**	0,9887	**2,78**	0,9973	**3,28**	0,9995
0,29	0,6141	**0,79**	0,7852	**1,29**	0,9015	**1,79**	0,9633	**2,29**	0,9890	**2,79**	0,9974	**3,29**	0,9995
0,30	0,6179	**0,80**	0,7881	**1,30**	0,9032	**1,80**	0,9641	**2,30**	0,9893	**2,80**	0,9974	**3,30**	0,9995
0,31	0,6217	**0,81**	0,7910	**1,31**	0,9049	**1,81**	0,9649	**2,31**	0,9896	**2,81**	0,9975	**3,31**	0,9995
0,32	0,6255	**0,82**	0,7939	**1,32**	0,9066	**1,82**	0,9656	**2,32**	0,9898	**2,82**	0,9976	**3,32**	0,9996
0,33	0,6293	**0,83**	0,7967	**1,33**	0,9082	**1,83**	0,9664	**2,33**	0,9901	**2,83**	0,9977	**3,33**	0,9996
0,34	0,6331	**0,84**	0,7995	**1,34**	0,9099	**1,84**	0,9671	**2,34**	0,9904	**2,84**	0,9977	**3,34**	0,9996
0,35	0,6368	**0,85**	0,8023	**1,35**	0,9115	**1,85**	0,9678	**2,35**	0,9906	**2,85**	0,9978	**3,35**	0,9996
0,36	0,6406	**0,86**	0,8051	**1,36**	0,9131	**1,86**	0,9686	**2,36**	0,9909	**2,86**	0,9979	**3,36**	0,9996
0,37	0,6443	**0,87**	0,8078	**1,37**	0,9147	**1,87**	0,9693	**2,37**	0,9911	**2,87**	0,9979	**3,37**	0,9996
0,38	0,6480	**0,88**	0,8106	**1,38**	0,9162	**1,88**	0,9699	**2,38**	0,9913	**2,88**	0,9980	**3,38**	0,9996
0,39	0,6517	**0,89**	0,8113	**1,39**	0,9177	**1,89**	0,9706	**2,39**	0,9916	**2,89**	0,9981	**3,39**	0,9997
0,40	0,6554	**0,90**	0,8159	**1,40**	0,9122	**1,90**	0,9713	**2,40**	0,9918	**2,90**	0,9981	**3,40**	0,9997
0,41	0,6591	**0,91**	0,8186	**1,41**	0,9207	**1,91**	0,9719	**2,41**	0,9920	**2,91**	0,9982	**3,41**	0,9997
0,42	0,6628	**0,92**	0,8212	**1,42**	0,9222	**1,92**	0,9726	**2,42**	0,9922	**2,92**	0,9982	**3,42**	0,9997
0,43	0,6664	**0,93**	0,8238	**1,43**	0,9236	**1,93**	0,9732	**2,43**	0,9925	**2,93**	0,9983	**3,43**	0,9997
0,44	0,6700	**0,94**	0,8264	**1,44**	0,9251	**1,94**	0,9738	**2,44**	0,9927	**2,94**	0,9984	**3,44**	0,9997
0,45	0,6736	**0,95**	0,8289	**1,45**	0,9265	**1,95**	0,9744	**2,45**	0,9929	**2,95**	0,9984	**3,45**	0,9997
0,46	0,6772	**0,96**	0,8315	**1,46**	0,9279	**1,96**	0,9750	**2,46**	0,9931	**2,96**	0,9985	**3,46**	0,9997
0,47	0,6803	**0,97**	0,8340	**1,47**	0,9292	**1,97**	0,9756	**2,47**	0,9932	**2,97**	0,9985	**3,47**	0,9997
0,48	0,6844	**0,98**	0,8365	**1,48**	0,9306	**1,98**	0,9761	**2,48**	0,9934	**2,98**	0,9986	**3,48**	0,9997
0,49	0,6879	**0,99**	0,8389	**1,49**	0,9319	**1,99**	0,9767	**2,49**	0,9936	**2,99**	0,9986	**3,49**	0,9998
0,50	0,6915	**1,00**	0,8413	**1,50**	0,9332	**2,00**	0,9772	**2,50**	0,9938	**3,00**	0,9986	**3,50**	0,9998

