T.C. ANADOLU ÜNİVERSİTESİ YAYINI NO: 2608
AÇIKÖĞRETİM FAKÜLTESİ YAYINI NO: 1576

İSTATİSTİK-I

Yazarlar
Yrd.Doç.Dr. Atilla ASLANARGUN (Ünite 1)
Doç.Dr. Berna YAZICI (Ünite 2)
Doç.Dr. Yeliz MERT KANTAR (Ünite 3, 4)
Prof.Dr. Birdal ŞENOĞLU (Ünite 5)
Yrd.Doç.Dr. İlhan USTA (Ünite 6)

Editörler
Prof.Dr. Embiya AĞAOĞLU
Yrd.Doç.Dr. Mahmut ATLAS

ANADOLU ÜNİVERSİTESİ

İçindekiler

Önsöz

Değerli Öğrenciler,

İktisat ve İşletme Fakültelerinin 2012-2013 öğretim yılı itibariyle kredili sisteme geçmesinden dolayı, önceki yıllarda tek bir kitap olarak okutulan İstatistik dersi, İstatistik I ve İstatistik II olmak üzere ikiye ayrılmıştır. Elinizdeki kitap, ikinci sınıfın birinci döneminde okutulan İstatistik I dersinin konularını içermektedir.

İstatistik, temel olarak, verilerin derlenip toparlanması, betimlenmesi ve bu verilerin kullanılarak çıkarsama yapılması amaçlanan bir bilimdir. Bu nedenle, İstatistik I kitabının konuları, bu amaca hizmet edecek şekilde düzenlenmiştir. Birinci ünitede, İstatistiğin temel kavramları, verilerin serilerle ifade edilmesi ve grafiklerle gösterilmesi konusu ele alınmıştır. İkinci ünite ise, verilere ilişkin çeşitli özet bilgilerin nasıl bulunacağı ve yorumlanacağı işlenmiştir. Bu kapsamda, ikinci ünite, merkezi eğilim ve değişkenlik ölçülerini içermektedir. İstatistiğinin temelini teşkil eden önemli konulardan biri olan Olasılık, iki kısma ayrılmıştır. Olasılığın temel konuları üçüncü ünitede; koşullu olasılık, bağımsız olaylar gibi diğer konular da dördüncü ünitede ele alınmıştır. Beşinci ünitede ise, kesikli rassal değişkenler ve bu rassal değişkenlere ilişkin bazı olasılık dağılımları verilmiştir. Kitabımızın son ünitesinde sürekli rassal değişkenler ele alınmış olup, İstatistikte önemli bir yere sahip olan normal dağılım bu ünitede işlenmiştir.

Değerli Öğrenciler,

Kitabımızı yazarken sade ve anlaşılır bir dil kullanmaya büyük özen gösterdik. Konuların sadece bilgi düzeyinde kalmaması için ünite içinde çokça örnek çözümü vermeye çalıştık. Örneklerin kolay anlaşılır ve konunun pekişmesine yardımcı olacak nitelikte olmalarına özellikle dikkat ettik. Bunun yanı sıra, Açıköğretim sistemine göre tasarlanmış diğer tüm kitaplarımızda olduğu gibi her bir ünitede "Sıra Sizde" ve "Kendimizi Sınayalım" bölümleri ile hem konunun anlaşılıp anlaşılmadığını değerlendirmeniz hem de sınavlarınıza yönelik bir ön çalışma yapmanızı amaçladık.

Kitabımızda hata olmaması için çok büyük bir titizlik gösterdik; fakat gözden kaçan hataların olabileceği de şüphesiz bir gerçektir. Bu konuda, sizlerin bize vereceği dönütler, ilerleyen baskılarda hataların düzeltilmesini sağlayacaktır.

Bu kitabın size ulaşmasında pek çok kişinin emeği geçmiştir. Burada, en büyük teşekkürü, ulusal ve uluslar arası standartta bir kitap yazmak için yoğun bir çaba harcayan yazar arkadaşlarım hak etmektedir. Hepsine, şahsım ve sizlerin adına teşekkür ederim. Başta rektörümüz Prof. Dr. Davut AYDIN olmak üzere, kitabın dizgisinden baskısına kadar görevli olan herkese ayrıca teşekkürlerimi sunarım.

Çalışmalarınızda başarılar dilerim.
Prof.Dr. Embiya AĞAOĞLU
Yrd.Doç.Dr. Mahmut ATLAS
Editörler

Amaçlarımız

Bu üniteyi tamamladıktan sonra;

- İstatistiğin temel kavramlarını tanımlayabilecek,
- İstatistiğin temel kavramlarını kendi içinde ayırt edebilecek,
- İstatistik serilerini açıklayabilecek,
- İstatistik serilerini grafik yardımıyla açıklayabileceksiniz.

Anahtar Kavramlar

- İstatistik Birimi
- Anakütle
- Örneklem
- Değişken (vasıf)
- Ölçme Düzeyi
- Veri Derleme
- Frekans Serisi
- Gruplandırılmış Seri
- Birikimli Seri
- Bileşik Seri
- Çubuk Grafik
- Histogram
- Frekans Poligonu

İçindekiler

İstatistik-I → Temel Kavramlar ve Seriler →

- GİRİŞ
- İSTATİSTİK BİRİMİ
- ANAKÜTLE
- DEĞİŞKEN
- ÖLÇME DÜZEYLERİ (ÖLÇEKLER)
- İSTATİSTİKSEL SERİLER
- GRAFİKLER

Temel Kavramlar ve Seriler

GİRİŞ

İstatistik kelimesi günlük hayatımızda her alanda sık kullanılan bir kavramdır. Bu nedenle, istatistik kavramına farklı farklı anlamlar yüklenir. Örneğin; günlük, haftalık, vb. zaman süreleri için borsa istatistikleri, hava raporlarına ilişkin istatistikler, televizyon kanallarının izlenmesine ilişkin istatistikler, enflasyon oranları, işsizlik oranları vb. Bu durumda, istatistiğe "sayı" anlamı verilmektedir.

Ancak, unutulmamalıdır ki "sayı" anlamında verilen istatistikler, sıradan sayılar değildir. Her biri günlük hayatımızdaki olaylara ilişkin bilimsel amaçlı olarak elde edilen ve kullanılan verilerdir.

İstatistik kavramının gerçek tanımlaması şöyle verilebilir: Araştırmanın amacına yönelik olarak var olan bir problemin belirlenmesi, istatistik birimlerinin tanımlanması ve bunlara ilişkin değişkenlerin belirlenmesi, değişkenlere ilişkin verilerin toplanması, bu verilerin seriler ve tablolar yardımıyla gösterilmesi, serilerin çözümlenmesi ve yorumlanması sürecini içeren bir yöntemler bilimidir.

İstatistik, betimsel istatistik ve çözümsel istatistik olarak ikiye ayrılabilir. Araştırmaya ilişkin verilerin toplanması, seriler yardımıyla gösterilmesi, sınıflandırılması ve grafikler yardımıyla gösterilmesi aşamalarını içeren sürece betimsel istatistik denir. Tablo ve grafikler yardımıyla gösterimi yapılan verilerin amaca uygun farklı istatistik teknikler yardımıyla çözümlenmesi (analizi), sonuçların modeller yardımıyla ifade edilmesi ve ileriye dönük tahminlerde (öngörü) bulunulması sürecine de çözümsel istatistik denir.

İSTATİSTİK BİRİMİ

Sayılabilir veya ölçülebilir özellikleri (değişkenleri) içeren, aralarında bir çok benzerlikler olmakla beraber farklılıklar da bulunan nesnelere veya olaylara *"istatistik birimi"* denir. Eğer, sayılamayan veya ölçülemeyen nesneler veya olaylar söz konusu olduğunda bunlar istatistik birimi oluşturmazlar. Örneğin; koku, renk, korku, sevinç vb.

İstatistik birimi; canlı, cansız, bir olgu, bir olay veya bir kurum olabilir.

Örneğin; İnsan, balık, çocuk, öğretmen vb. canlı istatistik birimleri; araba, ev, okul cansız istatistik birimleri; evlenme, boşanma, kavga vb. olay istatistik birimleri; hastane, okul, fakülte, vb. ise kurum istatistik birimleridir.

Anlaşılacağı üzere; istatistik birimleri sürekli var olabileceği gibi, belli bir anda da ortaya çıkabilir. Eğer, istatistik birimlerine istenilen bir anda ulaşılabiliyorsa bu birime “sürekli istatistik birimi” denir. Buna karşın, istatistik birimi belirli bir anda ortaya çıkıyorsa bu birime de “ani istatistik birimi” adı verilir.

Örneğin; eşya, konut, bina, öğrenci vb. sürekli istatistik birimleri iken; yağmur, kavga, deprem vb. ani istatistik birimleridir.

ANAKÜTLE

Araştırmaya ilişkin tanımlanan istatistik birimlerin tümünün oluşturduğu topluluğa *anakütle* denir. Anakütle istatistik birimlerinden oluştuğuna göre bunlardan farklı bir yapıya sahip olamaz. Örneğin; 2011-2012 öğretim yılında Eskişehir’de ikamet eden kayıtlı AÖF öğrencilerinin aylık harcamalarına ilişkin yapılan bir çalışmada, AÖF öğrencilerinin her biri istatistik birimi iken bu öğrencilerin tümünün oluşturduğu topluluğa anakütle denir. Bu araştırma için, AÖF anakütle olamaz. Çünkü AÖF sadece öğrencilerden oluşan bir topluluk değildir. Öğrencilerle birlikte çalışanlardan ve birçok yapıdan oluşur ve bir tüzel kişiliğe sahiptir.

Çalışmalarda her zaman anakütledeki tüm istatistik birimleri ile çalışmak mümkün ya da anlamlı olmayabilir. Bu nedenle anakütlenin tamamının yerine, bu anakütleden farklı tekniklerle oluşturulan daha az istatistik birimlerinden oluşan alt topluluklarla çalışılır. Bu alt topluluğa da **örneklem** denir.

Anakütlenin alt topluluğuna **örneklem** denir.

Araştırmanın amacına göre; anakütlenin gerek zaman gerekse mekân olarak sınırlandırılması gerekir.

Ele alınan örnekte, anakütleye mekân olarak Eskişehir, zaman olarak da 2011 - 2012 öğretim yılı sınırlandırması yapılmıştır.

Anakütle Türleri

Anakütle içerdiği birimlerin özelliklerine göre türlere ayrılır.