EK 5: t Değerleri Tablosu



Güven aralığı Tek Yanlı İki Yanlı	0,10 0,45 0,90	0,30 0,35 0,70	0,50 0,25 0,50	0,70 0,15 0,30	0,80 0,10 0,20	0,90 0,05 0,10	0,95 0,025 0,05	0,98 0,01 0,02	0,99 0,005 0,01
sd				**t Değerleri**					
1	0,1584	0,5095	1,0000	1,9626	3,0777	6,3137	12,7062	31,8210	63,6559
2	0,1421	0,4447	0,8165	1,3862	1,8856	2,9200	4,3027	6,9645	9,9250
3	0,1366	0,4242	0,7649	1,2498	1,6377	2,3534	3,1824	4,5407	5,8408
4	0,1338	0,4142	0,7407	1,1896	1,5332	2,1318	2,7765	3,7469	4,6041
5	0,1322	0,4082	0,7267	1,1558	1,4759	2,0150	2,5706	3,3649	4,0321
6	0,1311	0,4043	0,7176	1,1342	1,4398	1,9432	2,4469	3,1427	3,7074
7	0,1303	0,4015	0,7111	1,1192	1,4149	1,8946	2,3646	2,9979	3,4995
8	0,1297	0,3995	0,7064	1,1081	1,3968	**1,8595**	2,3060	2,8965	3,3554
9	0,1293	0,3979	0,7027	1,0997	1,3830	1,8331	2,2622	2,8214	3,2498
10	0,1289	0,3966	0,6998	1,0931	1,3722	1,8125	2,2281	2,7638	3,1693
11	0,1286	0,3956	0,6974	1,0877	1,3634	1,7959	2,2010	2,7181	3,1058
12	0,1283	0,3947	0,6955	1,0832	1,3562	1,7823	2,1788	2,6810	3,0545
13	0,1281	0,3940	0,6938	1,0795	1,3502	1,7709	2,1604	2,6503	3,0123
14	0,1280	0,3933	0,6924	1,0763	1,3450	1,7613	2,1448	2,6245	2,9768
15	0,1278	0,3928	0,6912	1,0735	1,3406	1,7531	2,1315	2,6025	2,9467
16	0,1277	0,3923	0,6901	1,0711	1,3368	1,7459	2,1199	2,5835	2,9208
17	0,1276	0,3919	0,6892	1,0690	1,3334	1,7396	2,1098	2,5669	2,8982
18	0,1274	0,3915	0,6884	1,0672	1,3304	1,7341	2,1009	2,5524	2,8784
19	0,1274	0,3912	0,6876	1,0655	1,3277	1,7291	2,0930	2,5395	2,8609
20	0,1273	0,3909	0,6870	1,0640	1,3253	1,7247	2,0860	2,5280	2,8453
21	0,1272	0,3906	0,6864	1,0627	1,3232	1,7207	2,0796	2,5176	2,8314
22	0,1271	0,3904	0,6858	1,0614	1,3212	1,7171	2,0739	2,5083	2,8188
23	0,1271	0,3902	0,6853	1,0603	1,3195	1,7139	2,0687	2,4999	2,8073
24	0,1270	0,3900	0,6848	1,0593	1,3178	1,7109	2,0639	2,4922	2,7970
25	0,1269	0,3898	0,6844	1,0584	1,3163	1,7081	2,0595	2,4851	2,7874
26	0,1269	0,3896	0,6840	1,0575	1,3150	1,7056	2,0555	2,4786	2,7787
27	0,1268	0,3894	0,6837	1,0567	1,3137	1,7033	2,0518	2,4727	2,7707
28	0,1268	0,3893	0,6834	1,0560	1,3125	1,7011	2,0484	2,4671	2,7633
29	0,1268	0,3892	0,6830	1,0553	1,3114	1,6991	2,0452	2,4620	2,7564
30	0,1267	0,3890	0,6828	1,0547	1,3104	1,6973	2,0423	2,4573	2,7500
40	0,1265	0,3881	0,6807	1,0500	1,3031	1,6839	2,0211	2,4233	2,7045
50	0,1263	0,3875	0,6794	1,0473	1,2987	1,6759	2,0086	2,4033	2,6778
60	0,1262	0,3872	0,6786	1,0455	1,2958	1,6706	2,0003	2,3901	2,6603
70	0,1261	0,3869	0,6780	1,0442	1,2938	1,6669	1,9944	2,3808	2,6479
80	0,1261	0,3867	0,6776	1,0432	1,2922	1,6641	1,9901	2,3739	2,6387
90	0,1260	0,3866	0,6772	1,0424	1,2910	1,6620	1,9867	2,3685	2,6316
100	0,1260	0,3864	0,6770	1,0418	1,2901	1,6602	1,9840	2,3642	2,6259
250	0,1258	0,3858	0,6755	1,0386	1,2849	1,6510	1,9695	2,3414	2,5956
500	0,1257	0,3855	0,6750	1,0375	1,2832	1,6479	1,9647	2,3338	2,5857

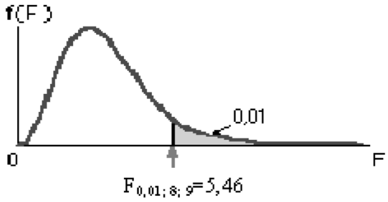
EK 6: Ki-Kare Değerleri



	0,99	**0,95**	**0,90**	**0,80**	**0,70**	**0,60**	**0,50**	**0,40**	**0,30**	**0,20**	**0,10**	**0,05**	**0,02**	**0,01**	**0,005**
sd							**Ki-Kare Değerleri**								
1	0,0002	0,0039	0,0158	0,0640	0,1485	0,2750	0,4549	0,7080	1,0742	1,6424	2,7055	3,8415	5,4119	6,6349	7,8794
2	0,0201	0,1026	0,2107	0,4460	0,7133	1,0220	1,3863	1,8330	2,4079	3,2189	4,6052	5,9915	7,8241	9,2101	10,5965
3	0,1148	0,3518	0,5844	1,0050	1,4237	1,8690	2,3660	2,9460	3,6649	4,6416	6,2514	7,8147	9,8374	11,3449	12,8381
4	0,2971	0,7107	1,0636	1,6490	2,1947	2,7530	3,3567	4,0450	4,8784	5,9886	7,7794	9,4877	11,6678	13,2767	14,8602
5	0,5543	1,1455	1,6103	2,3430	2,9999	3,6560	4,3515	5,1320	6,0644	7,2893	9,2363	11,0705	13,3882	15,0863	16,7496
6	0,8721	1,6354	2,2041	3,0700	3,8276	4,5700	5,3481	6,2110	7,2311	8,5581	**10,6446**	12,5916	15,0332	16,8119	18,5475
7	1,2390	2,1673	2,8331	3,8220	4,6713	5,4930	6,3458	7,2830	8,3834	9,8032	12,0170	14,0671	16,6224	18,4753	20,2777
8	1,6465	2,7326	3,4895	4,5940	5,5274	6,4230	7,3441	8,3510	9,5245	11,0301	13,3616	15,5073	18,1682	20,0902	21,9549
9	2,0879	3,3251	4,1682	5,3800	6,3933	7,3570	8,3428	9,4140	10,6564	12,2421	14,6837	16,9190	19,6790	21,6660	23,5893
10	2,5582	3,9403	4,8652	6,1790	7,2672	8,2950	9,3418	10,4730	11,7807	13,4420	15,9872	18,3070	21,1608	23,2093	25,14881
11	3,0535	4,5748	5,5778	6,9890	8,1479	9,2370	10,3410	11,5300	12,8987	14,6314	17,2750	19,6752	22,6179	24,7250	26,7569
12	3,5706	5,2260	6,3038	7,8070	9,0343	10,1820	11,3403	12,5840	14,0111	15,8120	18,5493	21,0261	24,0539	26,2170	28,2997
13	4,1069	5,8919	7,0415	8,6340	9,9257	11,1290	12,3398	13,6360	15,1187	16,9848	19,8119	22,3620	25,4715	27,6882	29,8193
14	4,6604	6,5706	7,7895	9,4670	10,8215	12,0780	13,3393	14,6850	16,2221	18,1508	21,0641	23,6848	26,8727	29,1412	31,3194
15	5,2294	7,2609	8,5468	10,3070	11,7212	13,0300	14,3389	15,7330	17,3217	19,3107	22,3071	24,9958	28,2595	30,5780	32,8015
16	5,8122	7,9616	9,3122	11,1520	12,6243	13,9830	15,3385	16,7800	18,4179	20,4651	23,5418	26,2962	29,6332	31,9999	34,2671
17	6,4077	8,6718	10,0852	12,0020	13,5307	14,9370	16,3382	17,8240	19,5110	21,6146	24,7690	27,5871	30,9950	33,4087	35,7184
18	7,0149	9,3904	10,8649	12,8570	14,4399	15,8930	17,3379	18,8680	20,6014	22,7595	25,9894	28,8693	32,3462	34,8052	37,1564
19	7,6327	10,1170	11,6509	13,7160	15,3517	16,8500	18,3376	19,9100	21,6891	23,9004	27,2036	30,1435	33,6874	36,1908	38,5821
20	8,2604	10,8508	12,4426	14,5780	16,2659	17,8090	19,3374	20,9510	22,7745	25,0375	28,4120	31,4104	35,0196	37,5663	39,9969
21	8,8972	11,5913	13,2396	15,4450	17,1823	18,7680	20,3372	21,9920	23,8578	26,1711	29,6151	32,6706	36,3434	38,9322	41,4009
22	9,5425	12,3380	14,0415	16,3140	18,1007	19,7290	21,3370	23,0310	24,9390	27,3015	30,8133	33,9245	37,6595	40,2894	42,7957
23	10,1957	13,0905	14,8480	17,1870	19,0211	20,6900	22,3369	24,0690	26,0184	28,4288	32,0069	35,1725	38,9683	41,6383	44,1814
24	10,8563	13,8484	15,6587	18,0620	19,9432	21,6520	23,3367	25,1060	27,0960	29,5533	33,1962	36,4150	40,2703	42,9798	45,5584
25	11,5240	14,6114	16,4734	18,9400	20,8670	22,6160	24,3366	26,1430	28,1719	30,6752	34,3816	37,6525	41,5660	44,3140	46,9280
26	12,1982	15,3792	17,2919	19,8200	21,7924	23,5790	25,3365	27,1790	29,2463	31,7946	35,5632	38,8851	42,8558	45,6416	48,2898
27	12,8785	16,1514	18,1139	20,7030	22,7192	24,5440	26,3363	28,2140	30,3193	32,9117	36,7412	40,1133	44,1399	46,9628	49,6450
28	13,5647	16,9279	18,9392	21,5880	23,6475	25,5090	27,3362	29,2490	31,3909	34,0266	37,9159	41,3372	45,4188	48,2782	50,9936
29	14,2564	17,7084	19,7677	22,4750	24,5770	26,4750	28,3361	30,2830	32,4612	35,1394	39,0875	42,5569	46,6926	49,5878	52,3355
30	14,9535	18,4927	20,2992	23,3640	25,5078	27,4420	29,3360	31,3160	33,5302	36,2502	40,2560	43,7730	47,9618	50,8922	53,6719