Somut ve Soyut Anakütle

Yapılacak çalışmada istatistik birimlerin tümüne ulaşılabiliyorsa bu birimlerden oluşan topluluğa “somut anakütle” denir. Örneğin “2011 - 2012 öğretim yılında Eskişehir’de ikamet eden AÖF öğrencilerinin oluşturduğu topluluk bir somut anakütledir.

İstatistik birimlerinin tümüne ulaşmanın olası olmadığı durumda oluşan topluluğa da “soyut anakütle” denir. Örneğin; Eskişehir’de bir yıl boyunca yapılan satışlar. Burada tüm birimlere ve bunlara ilişkin satışları elde etmek (ulaşmak) mümkün değildir. Bu nedenle bu kütle de soyut bir anakütledir.

Gerçek ve Varsayımsal Anakütle

Anakütle için diğer bir ayrım da “gerçek ve varsayımsal” olmalarıdır. Gerçekte var olan istatistik birimlerinden oluşan kütleye “gerçek anakütle” denir.

Örneğin; “Eskişehir’de yaşayan işçilerin oluşturduğu anakütle” gerçek bir anakütledir.

Gerçekte var olmadığı veya ortaya çıkmadığı hâlde, var olmaları ya da ortaya çıkmaları olası istatistik birimlerinden oluşan topluluğa da “varsayımsal anakütle” denir.

Örneğin; 10 kişilik bir aday grubu içinden atanacak 2 kişilik gruplar için 45 farklı seçim yapılabilir. Burada her farklı grup bir istatistik birimi ve bu istatistik birimlerinin oluşturduğu topluluk da varsayımsal anakütledir.

$$\binom{10}{2} = \frac{10!}{2! \cdot (10-2)!} = \frac{10 \cdot 9 \cdot 8!}{2 \cdot 8!} = 45$$

Hazır ve Hareketli Anakütle

İstatistik birimlerin sürekli ve ani birimler olmasına göre de anakütleler sınıflandırılabilir.

Sürekli istatistik birimlerinden oluşan topluluğa "hazır anakütle"; ani istatistik birimlerinden oluşan topluluğa da "hareketli anakütle" denir.

Örneğin; öğrenci, ev, tarla vb. sürekli istatistik birimleri olduğundan bunların oluşturduğu topluluklar hazır anakütledir. Doğum, ölüm, evlenme vb. ani istatistik birimleridir. Bu nedenle, bunların oluşturduğu topluluklar da hareketli anakütleyi oluşturur.

DEĞİŞKEN

İstatistik birimlerin sahip oldukları ve farklı değerler alabilen, diğer istatistik birimlerinden ayırt edilmesini sağlayan özelliklere *değişken* denir. Değişkenin aldığı değerlere de gözlem veya ölçüm değerleri denir.

Örneğin, öğrencilerin boyu ve göz renklerine ilişkin bir araştırma yapıldığında; öğrenci boyu ve göz rengi değişkenlerdir. Anakütledeki herhangi bir öğrencinin boyu 178 cm ve göz renginin kahverengi olarak ölçüldüğünde veya belirlendiğinde, 178 cm ve kahverengi, gözlem (ölçüm) değerleridir.

Değişkenler ölçülebilir, sayılabilir ve farklı özelliklerine göre türlere ayrılır.

Değişken Türleri

Değişkenlerin aldığı değerler sayısal olarak ifade edilebiliyorsa bu değişkene "sayısal değişken" (nicel değişken); sözel olarak ifade edilebiliyorsa "sözel değişken" (nitel değişken) denir.

Örneğin; öğrencilerin kilosu, boyu, dersten aldığı not, haftalık harçlığı vb. gibi ölçülebilen veya öğrenci sayısı, konut sayısı, çocuk sayısı vb. gibi sayılabilen değişkenler sayısal değişkenlerdir. Bunun yanında; göz rengi, cinsiyet, medeni durum vb. değişkenler sözel olarak ifade edilip sınıflara ayrıldığından bu değişkenler de sözel değişkenlerdir.

Sayısal değişkenler ölçülebiliyorsa sürekli değişken, sayılabiliyorsa kesikli değişken olarak adlandırılır.

Örneğin; öğrenci ağırlığı "ölçülebilir" ve sayısal bir değişken olduğuna göre sürekli değişken, konut sayısı da "sayılabilir" ve sayısal bir değişken olduğuna göre de kesikli değişkendir.

Değişkenler için bir diğer önemli sınıflama da "zaman değişkeni", "mekan değişkeni" ve "maddi değişken" sınıflamasıdır. Eğer bir değişkenin aldığı değer, zamana göre belirleniyorsa bu değişkene "zaman değişkeni"; mekana göre belirleniyorsa bu değişkene de "mekan değişkeni" denir. Bunların dışında kalan tüm değişkenlere de maddesel değişkenler denir.

Örneğin; aylara göre yağış miktarı, günlere göre borsa kapanış endeksi, haftalara göre ortalama sıcaklık, saatinde inen ve kalkan uçak sayısı zamana bağlı olduğu için zaman değişkenidir.

İllere göre seçmen sayısı, bölgelere göre işsizlik oranları, ülkelere göre kişi başına düşen milli gelir mekan değişkenlerine örnek oluşturur.

Öğrenci boyu, kilosu, günlük harcaması, not ortalaması vb. maddesel değişkene örnektir.

SIRA SİZDE 1

Öğrencinin ağırlığı, göz rengi, ailesindeki kişi sayısı, medeni durumu hangi tür değişkenlere örnektir?

ÖLÇME DÜZEYLERİ (ÖLÇEKLER)

Anakütle veya örneklemdeki istatistik birimlerin, ilgilenilen sayısal veya sözel özelliklerinin aldığı değerlerin, sayılar veya simgelerle gösterimine ölçme denir. Ölçme sonucu değişkenin aldığı değere de ölçüm denir.

Değişkenlere ilişkin ölçümler, değişkenin yapısına göre 4 farklı ölçme düzeyinde yapılır.

Sınıflayıcı Ölçme Düzeyi

Sınıflayıcı ölçek, ölçme düzeyi en düşük olan ve birimi olmayan ölçektir.

İstatistik birimlerin ilgilenilen değişkeninin aldığı değerler, sayı ve sembollerle gösterilir. Bu sayı ve semboller, birimlerin hangi sınıfta bulunduğunu ifade eder. Bu sayı ve simgeler arasında bir büyüklük veya küçüklük söz konusu olmadığından matematiksel işlemler yapılamaz. Bu sayı ve simgeler sadece, her bir birimin hangi sınıfa ait olduğunu gösterir. Ölçme düzeyleri içinde en kaba ve ölçme düzeyi en düşük olanıdır. Bir ölçü birimi yoktur.

Genellikle sosyal bilimlerde sözel değişkenlerle ilgili çalışmalarda kullanılır. Örneğin meslek, medeni durum, kullanılan kredi kartı tipi vb. değişkenler sınıflayıcı ölçme düzeyinde ölçülebilir.

Örnek 1. Anadolu Üniversitesi Fen Fakültesi kantininde bulunan 100 öğrencinin hangi bölümde okudukları Tablo 1.1'de verilmiştir.

Tablo 1.1 *Öğrencilerin Bölümlere Göre Dağılımı*

Bölümler	Öğrenci Sayısı
Matematik	26
İstatistik	18
Fizik	32
Kimya	15
Biyoloji	9
Toplam	**100**

Tablo 1.1'de, 100 öğrencinin Fen Fakültesinde hangi bölümde okuduklarına ilişkin bilgiler verilmiştir. Değişken, "bölüm" olduğu için sözel değişkendir. Bu değişken için sınıflayıcı ölçek kullanılır. Çünkü, bölümler arası bir sıralama veya üstünlük yoktur. Ayrıca, değişken birimi de yoktur.

Sıralayıcı Ölçme Düzeyi

Sıralayıcı ölçek sınıflamanın yanında sayı ve simgelerde, büyüklük ve küçüklük kavramının olduğu ölçektir.

Bu ölçme düzeyi, sınıflayıcı ölçme düzeyine ek olarak, değişkenin aldığı sayı ve simgelerin farklılığının yanında, bu sayı ve simgelere büyüklük ve küçüklük kavramlarına anlam kazandıran ölçme düzeyidir. İstatistik birimlerine sınıflama yanında önem veya değer sıralaması verir. Örneğin; öğrenim durumu, akademik ünvan, rütbe vb. değişkenler için sıralayıcı ölçme düzeyi kullanılır.

Örnek 2. A şirketinde çalışan 160 kişinin öğrenim durumları Tablo 1.2'de verilmiştir.

Tablo 1.2 *Çalışanların Öğrenim Durumlarına İlişkin Veriler*

Öğrenim Durumu	Çalışan Sayısı
Lise	5
Yüksek Okul	52
Üniversite	85
Yüksek Lisans	15
Doktora	3

Tablo 1.2'ye göre, çalışanlar öğrenim durumuna göre sıralanmıştır. Sıralamanın yanında sınıflama da olduğuna göre, öğrenim durumu için uygun olan ölçek, sıralayıcı ölçek olur.

Aralıklı Ölçme Düzeyi

Aralıklı ölçek, değişkenin aldığı sayısal değerlerin birimle ifade edildiği ve sayılar arasındaki farkın anlam kazandığı ölçektir.

Bu ölçme düzeyinde değişkenin aldığı sayısal değerler birimle ifade edilir ve sayılar arasındaki farklar anlamlıdır. Ancak bu ölçme düzeyi için kesin bir sıfır başlangıç noktası yoktur. Örneğin; sıcaklık ölçümleri, takvimler, zeka derecesi vb.

Örneğin, sıcaklık değişkeni için birim, santigrat derece (°C) olduğundaki sıfır başlangıç noktası (0°C) ile fahrenheit (°F) olduğundaki sıfır başlangıç noktası (0°F) farklı sıcaklık ölçümlerini verirler. Çünkü 0°C'ye karşı gelen değer 32°F dır.

Oransal Ölçme Düzeyi

Oransal ölçek, eşit aralıklı ölçeğe ek olarak gerçek bir sıfır noktasının olduğu ölçektir.

Aralıklı ölçme düzeyinde yapılan ölçüme ek olarak gerçek bir sıfır noktasının olduğu ölçme düzeyidir. Örneğin; gelir, harcama, ağırlık, hız vb. değişkenler bu ölçme düzeyinde ölçülür.

Örnek 3. Bir kafeteryada rastgele seçilen 45 kişinin günlük ortalama gelirleri Tablo 1.3'te verilmiştir.

Tablo 1.3 *Kafeteryada Bulunan 45 kişinin Ortalama Gelirleri*

Günlük Ortalama Gelir (₺)	Kişi Sayısı
1 - 21'den az	5
21 - 41'den az	14
41 - 61'den az	22
61 - 81'den az	3
81 - 101'den az	1

Günlük ortalama gelir değişkeni için sınıflama, sıralama ve aralıklı ölçme düzeyine ek olarak gerçek bir sıfır noktasına (₺0) sahip olduğundan oransal ölçme düzeyinde değerlendirilir.

Sınıflayıcı ve sıralayıcı ölçme düzeyine göre ölçülen değişkenler, sözel değişkenlerdir. Bu değişkenler için parametrik olmayan istatistik teknikler kullanılır.

Aralıklı ve oransal ölçme düzeyine göre ölçülen değişkenler sayısal değişkenlerdir. Bu değişkenler için de parametrik istatistik teknikler uygulanır.

Bu nedenle çalışmanın amacına göre belirlenen değişkenlerin ölçme düzeylerine göre istatistik teknikler kullanılır.

SIRA SİZDE 2

Cinsiyet, ders kodları, mezuniyet derecesi, sıcaklık, boy uzunluğu ve aylık gelir değişkenleri için hangi tip ölçek türü kullanılır?

İSTATİKSEL SERİLER

Liste

İlgilenilen değişkenin almış olduğu değerler, diğer bir değişkene göre veya rastgele sıralanmış ise bu tabloya “liste” denir.

Tablo 1.4’te görüldüğü gibi “2011 - 2012 öğretim yılı gece danışmanlık eğitimi alan 120 AÖF öğrencisinin, aylık harcamaları öğrenci numarasına göre verilmiştir. İlgilenilen değişken olan aylık harcamalar için bir sıralama yoktur ve bu tablo bir listedir. Listeye bakıldığında, ilgilenilen değişkenle ilgili bir sayısal çıkarsama yapmak zordur. Bu nedenle, öncelikle yapılması gereken, bu serinin bir istatistik serisi haline getirilmesidir. Eğer, aylık harcamalar küçükten büyüğe (veya büyükten küçüğe) sıralanarak düzenlenirse bu seri daha iyi anlaşılabilir duruma gelir.

Tablo 1.4
120 Öğrencinin Aylık Harcamaları (Liste)

Öğrenci No	Aylık Harcama (₺)	Öğrenci No	Aylık Harcama (₺)	Öğrenci No	Aylık Harcama (₺)	Öğrenci No	Aylık Harcama (₺)
1	250	31	650	61	820	91	535
2	200	32	800	62	650	92	750
3	500	33	320	63	750	93	820
4	650	34	800	64	500	94	500
5	600	35	820	65	550	95	600
6	650	36	420	66	875	96	875
7	300	37	820	67	650	97	980
8	650	38	800	68	900	98	400
9	420	39	700	69	535	99	1000
10	600	40	820	70	950	100	750
11	600	41	840	71	875	101	980
12	700	42	800	72	900	102	750
13	650	43	600	73	750	103	875
14	500	44	500	74	400	104	420
15	700	45	650	75	950	105	535
16	650	46	380	76	900	106	750
17	700	47	875	77	700	107	950
18	320	48	820	78	875	108	420
19	650	49	750	79	800	109	600
20	750	50	800	80	750	110	750
21	800	51	550	81	500	111	900
22	600	52	750	82	400	112	875
23	420	53	380	83	900	113	535
24	600	54	550	84	820	114	750

25	700	55	800	85	600	115	1050
26	800	56	700	86	875	116	500
27	750	57	650	87	900	117	600
28	800	58	820	88	750	118	875
29	380	59	650	89	875	119	750
30	700	60	400	90	400	120	550

Tablo 1.4
Devamı

Basit Seri

İlgilenilen değişkenin almış olduğu değerlerin küçükten büyüğe veya büyükten küçüğe sıralanmasıyla oluşan seriye "basit seri" denir.

Tablo 1.4'te verilen aylık harcamalar, öğrenci numarasına göre sıralanmış iken, küçükten büyüğe doğru sıralanarak oluşan basit seri Tablo 1.5'te görüldüğü gibidir. Tablo 1.5'e bakıldığında en az ve en fazla harcamalar görülebilir.

Ancak, gözlem sayısının çok fazla olduğu basit seri ile hangi değer çevresinde yığılma olduğunu görmek zordur. Bu nedenle, basit serinin frekans veya gruplandırılmış seriye dönüştürülmesi daha yararlı olabilir.

Aylık Harcamalar (₺)							
200	420	550	650	700	750	820	875
250	420	550	650	700	750	820	875
300	420	550	650	700	750	820	900
320	500	600	650	750	800	820	900
320	500	600	650	750	800	820	900
380	500	600	650	750	800	820	900
380	500	600	650	750	800	840	900
380	500	600	650	750	800	875	900
400	500	600	650	750	800	875	950
400	500	600	650	750	800	875	950
400	535	600	700	750	800	875	950
400	535	600	700	750	800	875	980
400	535	600	700	750	800	875	980
420	535	650	700	750	820	875	1000
420	550	650	700	750	820	875	1050

Tablo 1.5
120 Öğrencinin Aylık Harcamaları (Basit Seri - Küçükten Büyüğe Sıralanmış)

Frekans Serisi

İlgilenilen değişkenin almış olduğu farklı değerlerin küçükten büyüğe sıralanması ve bu değerlerin karşısına kaç kez tekrar ettiğinin (frekansı) yazılmasıyla oluşturulan istatistik serisine "frekans serisi" denir.

Tablo 1.5'teki basit serinin frekans serisine dönüştürülmüş biçimi Tablo 1.6'da görüldüğü gibidir.

Tablo 1.6
120 Öğrencinin Aylık Harcamaları (Frekans Serisi)

Aylık Harcamalar (₺)	**Frekanslar (f_i)**
200	1
250	1
300	1
320	2
380	3
400	5
420	5
500	7
535	4
550	4
600	10
650	12
700	8
750	15
800	10
820	8
840	1
875	10
900	6
950	3
980	2
1000	1
1050	1
Toplam Frekans	**120**

Tablo 1.6'ya bakıldığında farklı gözlem değerlerinin kaç kez ortaya çıktığı ve bunun yanında, en fazla tekrar eden gözlem değerleri kolaylıkla görülebilir.

Ancak, gözlem sayısı ve/veya farklı gözlem değerlerinin sayısı çok fazla olduğunda, frekans serisini kavramak da zorlaşabilir. Bu durumda frekans serisi sınıflara (aralıklara) ayrılarak "gruplandırılmış seri" haline dönüştürülür.

Gruplandırılmış Seri

İlgilenilen değişken değerlerinin, belirlenen sınıflara (aralıklara) ayrılması ve bu sınıflara giren gözlem sayısının ayrı bir sütuna yerleştirilmesiyle oluşan seriye "gruplandırılmış seri" denir.

Frekans serisini, gruplandırılmış seri haline dönüştürürken sınıf aralığı büyüklüğü Sturges Kuralı ya da araştırmacının görüşüne bağlı olarak belirlenir. Sınıf aralığının keyfî olarak belirlenmesinde göz önüne alınması gereken kriter birbirine yakın gözlem değerlerinin bir araya getirilmesidir. Genellikle, en az 5 en fazla 20 sınıf sayısı önerilir.

Sturges kuralı yardımıyla sınıf sayısı ve buna bağlı olarak da sınıf aralığı büyüklüğü belirlenir. Gruplandırılmış seride her sınıfın en küçük değerine alt sınır değeri, en büyük değerine de üst sınır değeri denir. Her bir sınıftaki üst sınır ile alt sınır arasındaki farka da sınıf aralığı büyüklüğü (genişliği) denir.

x_{enb} = *En büyük gözlem değeri*

x_{enk} = *En küçük gözlem değeri*

k = *Sınıf Sayısı*

h = *Sınıf Aralığı Büyüklüğü*

$$k = 1 + (3.322) \cdot log(N)$$
$$h = (x_{enb} - x_{enk}) / k$$

Tablo 1.6'da verilen frekans serisi gruplandırılmış seriye dönüştürülmek istendiğinde;

$x_{enb} = 1050$, $x_{enk} = 200$, $N = 120$

$$k = 1 + (3.322) \cdot log(120)$$
$$k = 1 + (3.322) \cdot (2.079)$$
$$k = 1 + 6.906 = 7.906$$
$$k \approx 8$$

Bu örnek için sınıf aralığı büyüklüğü (h);

$$h = \frac{1050 - 200}{8} = 106.25$$
$$h \approx 107$$

Aylık Harcama (₺)	Frekanslar (f_i)
200 - 307'den az	3
307 - 414'den az	10
414 - 521'den az	12
521 - 628'den az	18
628 - 735'den az	20
735 - 842'den az	34
842 - 949'dan az	16
949 - 1056'dan az	7
Toplam	**120**

Tablo 1.7
120 Öğrenciye İlişkin Aylık Harcamalar (Gruplandırılmış Seri)

Gruplandırılmış seri oluşturulurken dikkat edilmesi gereken bir diğer konu da değişkenin sürekli veya kesikli olması durumunda sınıfların belirlenmesidir.

Eğer değişken sürekli ise sınıflar belirlenirken her sınıfın üst sınır değeri sınıfa dahil edilmez, fakat o izleyen sınıfın alt sınır değeri olur. Tablo 1.7'de gruplandırılmış seriye bakıldığında görüldüğü gibi sınıf aralığı büyüklüğü 107 olarak belirlenmiştir. Buna göre ilk sınıfın alt sınır değeri 200 ve üst sınır değeri 200+107=307'den az olur. Tablo 1.7'de frekans serisine bakıldığında 200 ile 307'den az 3 gözlem değeri bulunmaktadır. Benzer şekilde diğer sınıflar ve frekanslar da belirlenir.

Kesikli değişken söz konusu olduğunda ise Tablo 1.8'de görüldüğü gibi bir sınıfın üst sınır değeri o sınıfa aittir.

Tablo 1.8
Kesikli Değişkene İlişkin Frekans Serisi

Sınıflar	Frekanslar (f_i)
0 - 20	5
21 - 40	15
41 - 60	25
61 - 80	12
81 - 100	3
Toplam	**60**

Gruplandırılmış seride değişkene ilişkin birçok bilgiye görsel olarak da ulaşılabilir. Değişkenin aldığı en küçük ve en büyük değer, değerlerin nerede yoğunlaştığı, serinin dağılımı kabaca görülebilir.