EK 7: F Değerleri Tablosu

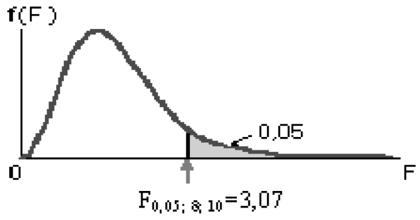
α=0,01 İçin



sd₁	sd₂																		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**12**	**15**	**20**	**24**	**30**	**40**	**60**	**120**	**300**
1	4052	4999	5404	5624	5764	5859	5928	5981	6022	6056	6107	6157	6208	6234	6260	6286	6313	6340	6355
2	98,50	99,00	99,16	99,25	99,30	99,33	99,35	99,37	99,39	99,39	99,41	99,43	99,44	99,45	99,46	99,47	99,48	99,49	99,49
3	34,12	30,82	29,46	28,71	28,23	27,91	27,67	27,48	27,34	27,22	27,05	26,87	26,69	26,59	26,50	26,41	26,31	26,22	26,16
4	21,19	18,00	16,69	15,97	15,52	15,20	14,97	14,79	14,65	14,54	14,37	14,19	14,01	13,92	13,83	13,74	13,65	13,55	13,50
5	16,26	13,27	12,06	11,39	10,96	10,67	10,45	10,28	10,15	10,05	9,88	9,72	9,55	9,46	9,37	9,29	9,20	9,11	9,05
6	13,75	10,93	9,78	9,14	8,74	8,46	8,26	8,10	7,97	7,87	7,71	7,55	7,39	7,31	7,22	7,14	7,05	6,96	6,91
7	12,25	9,55	8,45	7,84	7,46	7,19	6,99	6,84	6,71	6,62	6,46	6,31	6,15	6,07	5,99	5,90	5,82	5,73	5,68
8	11,26	8,65	7,59	7,01	6,63	6,37	6,17	6,02	5,91	5,81	5,66	5,51	5,35	5,27	5,19	5,11	5,03	4,94	4,89
9	10,56	8,02	6,99	6,42	6,05	5,80	5,61	**5,46**	5,35	5,25	5,11	4,96	4,80	4,72	4,64	4,56	4,48	4,39	4,34
10	10,04	7,56	6,55	5,99	5,63	5,38	5,20	5,05	4,94	4,84	4,70	4,55	4,40	4,32	4,24	4,16	4,08	3,99	3,94
11	9,65	7,21	6,21	5,66	5,31	5,06	4,88	4,74	4,63	4,53	4,39	4,25	4,09	4,02	3,94	3,86	3,77	3,69	3,63
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,49	4,38	4,29	4,15	4,01	3,85	3,78	3,70	3,61	3,53	3,44	3,39
13	9,07	6,70	5,73	5,20	4,86	4,62	4,44	4,30	4,19	4,10	3,96	3,81	3,66	3,58	3,50	3,42	3,34	3,25	3,20
14	8,86	6,51	5,56	5,03	4,69	4,45	4,27	4,14	4,03	3,93	3,80	3,65	3,50	3,42	3,34	3,26	3,18	3,09	3,04
15	8,68	6,36	5,41	4,89	4,55	4,31	4,14	4,00	3,89	3,80	3,66	3,52	3,37	3,29	3,21	3,13	3,04	2,95	2,90
16	8,53	6,23	5,29	4,77	4,43	4,20	4,02	3,89	3,78	3,69	3,55	3,40	3,25	3,18	3,10	3,01	2,93	2,84	2,79
17	8,40	6,11	5,18	4,66	4,33	4,10	3,92	3,79	3,68	3,59	3,45	3,31	3,16	3,08	3,00	2,92	2,83	2,74	2,69
18	8,29	6,01	5,09	4,57	4,24	4,01	3,84	3,70	3,59	3,50	3,37	3,22	3,07	2,99	2,91	2,83	2,74	2,65	2,60
19	8,19	5,93	5,01	4,50	4,17	3,93	3,76	3,63	3,52	3,43	3,29	3,15	3,00	2,92	2,84	2,76	2,67	2,58	2,52
20	8,09	5,84	4,93	4,43	4,10	3,87	3,69	3,56	3,45	3,36	3,23	3,08	2,93	2,85	2,77	2,69	2,60	2,51	2,46
21	8,01	5,78	4,87	4,36	4,04	3,81	3,63	3,50	3,39	3,30	3,17	3,03	2,87	2,80	2,72	2,63	2,54	2,45	2,39
22	7,94	5,71	4,81	4,31	3,98	3,75	3,58	3,45	3,34	3,25	3,12	2,97	2,82	2,74	2,66	2,58	2,49	2,40	2,34
23	7,88	5,66	4,76	4,26	3,93	3,71	3,53	3,40	3,29	3,21	3,07	2,93	2,78	2,70	2,62	2,53	2,44	2,35	2,29
24	7,82	5,61	4,71	4,21	3,89	3,66	3,49	3,36	3,25	3,16	3,03	2,88	2,73	2,65	2,57	2,49	2,40	2,30	2,25
25	7,76	5,56	4,67	4,17	3,85	3,62	3,45	3,32	3,21	3,12	2,99	2,85	2,69	2,62	2,53	2,45	2,36	2,26	2,21
26	7,72	5,52	4,63	4,14	3,81	3,59	3,42	3,28	3,18	3,09	2,95	2,81	2,66	2,58	2,50	2,41	2,32	2,23	2,17
27	7,67	5,48	4,60	4,10	3,78	3,55	3,38	3,25	3,14	3,06	2,92	2,78	2,63	2,55	2,46	2,38	2,29	2,19	2,13
28	7,63	5,45	4,56	4,07	3,75	3,52	3,35	3,22	3,11	3,03	2,89	2,75	2,60	2,52	2,43	2,35	2,26	2,16	2,10
29	7,59	5,42	4,53	4,04	3,72	3,49	3,33	3,19	3,09	3,00	2,86	2,72	2,57	2,49	2,41	2,32	2,23	2,13	2,07
30	7,56	5,39	4,51	4,01	3,69	3,47	3,30	3,17	3,06	2,97	2,84	2,70	2,54	2,46	2,38	2,29	2,20	2,14	2,04
40	7,31	5,18	4,31	3,82	3,51	3,29	3,12	2,99	2,88	2,80	2,66	2,52	2,36	2,88	2,20	2,11	2,01	1,91	1,85
60	7,07	4,97	4,12	3,64	3,33	3,11	2,95	2,82	2,71	2,63	2,49	2,35	2,19	2,11	2,02	1,93	1,83	1,72	1,65
120	6,85	4,78	3,94	3,47	3,17	2,95	2,79	2,66	2,55	2,47	2,33	2,19	2,03	1,95	1,86	1,76	1,65	1,53	1,44
300	6,72	4,68	3,84	3,38	3,07	2,86	2,69	2,57	2,46	2,38	2,24	2,09	1,94	1,85	1,76	1,66	1,54	1,41	1,31