Bunun yanında frekans serisinin gruplandırılmış seriye dönüştürülmesinde bilgi kaybı olmaktadır. Çünkü frekans serisinde gerçek gözlem değerleriyle çalışılırken gruplandırılmış seride sınıf aralığında bulunan gözlem değerleri için kesin değerler yoktur. Bu nedenle izleyen bölümde aritmetik ortalama, mod, medyan vb. istatistikler hesaplanırken yapılacak varsayımlarla gözlem değerleri için yaklaşık değerlerle çalışılır.

SIRA SİZDE 3

Bir mağazaya gelen 365 günlük müşteri sayısı dağılımı izleyen tabloda verilmiştir.

Müşteri Sayısı	Gün Sayısı (f_i)
0 - 4	170
5 - 9	140
10 - 14	30
15 - 19	15
20 - 24	10
Toplam	**365**

Buna göre;
a. (10 - 14) müşterinin geldiği gün sayısı (frekansı) kaçtır?
b. Gruplandırılmış serinin sınıf aralığı büyüklüğü kaçtır?
c. 15 ve 15'ten fazla müşterinin geldiği gün sayısı kaçtır?

Birikimli Seriler

Gruplandırılmış seride, sınıf aralıklarındaki frekanslar birbirine eklenerek oluşturulan seriye **"-den az"** serisi denir.

Gruplandırılmış seride, sınıf aralıklarındaki frekansların toplam frekanstan eksiltilmesiyle oluşturulan seriye **"-den çok"** serisi denir.

Çalışmalarda, bazen istatistik serilerinde belli bir "değerden az" veya "değerden çok" gözlem sayılarına kolayca ulaşılmak istenebilir. Bu durumda, "birikimli seriler" söz konusudur. Gruplandırılmış seride her sınıfın frekansına, izleyen sınıfların frekansları eklenerek (veya toplam frekanstan eksiltilerek) oluşturulan seriye birikimli seri; birikimli seride her sınıfa karşı gelen frekansa da birikimli frekans denir.

Birikimli seriler **"-den az"** ve **"-den çok"** olmak üzere iki şekilde oluşturulur. Eğer, seride frekanslar eklenerek küçükten büyüğe doğru oluşturuluyorsa "-den az" serisi; büyükten küçüğe doğru eksiltilerek oluşturuluyorsa "-den çok" serisi söz konusudur.

Örnek 4: Aylık harcamalara ilişkin gruplandırılmış seri Tablo 1.7 için “-den az” ve “ -den çok” serileri Tablo 1.9’da görüldüğü gibidir.

Tablo 1.9
Gruplandırılmış Seri İçin Birikimli Seriler

Aylık Harcama (₺)	Frekanslar (f_i)	$\sum (f_i)$ (-den az)	$\sum (f_i)$ (-den çok)
200 - 307’den az	3	3	120
307 - 414’den az	10	10+3=13	120-3=117
414 - 521’den az	12	12+13=25	117-10=107
521 - 628’den az	18	18+25=43	107-12=95
628 - 735’den az	20	20+43=63	95-18=77
735 - 842’den az	34	34+63=97	77-20=57
842 - 949’dan az	16	16+97=113	57-34=23
949 - 1056’dan az	7	7+113=120	23-16=7
Toplam	**120**		

Yorum: “-den az” serisi için; aylık harcaması ₺307’den az 3 öğrenci, ₺414’den az 13 öğrenci vb. şekilde ₺1056’den az 120 öğrenci vardır.

“-den çok” serisi için; aylık harcaması 200 ve ₺200’den çok 120 öğrenci, 307 ve ₺307’den çok 117 öğrenci vb. şekilde 949 ve ₺949’den çok 7 öğrenci vardır.

Genelde gruplandırılmış seriler için oluşturulan birikimli seriler, nadir olarak frekans serileri için de kullanılır.

Örnek 5: 40 öğrencinin günlük harcamalarına ilişkin frekans serisi izleyen tabloda verilmiştir. Buna göre, “-den az” ve “-den çok” serileri oluşturulmuş ve yorumlanmıştır.

X_i	(f_i)	$\sum (f_i)$ (-den az)	$\sum (f_i)$ (-den çok)
5	2	2	40
10	3	3+2=5	40-2=38
18	5	5+5=10	38-3=5
20	10	10+10=20	35-5=30
23	12	12+20=32	30-10=20
25	8	8+32=40	20-12=8
Toplam	**40**		

Yorum: “-den az” serisi için günlük harcaması 5 ve ₺5’den az olan 2 öğrenci, 10 ve ₺10’den az 5 öğrenci vb. şekilde 25 ve ₺25’den az 40 öğrenci vardır.

“-den çok” serisi için; günlük harcaması 5 ve ₺5’den çok 40 öğrenci, 10 ve ₺10’den çok 38 öğrenci vb. şekilde 25 ve ₺25’den çok 8 öğrenci vardır.

SIRA SİZDE 4

Aşağıdaki tabloda -den az serisi veriliyor. Buna göre 40 ve 40’tan çok kaç tane gözlem değeri vardır?

Sınıflar	$\sum (f_i)$ (-den az)
10 - 20’den az	5
20 - 30’dan az	12
30 - 40’dan az	20
40 - 50’den az	24
50 - 60’dan az	27
60 - 70’den az	30

SIRA SİZDE 5

30 öğrencinin istatistik dersinden aldığı başarı notları izleyen tabloda verilmiştir. Buna göre, “-den az” ve -den çok serilerini oluşturun.

Başarı Notları	Frekanslar (f_i)
AA	2
BB	7
CC	11
DD	3
FF	7
Toplam	30

Bileşik Seriler

İstatistik birimlerin, iki veya daha fazla değişkene göre aldığı değerleri birlikte gösteren serilere “bileşik seri” denir. Değişken sayısı birden fazla olduğu için bu durumda değişkenlerden herhangi birine göre gözlem değerleri küçükten büyüğe doğru sıralanır ve diğer değişken değerleri de yeni sıralamaya göre düzenlenir. Bileşik seriler de basit, frekans ve gruplandırılmış seri olarak gösterilebilir.

Basit bileşik seride, istatistik birimlerin iki değişken için aldığı değerler, değişkenlerden birine göre sıralanır ve ikinci sütuna da bu sıralamaya göre diğer değişkenin aldığı değerler yazılır.

Tablo 1.10’da 8 öğrencinin ağırlığı ve boy uzunluğuna ilişkin basit bileşik seri örneği verilmiştir. Bu tabloda sıralama, öğrenci ağırlığı değişkenine göre yapılmıştır.

Tablo 1.10
Öğrencilerin Boy Uzunluğu ve Kilolarına İlişkin Bileşik Seri

Öğrencilerin Ağırlığı (kg)	Öğrencilerin Boy Uzunluğu (cm)
52	157
55	152
62	161
68	178
71	165
76	170
85	187
97	193

Bileşik frekans serisinde, istatistik birimlerin iki değişken için aldığı farklı değerlere göre bir tablo oluşturulur. Tabloyu oluşturmak için değişkenlerden birinin aldığı farklı değerler yatay sütuna, diğeri de düşey sütuna küçükten büyüğe doğru sıralanır. Daha sonra oluşan tabloda gözelere karşı gelen gözlem sayıları yazılır.

Örneğin, AÜ AÖF öğrencileri ile yapılan bir çalışmada 970 öğrenci için aylık gelirleri ve barınma giderleri Tablo 1.11'de bileşik frekans serisi şeklinde verilmiştir.

Tablo 1.11
Aylık Gelir ve Barınma Giderlerine İlişkin Bileşik Frekans Serisi

	Aylık Gelir (₺)								
Barınma Gideri (₺)	300	400	500	600	700	800	900	1000	Toplam
75	50	60	-	-	-	-	-	-	110
125	15	100	150	5	-	-	-	-	270
200	3	80	70	15	-	-	-	-	168
250	-	32	90	70	5	-	-	-	197
350	-	-	15	65	40	2	3	5	130
400	-	-	-	10	40	3	2	10	65
500	-	-	-	6	5	2	-	7	20
600	-	-	-	-	3	5	-	2	10
Toplam	68	272	325	171	93	12	5	24	**970**

Tablo 1.11'e bakıldığında; örneğin aylık geliri ₺700 ve barınma gideri ₺250 olan 5 öğrenci ya da barınma gideri ₺400 ve aylık geliri ₺1000 olan 10 öğrenci bulunmaktadır.

Bileşik gruplandırılmış seride ise istatistik biriminin iki değişkene göre almış olduğu değerler yerine farklı sınıf aralıkları kullanılmaktadır. Buna göre değişkenlerden biri için yatay sütuna, diğeri için de düşey sütuna farklı sınıf aralıkları küçükten büyüğe doğru sıralanır. Oluşan tablodaki gözelere ilgili gözlem sayıları yazılır.

Örneğin Tablo 1.11'deki bileşik frekans serisi ele alındığında, izleyen Tablo 1.12'deki bileşik gruplandırılmış seri oluşturulabilir.

Tablo 1.12
Aylık Gelir ve Barınma Giderlerine İlişkin Gruplandırılmış Seri

	Aylık Gelir (₺)			
Barınma Gideri (₺)	**300 - 600'den az**	**600 - 900'den az**	**900 - 1200'den az**	**Toplam**
75 - 250'den az	528	20	-	548
250 - 500'den az	137	235	20	392
500 - 750'den az	-	21	9	30
Toplam	665	276	29	970

Tablo 1.12'ye bakıldığında aylık geliri (₺600-900) arasında olup barınma gideri (₺250-500) arasında olan 235 öğrenci bulunmaktadır.

GRAFİKLER

Günlük hayatta olsun, bilimsel çalışmalarda olsun, ilgilenilen olaya ilişkin sayılar veya serilerin grafikle gösterimi daha yaygın olarak kullanılmaktadır. Çünkü, sayılara veya serilere bakıldığında görülemeyen birçok artış, azalış, eğilim gibi davranışlar grafik gösterimle kolayca görülebilir. Bu nedenle, istatistiksel serilerin grafik gösterimleri, tek tek ele alınacak ve nasıl çizileceği üzerinde sırasıyla durulacaktır.

Dairesel (Pasta) Grafikler

Dairesel (Pasta) grafik, sözel değişkenlerin oransal frekanslarını göstermek için kullanılan grafiktir.

Araştırmalarda sınıflayıcı ya da sıralayıcı ölçeklendirilmiş sözel değişkenler için kullanılan bir grafiktir. Değişkenin aldığı farklı değerlerin frekansları, parçaları; frekanslar toplamı, dağılımın bütününü göstermek üzere daire şeklinde çizilir. Oluşan bu grafiğe, **dairesel (pasta)** grafik denir.