EK 7 devam: F Değerleri Tablosu

α=0,05 İçin



sd_1	**sd_2**																		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**12**	**15**	**20**	**24**	**30**	**40**	**60**	**120**	**300**
1	161	200	216	225	230	234	237	230	241	242	244	246	248	249	250	251	252	253	254
2	18,50	19,0	19,2	19,2	19,3	19,3	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,5	19,5	19,5	19,5	19,5	19,5
3	10,10	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,70	8,66	8,64	8,62	8,59	8,57	8,55	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,86	5,80	5,77	5,75	5,72	5,69	5,66	5,64
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,62	4,56	4,53	4,50	4,46	4,43	4,40	4,38
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,94	3,87	3,84	3,81	3,77	3,74	3,70	3,68
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,57	3,51	3,44	3,41	3,38	3,34	3,30	3,27	3,25
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,28	3,22	3,15	3,12	3,08	3,04	3,01	2,97	2,95
9	5,12	4,26	3,86	3,63	3,48	3,37	3,20	3,23	3,18	3,14	3,07	3,01	2,94	2,90	2,86	2,83	2,79	2,75	2,72
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	**3,07**	3,02	2,98	2,91	2,85	2,77	2,74	2,70	2,66	2,62	2,58	2,56
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,79	2,72	2,65	2,61	2,57	2,53	2,49	2,45	2,42
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,69	2,62	2,54	2,51	2,47	2,43	2,38	2,34	2,31
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,60	2,53	2,46	2,42	2,38	2,34	2,30	2,25	2,23
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,53	2,46	2,39	2,35	2,31	2,27	2,22	2,18	2,15
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,48	2,40	2,33	2,29	2,25	2,20	2,16	2,11	2,09
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42	2,35	2,28	2,24	2,19	2,15	2,11	2,06	2,03
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,38	2,31	2,23	2,19	2,15	2,10	2,06	2,01	1,98
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,27	2,19	2,15	2,11	2,06	2,02	1,97	1,94
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,31	2,23	2,16	2,11	2,07	2,03	1,98	1,93	1,90
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,28	2,20	2,12	2,08	2,04	1,99	1,95	1,90	1,87
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25	2,18	2,10	2,05	2,01	1,96	1,92	1,87	1,83
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,23	2,15	2,07	2,03	1,98	1,94	1,89	1,84	1,80
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,20	2,13	2,05	2,01	1,96	1,91	1,86	1,81	1,78
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,18	2,11	2,03	1,98	1,94	1,89	1,84	1,79	1,76
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,16	2,09	2,01	1,96	1,92	1,87	1,82	1,77	1,73
26	4,22	3,37	2,97	2,74	2,58	2,47	2,38	2,32	2,26	2,21	2,14	2,07	1,98	1,94	1,90	1,85	1,80	1,75	1,71
27	4,21	3,35	2,96	2,72	2,57	2,45	2,37	2,30	2,25	2,20	2,13	2,06	1,97	1,93	1,88	1,84	1,79	1,73	1,70
28	4,19	3,34	2,94	2,71	2,55	2,44	2,35	2,29	2,24	2,19	2,11	2,04	1,96	1,91	1,87	1,82	1,77	1,71	1,68
29	4,18	3,32	2,93	2,70	2,54	2,43	2,34	2,28	2,22	2,18	2,10	2,02	1,94	1,90	1,85	1,80	1,75	1,70	1,66
30	4,17	3,31	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,09	2,01	1,93	1,89	1,84	1,79	1,74	1,68	1,65
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,00	1,92	1,84	1,79	1,74	1,69	1,64	1,58	1,54
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,92	1,84	1,75	1,70	1,65	1,59	1,53	1,47	1,42
120	3,92	3,07	2,68	2,45	2,29	2,18	2,09	2,02	1,96	1,91	1,83	1,75	1,66	1,61	1,55	1,50	1,43	1,35	1,30
300	3,87	3,02	2,64	2,40	2,24	2,13	2,04	1,97	1,91	1,86	1,79	1,70	1,60	1,55	1,49	1,43	1,36	1,27	1,21