Bu grafiği çizmek için değişkenin aldığı her farklı gözlem değerinin oransal frekansının belirlenmesi gerekir.

$f_i^{'} = f_i / N$ $\qquad$ $f_i^{'}$ = oransal frekans

Daha sonra oransal frekanslar 360 ile çarpılarak her farklı değer için dairedeki payı (açısı) bulunur. Böylece bu açılara göre grafik çizilir.

$\alpha_i = f_i^{'} \cdot 360°$ $\qquad$ α_i = oransal frekans derecesi

DİKKAT

Dairesel grafik, oransal frekanslar esas alınarak açılar yardımıyla çizilir.

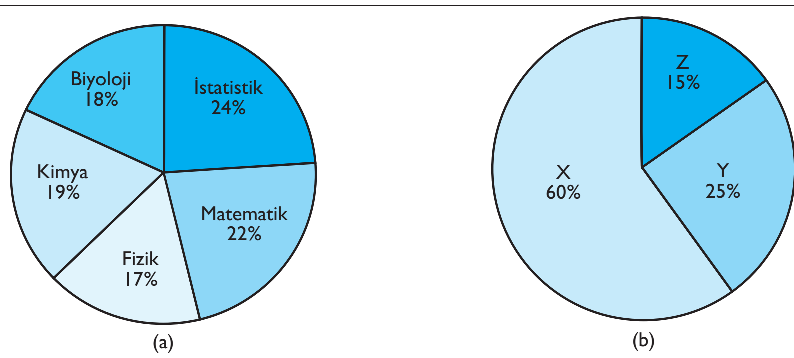
Örnek 6: AÜ Fen Fakültesi bölümlerinde kayıtlı öğrenci sayılarının dağılımının dairesel (pasta) grafik yardımıyla gösterimi

Tablo 1.13
Bölümlere Göre Öğrenci Sayısı ve Oransal Frekansları

BÖLÜMLER	Öğrenci Sayısı (f_i)	Oransal Frekanslar ($f_i^{'}$)	Açı (α_i)
İstatistik	600	$\frac{600}{2500} = 0.24$	(0.24)·360°=86.4°
Matematik	550	$\frac{550}{2500} = 0.22$	(0.22)·360°=79.2°
Fizik	425	$\frac{425}{2500} = 0.17$	(0.17)·360°=61.2°
Kimya	475	$\frac{475}{2500} = 0.19$	(0.19)·360°=68.4°
Biyoloji	450	$\frac{450}{2500} = 0.18$	(0.18)·360°=64.8°
Toplam	2500	**1.00**	360°

Tablo 1.13'e bakıldığında, Fen Fakültesindeki İstatistik bölümü öğrenci sayısının 600 ve fakültedeki öğrenci payının (oransal frekansın) 600/2500=0.24 olduğu hesaplanmıştır. Buna göre, İstatistik bölümü öğrencilerinin fakültedeki oranı %24'tür. Bu oransal frekansın 360°'lik pastadaki payı α_i= (0.24).360°=86.4°'dir. Diğer bölümler için de hesaplamalar benzer şekilde yapılarak, dairesel grafikteki payları bulunur ve grafikteki yerini alır (Şekil 1.1).

Şekil 1.1
Bölümlere Göre Öğrenci Sayısı (Pasta Grafik)



SIRA SİZDE 6

Sözel bir değişkene ilişkin pasta grafiği şekil 1.1 (b)'de verilmiştir. Z değişkeninin 9 adet gözlem değeri olduğuna göre toplam gözlem sayısı kaçtır?

Kartezyen Koordinatlı Grafikler

Zaman Serilerinin Grafik Yardımıyla Gösterimi

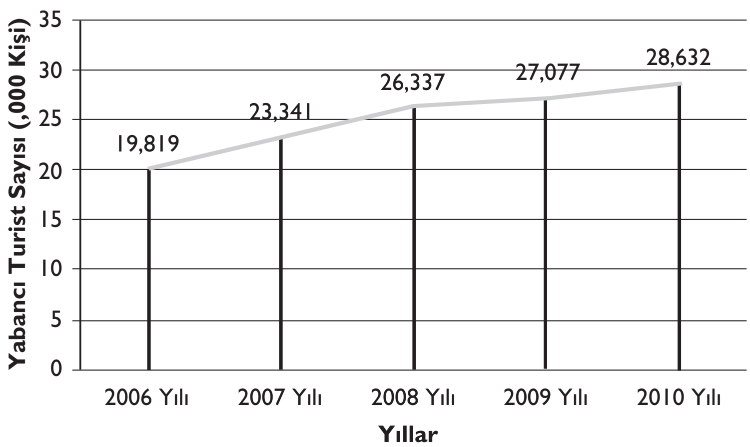
Zamana bağlı bir değişkenin almış olduğu değerlerden oluşan zaman serisinin gösteriminde kartezyen koordinatlı grafik kullanılır. Bu grafikte, yatay eksen zaman değişkenini ve düşey eksen de ilgilenilen değişkeni gösterir. Zaman serilerinin grafikle gösterimi, serilerin içerdiği özelliklerin görsel olarak kolaylıkla görülmesini sağlar. Bu konu ayrıntılı olarak zaman serisi bölümünde ele alınacaktır.

Örnek 7: 2006 - 2010 yılları arası ülkemize gelen yabancı turistlerin sayısı Tablo 1.14'te ve zaman serisi grafiği de Şekil 1.2'de verilmiştir.

Yıllar	Yabancı Turist Sayısı (1000 Kişi)
2006	19.819
2007	23.341
2008	26.337
2009	27.077
2010	28.632

Tablo 1.14
2006 - 2011 Yılları Arasında Ülkemize Gelen Yabancı Turist Sayısı

Tablo 1.14'e bakıldığında, ülkemize 2006 yılında 19.819.000; 2007 yılında 23.341.000; 2008 yılında 26.337.000; 2009 yılında 27.077.000 ve 2010 yılında 28.632.000 yabancı turist giriş yapmıştır. Bu tabloya ilişkin zaman serisi grafiği Şekil 1.2'de görüldüğü gibidir.



Şekil 1.2
2006 - 2011 Yılları Arasında Ülkemize Gelen Yabancı Turist Sayısına İlişkin Zaman Serisi Grafiği

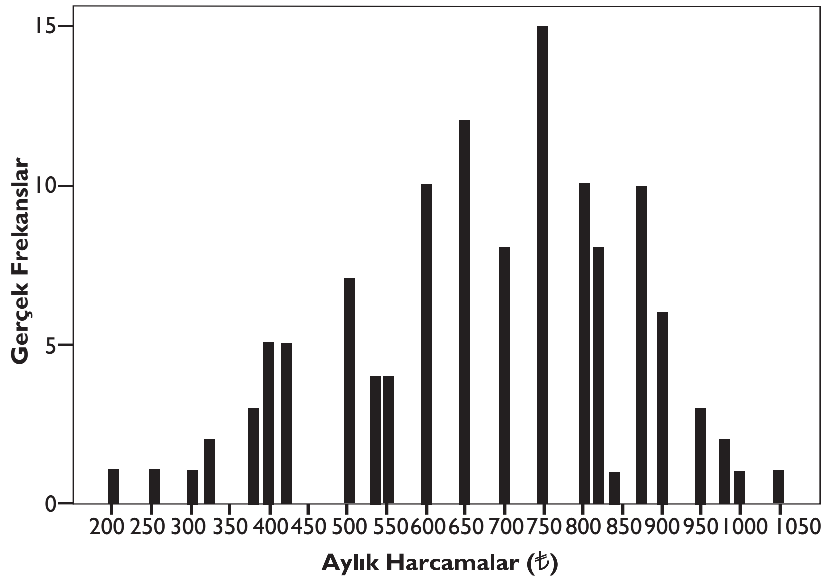
Frekans Serilerinin Grafik Yardımıyla Gösterimi

Frekans serileri çubuk grafikler yardımıyla gösterilir. **Çubuk grafik,** yatay eksende gözlem değerleri ve düşey eksende gözlem değerlerinin kaç kez tekrar ettiğini gösteren frekanslar yardımıyla oluşturulur. Örneğin, Tablo 1.6'da verilen aylık harcamalara ilişkin frekans serisinin çubuk grafiği Şekil 1.3'te gösterilmiştir.

Çubuk grafik, frekans serisinin grafikle gösteriminde kullanılır.

Şekil 1.3

120 Öğrencinin Aylık Harcamalarına İlişkin Çubuk Grafik



Gruplandırılmış Serilerin Grafik Yardımıyla Gösterimi

Histogram, gruplandırılmış seride sınıf aralıklarında bulunan gözlem sayılarını alanlar yardımıyla (dikdörtgenlerle) gösteren grafiktir.

Gruplandırılmış serilerin grafikle gösteriminde **histogram** kullanılır. Histogram, gruplandırılmış seride ilgili sınıf aralığında bulunan frekansları alanlar yardımıyla gösteren ve dikdörtgenlerden oluşan bir grafiktir. Bu nedenle, bu grafiğin çizilmesinde ilgili frekans değerleri ayarlanır. Daha sonra, yatay eksende sınıf aralıkları ve düşey eksende ayarlanmış frekanslar kullanılarak histogram oluşturulur.

$$f_i' = f_i / h$$

f_i' = *ayarlanmış frekans*
f_i = *frekans*
h = *sınıf aralığı büyüklüğü*

Örneğin, Tablo 1.7'de verilen aylık harcamalara ilişkin gruplandırılmış serinin ayarlanmış frekansları Tablo 1.15'te görüldüğü gibi hesaplanmıştır.