EK 8: Student Genişliği Değerleri

p=0,95 için

	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	17,97	26,98	32,82	37,08	40,41	43,12	45,40	47,36	49,07	50,59	51,96	53,20	54,33	55,36	56,32	57,22	58,04	58,83	59,56
2	6,08	8,33	9,80	10,88	11,74	12,44	13,03	13,54	13,99	14,39	14,75	15,08	15,38	15,65	15,91	16,14	16,37	16,57	16,77
3	4,50	5,91	6,82	7,50	8,04	8,48	8,85	9,18	9,46	9,72	9,95	10,15	10,35	10,52	10,69	10,84	10,98	11,11	11,24
4	3,93	5,04	5,76	6,29	6,71	7,05	7,35	7,60	7,83	8,03	8,21	8,37	8,52	8,66	8,79	8,91	9,03	9,13	9,23
5	3,64	4,60	5,22	5,67	6,03	6,33	6,58	6,80	6,99	7,17	7,32	7,47	7,60	7,72	7,83	7,93	8,03	8,12	8,21
6	3,46	4,34	4,90	5,30	5,63	5,90	6,12	6,32	6,49	6,65	6,79	6,92	7,03	7,14	7,24	7,34	7,43	7,51	7,59
7	3,34	4,16	4,68	5,06	5,36	5,61	5,82	6,00	6,16	6,30	6,43	6,55	6,66	6,76	6,85	6,94	7,02	7,10	7,17
8	3,26	4,04	4,53	4,89	5,17	5,40	5,60	5,77	5,92	6,05	6,18	6,29	6,39	6,48	6,57	6,65	6,73	6,80	6,87
9	3,20	3,95	4,41	4,76	5,02	5,24	5,43	5,59	5,74	5,87	5,98	6,09	6,19	6,28	6,36	6,44	6,51	6,58	6,64
10	3,15	3,88	4,33	4,65	4,91	5,12	5,30	5,46	5,60	5,72	5,83	5,93	6,03	6,11	6,19	6,27	6,34	6,40	6,47
11	3,11	3,82	4,26	4,57	4,82	5,03	5,20	5,35	5,49	5,61	5,71	5,81	5,90	5,98	6,06	6,13	6,20	6,27	6,33
12	3,08	3,77	4,20	4,51	4,75	4,95	5,12	5,27	5,39	5,51	5,61	5,71	5,80	5,88	5,95	6,02	6,09	6,15	6,21
13	3,06	3,73	4,15	4,45	4,69	4,88	5,05	5,19	5,32	5,43	5,53	5,63	5,71	5,79	5,86	5,93	5,99	6,05	6,11
14	3,03	3,70	4,11	4,41	4,64	4,83	4,99	5,13	5,25	5,36	5,46	5,55	5,64	5,71	5,79	5,85	5,91	5,97	6,03
15	3,01	3,67	4,08	4,37	4,59	4,78	4,94	5,08	5,20	5,31	5,40	5,49	5,57	5,65	5,72	5,78	5,85	5,90	5,96
16	3,00	3,65	4,05	4,33	4,56	4,74	4,90	5,03	5,15	5,26	5,35	5,44	5,52	5,59	5,66	5,73	5,79	5,84	5,90
17	2,98	3,62	4,02	4,30	4,52	4,70	4,86	4,99	5,11	5,21	5,31	5,39	5,47	5,54	5,61	5,67	5,73	5,79	5,84
18	2,97	3,61	4,00	4,28	4,49	4,67	4,82	4,96	5,07	5,17	5,27	5,35	5,43	5,50	5,57	5,63	5,69	5,74	5,79
19	2,96	3,59	3,98	4,25	4,47	4,65	4,79	4,92	5,04	5,14	5,23	5,31	5,39	5,46	5,53	5,59	5,65	5,70	5,75
20	2,95	3,58	3,96	4,23	4,45	4,62	4,77	4,90	5,01	5,11	5,20	5,28	5,36	5,43	5,49	5,55	5,61	5,66	5,71
24	2,92	3,53	3,90	4,17	4,37	4,54	4,68	4,81	4,92	5,01	5,10	5,18	5,25	5,32	5,38	5,44	5,49	5,55	5,59
30	2,89	3,49	3,85	4,10	4,30	4,46	4,60	4,72	4,82	4,92	5,00	5,08	5,15	5,21	5,27	5,33	5,38	5,43	5,47
40	2,86	3,44	3,79	4,04	4,23	4,39	4,52	4,63	4,73	4,82	4,90	4,98	5,04	5,11	5,16	5,22	5,27	5,31	5,36
60	2,83	3,40	3,74	3,98	4,16	4,31	4,44	4,55	4,65	4,73	4,81	4,88	4,94	5,00	5,06	5,11	5,15	5,20	5,24
120	2,80	3,36	3,68	3,92	4,10	4,24	4,36	4,47	4,56	4,64	4,71	4,78	4,84	4,90	4,95	5,00	5,04	5,09	5,13
∞	2,77	3,31	3,63	3,86	4,03	4,17	4,29	4,39	4,47	4,55	4,62	4,68	4,74	4,80	4,85	4,89	4,93	4,97	5,01