Tablo 1.15
120 Öğrencinin Aylık Harcamalara İlişkin Ayarlanmış Frekanslar

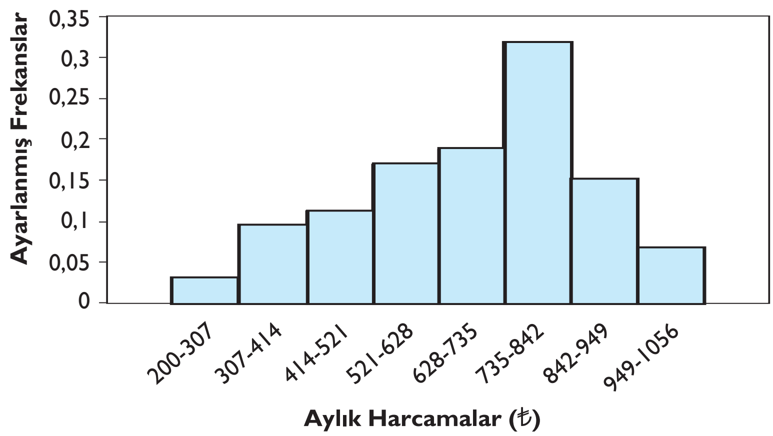
Aylık Harcama (₺)	Frekanslar (f_i)	Ayarlanmış Frekanslar ($f_i' = f_i / h$)
200 - 307'den az	3	$\frac{3}{107} = 0.028$
307 - 414'den az	10	$\frac{10}{107} = 0.093$
414 - 521'den az	12	$\frac{12}{107} = 0.112$
521 - 628'den az	18	$\frac{18}{107} = 0.168$
628 - 735'den az	20	$\frac{20}{107} = 0.187$
735 - 842'den az	34	$\frac{34}{107} = 0.318$

842 - 949'dan az	16	$\frac{16}{107} = 0.149$
949 - 1056'dan az	7	$\frac{7}{107} = 0.065$
Toplam	**120**	

Tablo 1.15
Devamı

Histogram, Tablo 1.15'te verilen sınıf aralıkları yatay eksende ve bu sınıf aralıklarına ilişkin ayarlanmış frekanslar düşey eksende olmak üzere Şekil 1.4'te görüldüğü gibi çizilir.

Şekil 1.4



Aylık Harcamalara İlişkin Histogram (Ayarlanmış Frekanslar ile)

Histogram çiziminde sınıf aralıkları eşit olabileceği gibi farklı da olabilir. Önemli olan ilgili sınıf aralığında bulunan gözlem sayısının alanlar yardımıyla gösterilmesidir. Histogramda dikdörtgenlerin alanları toplamı, araştırmada ele alınan toplam gözlem sayısını verir.

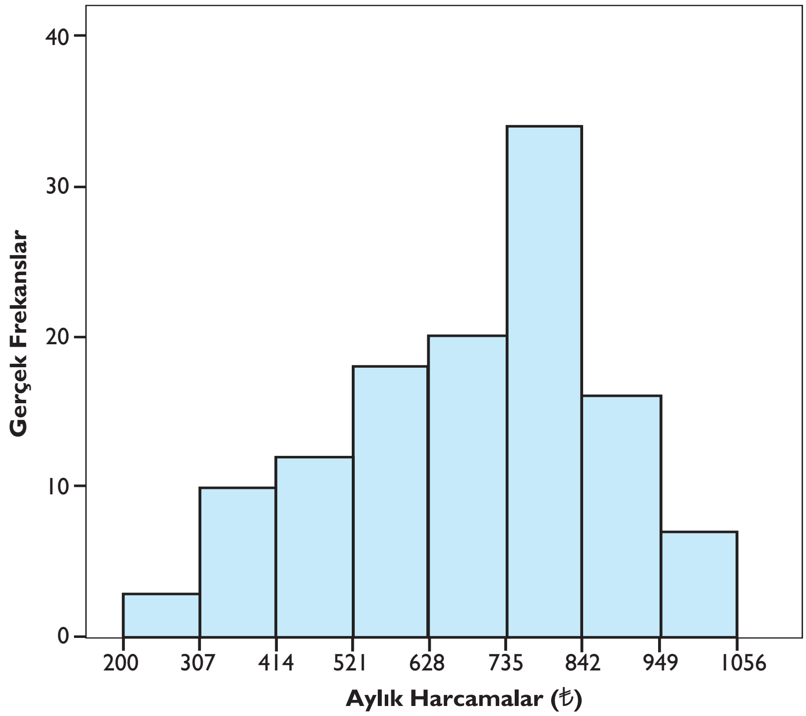
Örneğin Şekil 1.4'e bakıldığında,

(200-307) aralığı için; f_i = (307-200) · (0.028) = 2.996 ≅ 3
(307-414) aralığı için; f_i = (414-307) · (0.093) = 9.951 ≅ 10
(414-521) aralığı için; f_i = (521-414) · (0.112) = 11.984 ≅ 12
(521-628) aralığı için; f_i = (628-521) · (0.168) = 17.976 ≅ 18
(628-735) aralığı için; f_i = (735-628) · (0.187) = 20.009 ≅ 20
(735-842) aralığı için; f_i = (842-735) · (0.318) = 34.026 ≅ 34
(842-949) aralığı için; f_i = (949-842) · (0.149) = 15.943 ≅ 16
(949-1056) aralığı için; f_i = (1056-949) · (0.065) = 6.955 ≅ 7

Ancak, uygulamada kolaylık olması nedeniyle histogram gösteriminde ayarlanmış frekanslar yerine doğrudan gerçek frekans değerleri de kullanılır. Örneğin, Tablo 1.7'de verilen aylık harcamalara ilişkin gruplandırılmış serinin histogramı gerçek frekanslar yardımıyla çizildiğinde, Şekil 1.5'te görüldüğü gibidir.

Şekil 1.5

Aylık Harcamalara İlişkin Histogram (Gerçek Frekanslar ile)

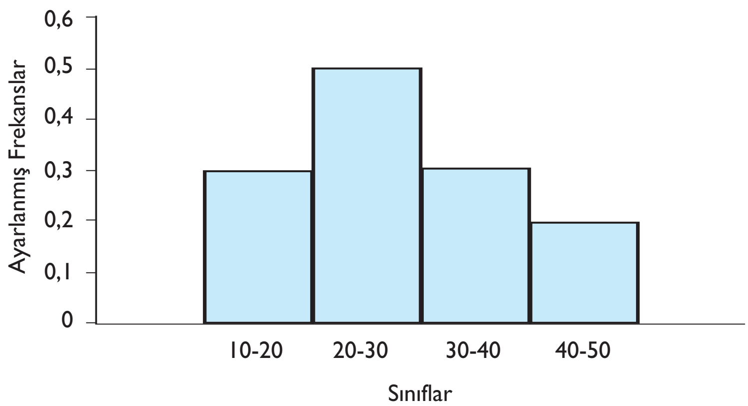


SIRA SİZDE 7

Bir histogram oluşturulurken dikkat edilmesi gereken noktaları belirleyiniz.

SIRA SİZDE 8

Aşağıdaki histograma göre (20 - 30) sınıf aralığında bulunan gözlem sayısı (frekans) nedir?



SIRA SİZDE 9

AÖF öğrencilerinin boy uzunlukları ile ilgili yapılan bir çalışmada, Eskişehir Merkez AÖF Bürosuna kayıtlı 100 adet öğrencinin ölçülen boy uzunlukları izleyen tabloda verilmiştir.

Boy Uzunluğu Sınıfları (cm)	Frekanslar (f_i)
150 - 160'den az	10
160 - 170'den az	20
170 - 180'den az	35
180 - 190'den az	25
190 - 200'den az	10
Toplam	**100**

a. Yukarıda verilen tabloya ne ad verilir?
b. Bu araştırma için; anakütle, örneklem, istatistik birimi ve değişkeni belirleyiniz.
c. Histogramını çiziniz (Ayarlanmış frekansları kullanınız.).
d. Histogramını çiziniz (Gerçek frekansları kullanınız.).
e. "-den az" ve "-den çok" birikimli serilerini oluşturunuz.

DİKKAT

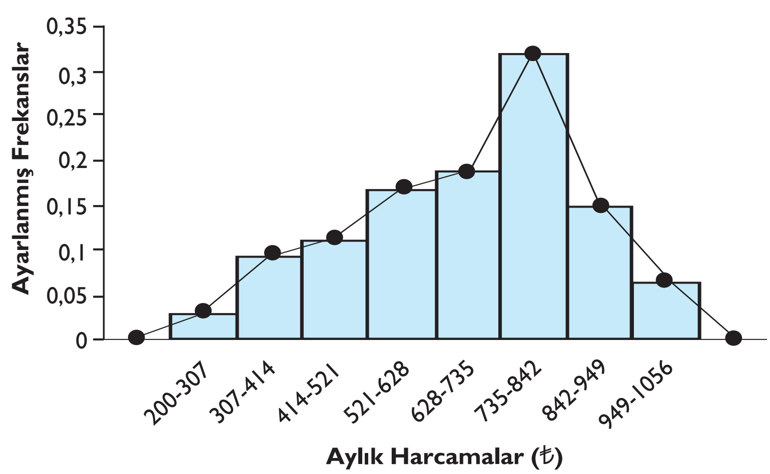
Çubuk grafik gösteriminde çubukların boyu ilgili değere sahip gözlem sayısını gösterirken histogramda dikdörtgenlerin alanı ilgili sınıf aralığında bulunan gözlem sayısını verir.

Frekans Poligonu

Gruplandırılmış serilerin gösteriminde kullanılan bir diğer grafik de **frekans poligonudur.** Histogram çizildikten sonra her sınıf aralığındaki dikdörtgenlerin üst orta noktaları belirlenir. Daha sonra, ilk sınıf aralığındaki dikdörtgenin sol orta noktası ve son sınıf aralığındaki dikdörtgenin sağ orta noktası belirlenir. Belirlenen tüm noktalar birleştirilerek oluşturulan eğriye frekans poligonu denir. Şekil 1.6'da görüldüğü gibi frekans poligonu altında kalan alan da toplam gözlem sayısını verir.

Histogram çizildikten sonra, her bir sınıf aralığındaki dikdörtgenlerin üst orta noktalarının birleştirilmesiyle oluşturulan eğriye **freakans poligonu** denir.

Örneğin, aylık harcamalara ilişkin gruplandırılmış seri için (Tablo 1.7) oluşturulan frekans poligonu Şekil 1.6'da görüldüğü gibidir.



Şekil 1.6

Aylık Harcamalara İlişkin Frekans Poligonu

"-den az" grafiği, birikimli serilerde yatay eksende sınıfların üst noktaları ve düşey eksende karşı gelen birikimli frekansların belirledikleri noktaların birleştirilmesiyle oluşan ve belli bir değerden az gözlem sayısını gösteren grafiktir.

"-den çok" grafiği, birikimli serilerde yatay eksende sınıfların alt noktaları ve düşey eksende karşı gelen birikimli frekansların belirledikleri noktaların birleştirilmesiyle oluşan ve belli bir değer ve değerden çok gözlem sayısını gösteren grafiktir.