EK 8 devam: Student Genişliği Değerleri

p=0,99 için

	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	90,03	135,0	164,3	185,6	202,2	215,8	227,2	237,0	245,6	253,2	260,0	266,2	271,8	277,0	281,8	286,3	290,4	294,3	298,0
2	14,04	19,02	22,29	24,72	26,63	28,20	29,53	30,68	31,69	32,59	33,40	34,13	34,81	35,43	36,00	36,53	37,03	37,50	37,95
3	8,26	10,62	12,17	13,33	14,24	15,00	15,64	16,20	16,69	17,13	17,53	17,89	18,22	18,52	18,81	19,07	19,32	19,55	19,77
4	6,51	8,12	9,17	9,96	10,58	11,10	11,55	11,93	12,27	12,57	12,84	13,09	13,32	13,53	13,73	13,91	14,08	14,24	14,40
5	5,70	6,98	7,80	8,42	8,91	9,32	9,67	9,97	10,24	10,48	10,70	10,89	11,08	11,24	11,40	11,55	11,68	11,81	11,93
6	5,24	6,33	7,03	7,56	7,97	8,32	8,61	8,87	9,10	9,30	9,48	9,65	9,81	9,95	10,08	10,21	10,32	10,43	10,54
7	4,95	5,92	6,54	7,01	7,37	7,68	7,94	8,17	8,37	8,55	8,71	8,86	9,00	9,12	9,24	9,35	9,46	9,55	9,65
8	4,75	5,64	6,20	6,62	6,96	7,24	7,47	7,68	7,86	8,03	8,18	8,31	8,44	8,55	8,66	8,76	8,85	8,94	9,03
9	4,60	5,43	5,96	6,35	6,66	6,91	7,13	7,33	7,49	7,65	7,78	7,91	8,03	8,13	8,23	8,33	8,41	8,49	8,57
10	4,48	5,27	5,77	6,14	6,43	6,67	6,87	7,05	7,21	7,36	7,49	7,60	7,71	7,81	7,91	7,99	8,08	8,15	8,23
11	4,39	5,15	5,62	5,97	6,25	6,48	6,67	6,84	6,99	7,13	7,25	7,36	7,46	7,56	7,65	7,73	7,81	7,88	7,95
12	4,32	5,05	5,50	5,84	6,10	6,32	6,51	6,67	6,81	6,94	7,06	7,17	7,26	7,36	7,44	7,52	7,59	7,66	7,73
13	4,26	4,96	5,40	5,73	5,98	6,19	6,37	6,53	6,67	6,79	6,90	7,01	7,10	7,19	7,27	7,35	7,42	7,48	7,55
14	4,21	4,89	5,32	5,63	5,88	6,08	6,26	6,41	6,54	6,66	6,77	6,87	6,96	7,05	7,13	7,20	7,27	7,33	7,39
15	4,17	4,84	5,25	5,56	5,80	5,99	6,16	6,31	6,44	6,55	6,66	6,76	6,84	6,93	7,00	7,07	7,14	7,20	7,26
16	4,13	4,79	5,19	5,49	5,72	5,92	6,08	6,22	6,35	6,46	6,56	6,66	6,74	6,82	6,90	6,97	7,03	7,09	7,15
17	4,10	4,74	5,14	5,43	5,66	5,85	6,01	6,15	6,27	6,38	6,48	6,57	6,66	6,73	6,81	6,87	6,94	7,00	7,05
18	4,07	4,70	5,09	5,38	5,60	5,79	5,94	6,08	6,20	6,31	6,41	6,50	6,58	6,65	6,73	6,79	6,85	6,91	6,97
19	4,05	4,67	5,05	5,33	5,55	5,73	5,89	6,02	6,14	6,25	6,34	6,43	6,51	6,58	6,65	6,72	6,78	6,84	6,89
20	4,02	4,64	5,02	5,29	5,51	5,69	5,84	5,97	6,09	6,19	6,28	6,37	6,45	6,52	6,59	6,65	6,71	6,77	6,82
24	3,96	4,55	4,91	5,17	5,37	5,54	5,69	5,81	5,92	6,02	6,11	6,19	6,26	6,33	6,39	6,45	6,51	6,56	6,61
30	3,89	4,45	4,80	5,05	5,24	5,40	5,54	5,65	5,76	5,85	5,93	6,01	6,08	6,14	6,20	6,26	6,31	6,36	6,41
40	3,82	4,37	4,70	4,93	5,11	5,26	5,39	5,50	5,60	5,69	5,76	5,83	5,90	5,96	6,02	6,07	6,12	6,16	6,21
60	3,76	4,28	4,59	4,82	4,99	5,13	5,25	5,36	5,45	5,53	5,60	5,67	5,73	5,78	5,84	5,89	5,93	5,97	6,01
120	3,70	4,20	4,50	4,71	4,87	5,01	5,12	5,21	5,30	5,37	5,44	5,50	5,56	5,61	5,66	5,71	5,75	5,79	5,83
∞	3,64	4,12	4,40	4,60	4,76	4,88	4,99	5,08	5,16	5,23	5,29	5,35	5,40	5,45	5,49	5,54	5,57	5,61	5,65