Birikimli Serilerin Grafik Yardımıyla Gösterimi

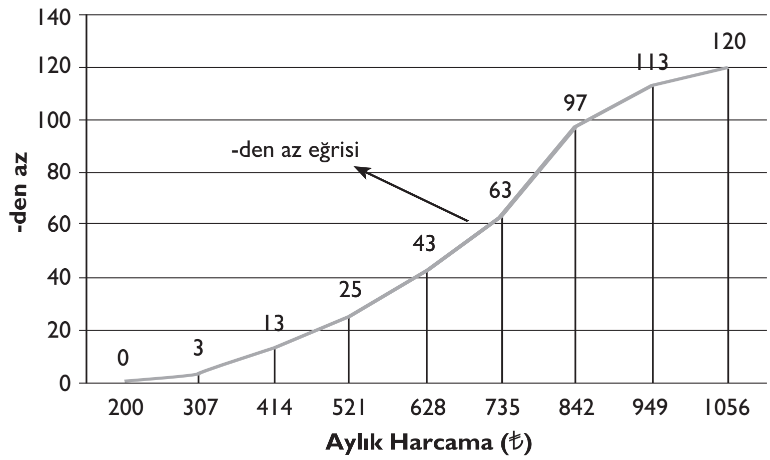
Birikimli serilerin grafikle gösteriminde **"-den az**" ve **"-den çok"** grafiklerinden yararlanılır. "-den az" grafiği, yatay eksende sınıfların üst noktaları ve düşey eksende karşı gelen birikimli frekansların belirledikleri noktaların birleştirilmeleriyle; "-den çok" grafiği de yatay eksende sınıf alt noktaları ve düşey eksende karşı gelen birikimli frekansların belirledikleri noktaların birleştirilmeleriyle oluşturulan grafiklerdir.

"-den az" grafiğinde oluşan eğriye "-den az eğrisi", "-den çok" grafiğinde oluşan eğriye de "-den çok eğrisi" denir. Bu eğrilere dayanarak, teorik olarak belirlenen herhangi bir değerden çok veya herhangi bir değerden az gözlem sayısı bulunabilir.

Örneğin, aylık harcamalara ilişkin gruplandırılmış seri için (Tablo 1.9) oluşturulan "-den az" birikimli serisi için "-den az" grafiği Şekil 1.7'de; "-den çok" birikimli serisi için "-den çok" grafiği Şekil 1.8'de görüldüğü gibidir.

Şekil 1.7

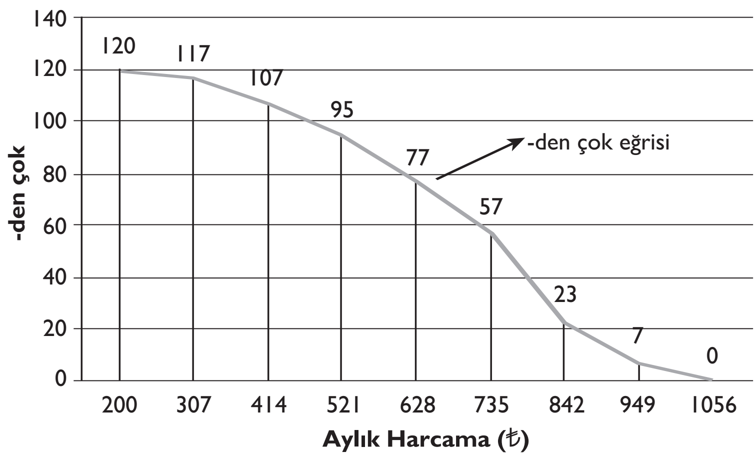
Aylık Harcamalara İlişkin -Den Az Birikimli Seri Grafiği



Şekil 1.7'ye bakıldığında herhangi bir değerden az gözlem sayısına kolayca ulaşılabilir. Örneğin, aylık harcaması ₺200'den az 0; ₺307'den az 3; ₺414'den az 13 kişi vb. şekilde diğer aylık harcama değerleri için de, aynı şekilde istenen kişi sayısı bulunabilir.

Şekil 1.8

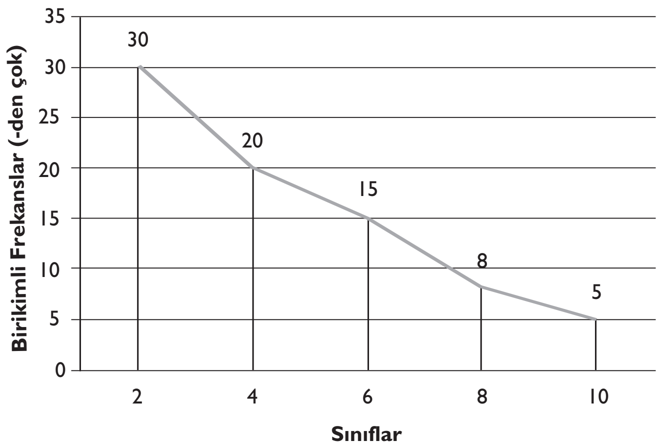
Aylık Harcamalara İlişkin -Den Çok Birikimli Seri Grafiği



Şekil 1.8'e bakıldığında herhangi bir değer ve değerden çok gözlem sayısına kolayca ulaşılabilir. Örneğin, aylık harcaması 200 ve 2₺00'den çok 120; 307 ve ₺307'den çok 117; 414 ve ₺414'den çok 107 kişi vb. şekilde diğer aylık harcama değerleri için de, aynı şekilde istenen kişi sayısı bulunabilir.

SIRA SİZDE 10

Aşağıdaki grafiğe göre (6 - 8) sınıf aralığında kaç tane gözlem bulunur?

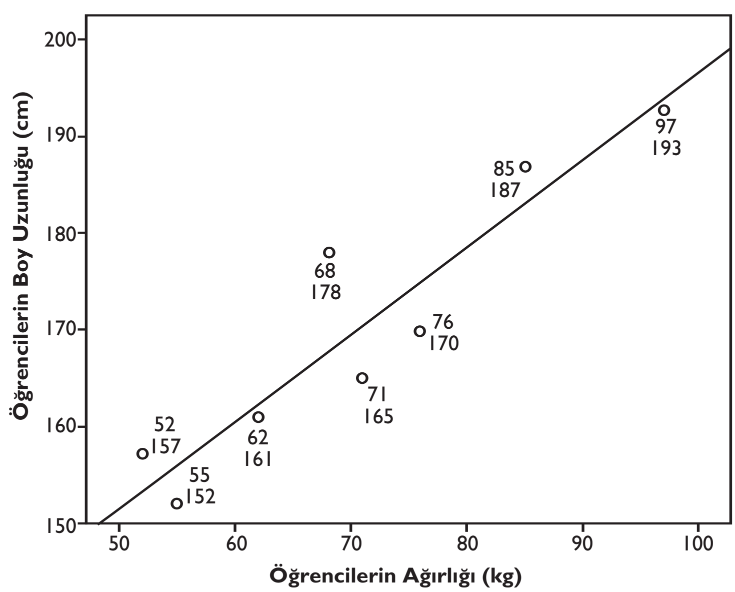


Bileşik Serilerin Grafik Yardımıyla Gösterimi

Bileşik serilerde değişkenler arasındaki ilişkiyi gösteren grafiğe **serpilme grafiği (diyagramı)** denir. Serpilme grafiğinde yatay eksende değişkenlerden biri, düşey eksende bir diğeri olmak üzere, istatistik birimine ilişkin iki değişken için sayı çifti işaretlenir. İşaretlenen bu noktaların dağılması (serpilmesi) değişkenler arasındaki ilişki hakkında görsel olarak bilgi verir.

Serpilme grafiği (diyagramı), bileşik serilerde değişkenler arasındaki ilişkiyi gösteren grafiktir.

Örneğin, Tablo 1.10'da verilen öğrencilerin boyu ve ağırlığına ilişkin bileşik seri için serpilme grafiği Şekil 1.9'da görüldüğü gibidir. Bu grafikte, yatay eksende öğrencilerin ağırlığı, düşey eksende öğrencilerin boy uzunluğu değişken olarak alınmıştır.



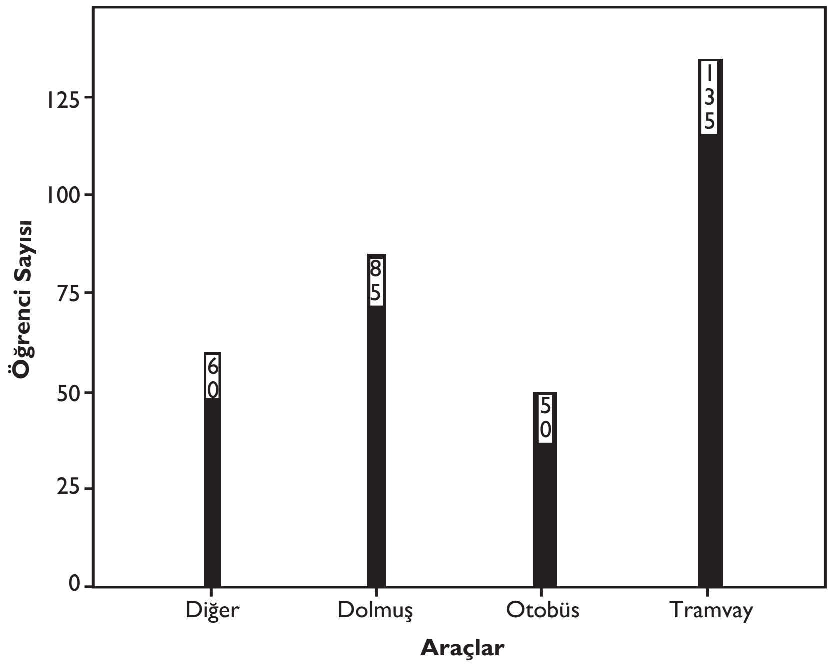
Şekil 1.9

Öğrencilerin Boy Uzunluğu ve Ağırlığına İlişkin Serpilme Grafiği

Şekil 1.9’daki grafiğe bakıldığında öğrencilerin ağırlıkları ile boy uzunlukları arasında doğrusal bir ilişkinin olduğu görülür.

SIRA SİZDE 11

İstatistik bölümünde okuyan öğrencilerin okula hangi araçla geldiklerine ilişkin veriler izleyen grafikte verilmiştir.



a. Buna göre, istatistik bölümünde okuyan kaç tane öğrenci vardır?
b. Bu grafiğe ne ad verilir?

Özet

AMAÇ 1

İstatistiğin temel kavramlarını tanımlamak.

Yapılacak araştırmada istatistik birimi, anakütleyi, örneklemi, değişkeni, değişkenlerin ölçme düzeyini belirlemektir.

İstatistik birimi: Sayılabilir veya ölçülebilir özellikleri (değişkenleri) içeren, aralarında bir çok benzerlikler olmakla beraber farklılıklar da bulunan nesnelere veya olaylara “istatistik birimi” denir.

Anakütle: Araştırmaya ilişkin tanımlanan istatistik birimlerin tümünün oluşturduğu topluluğa anakütle denir.

Örneklem: Anakütlenin alt topluluğuna örneklem denir.

Değişken: İstatistik birimlerin sahip oldukları ve farklı değerler alabilen, diğer istatistik birimlerinden ayırt edilmesini sağlayan özelliklere değişken denir.

Ölçme ve Ölçüm: Anakütle veya örneklemdeki istatistik birimlerin, ilgilenilen sayısal veya sözel özelliklerinin aldığı değerlerin, sayılar veya simgelerle gösterimine ölçme denir. Ölçme sonucu değişkenin aldığı değere de ölçüm denir.

AMAÇ 2

İstatistiğin temel kavramlarını kendi içinde ayırt etmek.

İstatistiğin temel kavramlarını kendi içindeki özelliklerine göre farklı türlere ayırabilmektir.

İstatistik birimi, canlı, cansız bir olgu, bir olay veya bir kurum olabilir. İstatistik birimine istenilen bir an ulaşılabiliyorsa sürekli istatistik birimi; belli bir anda ortaya çıkıyorsa ani istatistik birimi adını alır.

İstatistik birimlerin tümüne ulaşılabiliyorsa bu birimlerden oluşan topluluğa somut anakütle; ulaşılamadığı durumlardan oluşan topluluğa da soyut anakütle denir.

Gerçekte var olan istatistik birimlerinden oluşan topluluğa gerçek anakütle; var olmaları ya da ortaya çıkmaları olası istatistik birimlerinden oluşan topluluğa da varsayımsal anakütle denir.

Sürekli istatistik birimlerinden oluşan topluluğa hazır anakütle; ani istatistik birimlerinden oluşan topluluğa da hareketli anakütle denir.

Değişkenlerin aldığı değerler sayısal olarak ifade ediliyorsa bu değişkene sayısal değişken; sözel olarak ifade ediliyorsa sözel değişken denir.

Sınıflayıcı ölçek: Ölçme düzeyi en düşük olan ve birimi olmayan ölçektir.

Sıralayıcı ölçek: Sınıflamanın yanında sayı ve simgelerde, büyüklük ve küçüklük kavramının olduğu ölçektir.

Aralıklı ölçek: Değişkenin aldığı sayısal değerlerin birimle ifade edildiği ve sayılar arasındaki farkın anlam kazandığı ölçektir.

Oransal ölçek: Eşit aralıklı ölçeğe ek olarak gerçek bir sıfır noktasının olduğu ölçektir.

AMAÇ 3

İstatistik serilerini açıklamak.

Liste: İlgilenilen değişkenin almış olduğu değerler, diğer bir değişkene göre veya rastgele sıralanmış ise bu tabloya "liste" denir.

Basit seri: İlgilenilen değişkenin almış olduğu değerlerin küçükten büyüğe veya büyükten küçüğe sıralanmasıyla oluşan seriye basit seri denir.

Frekans serisi: İlgilenilen değişkenin almış olduğu farklı değerlerin küçükten büyüğe sıralanması ve bu değerlerin karşısına kaç kez tekrar ettiğinin (frekansı) yazılmasıyla oluşturulan istatistik serisine frekans serisi denir.

Gruplandırılmış seri: İlgilenilen değişken değerlerinin, belirlenen sınıflara (aralıklara) ayrılması ve bu sınıflara giren gözlem sayısının ayrı bir sütuna yerleştirilmesiyle oluşan seriye gruplandırılmış seri denir.

Birikimli seri: Gruplandırılmış seride her sınıfın frekansına, sırası ile izleyen sınıfların frekansları eklenerek veya toplam frekanstan eksiltilerek oluşturulan seriye birikimli seri; birikimli seride her sınıfa karşı gelen frekansa da birikimli frekans denir.

-den az serisi: Gruplandırılmış seride, sınıf aralıklarındaki frekanslar birbirine eklenerek oluşturulan seriye "-den az" serisi denir.

"-den çok" serisi: Gruplandırılmış seride, sınıf aralıklarındaki frekansların toplam frekanstan eksiltilmesiyle oluşturulan seriye "-den çok" serisi denir.

Bileşik seri: İstatistik birimlerin, iki veya daha fazla değişkenine göre aldığı değerleri birlikte gösteren serilere bileşik seri denir.

AMAÇ 4

İstatistik serileri grafik yardımı ile açıklamak.

Dairesel (Pasta) grafik, sözel değişkenlerin oransal frekanslarını göstermek için kullanılan grafiktir.

Çubuk grafik, frekans serisinin grafikle gösteriminde kullanılır.

Histogram, gruplandırılmış seride sınıf aralıklarında bulunan gözlem sayılarını alanlar yardımıyla (dikdörtgenlerle) gösteren grafiktir.

Frekans poligonu, histogram çizildikten sonra, her bir sınıf aralığındaki dikdörtgenlerin üst noktalarının birleştirilmesiyle oluşturulan eğridir. Altında kalan alan ise toplam gözlem sayısını verir.

"-den az" grafiği, birikimli serilerde yatay eksende sınıfların üst noktaları ve düşey eksende karşı gelen birikimli frekansların belirledikleri noktaların birleştirilmesiyle oluşan ve belli bir değerden az gözlem sayısını gösteren grafiktir.

"-den çok" grafiği, birikimli serilerde yatay eksende sınıfların alt noktaları ve düşey eksende karşı gelen birikimli frekansların belirledikleri noktaların birleştirilmesiyle oluşan ve belli bir değer ve değerden çok gözlem sayısını gösteren grafiktir.

Serpilme grafiği (diyagramı), bileşik serilerde değişkenler arasındaki ilişkiyi gösteren grafiktir.

Kendimizi Sınayalım

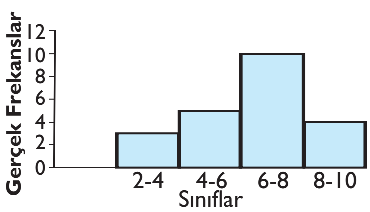
1. Aşağıdakilerden hangisi iki değişken için belirlenen gözlem değerlerini aynı anda veren grafiktir?

a. Histogram
b. Frekans poligonu
c. Zaman serisi grafiği
d. Serpilme diyagramı
e. Dairesel grafik

2. Aşağıdakilerden hangisi ölçek türlerinden biri **değildir?**

a. Sınıflayıcı ölçek
b. Birikimli ölçek
c. Oransal ölçek
d. Aralıklı ölçek
e. Sıralayıcı ölçek

3. Verilen histograma göre toplam gözlem sayısı (frekans) nedir?



a. 7
b. 8
c. 10
d. 15
e. 22

4. Histogramlarla ilgili aşağıdakilerden hangisi **yanlıştır?**

a. Gruplandırılmış serilerin grafikle gösteriminde kullanılır.
b. Dikdörtgenler kullanılarak çizilir.
c. Sınıf aralıkları eşit olmak zorunda değildir.
d. Dikdörtgenlerin alanı ilgili sınıf aralığındaki gözlem sayılarını verir.
e. Bileşik serilerin grafikle gösteriminde kullanılır.

5. Aşağıdaki veri seti hangi tür serilere örnek oluşturur?

Aylar	Ortalama Sıcaklık (°C) (Eskişehir)
Ocak	-0.1
Şubat	1.9
Mart	5.5
Nisan	10.6
Mayıs	15.2
Haziran	19.0
Temmuz	21.4
Ağustos	21.1
Eylül	17.3
Ekim	11.9
Kasım	6.7
Aralık	2.4

a. Zaman Serisi
b. Bileşik Seri
c. Birikimli Seri
d. Gruplandırılmış Seri
e. Mekan Serisi

6. Öğrencilerin (göz renklerine) ilişkin bir çalışmadaki "göz rengi" değişkeni ne tür bir değişkendir?

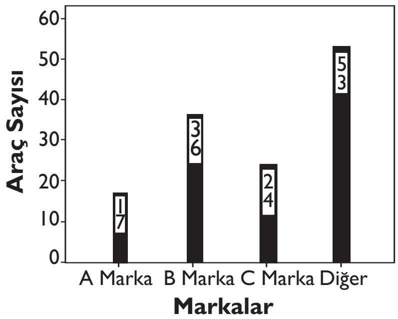
a. Nicel
b. Zaman
c. Sayısal
d. Sözel
e. Ani

7. Bir frekans dağılımına ilişkin "-den az" serisi aşağıda verilmiştir. Gerçek frekansları bulunuz.

Sınıflar	-den az
0 - 4'den az	7
4 - 8'den az	10
8 - 12'den az	18
12 - 16'dan az	20
16 - 20'den az	25

a.	b.	c.	d.	e.
f_i	f_i	f_i	f_i	f_i
5	25	5	7	7
2	18	7	3	8
8	15	15	8	3
3	7	18	2	5
7	5	25	5	2

8. ve 9. Soruları aşağıdaki grafiğe göre yanıtlayınız. Bir alışveriş merkezinin otoparkında bulunan araçların markalara göre dağılımı izleyen grafikte verilmiştir.



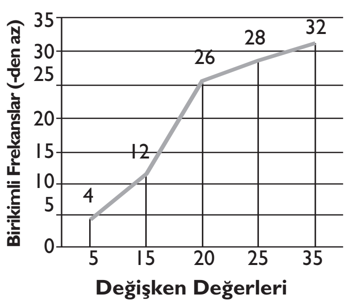
8. Buna göre; C marka araçtan kaç tane vardır?

a. 17
b. 24
c. 36
d. 53
e. 130

9. Bu grafiğe ne ad verilir?

a. Çubuk
b. Histogram
c. Pasta
d. Serpilme Diyagramı
e. "-den az" grafiği

10. Aşağıda verilen "-den az" grafiğine göre değeri 15'den az gözlem sayısı kaç tanedir?



a. 3
b. 5
c. 12
d. 15
e. 22

Kendimizi Sınayalım Yanıt Anahtarı

1. d Yanıtımız yanlış ise "Birikimli Serilerin Grafik Yardımıyla Gösterimi" konusunu gözden geçiriniz.
2. b Yanıtımız yanlış ise "Ölçme Düzeyleri (Ölçekler)" konusunu gözden geçiriniz.
3. e Yanıtımız yanlış ise " Gruplandırılmış Serilerin Grafik Yardımıyla Gösterimi" konusunu gözden geçiriniz.
4. e Yanıtımız yanlış ise, "Gruplandırılmış Serilerin Grafik Yardımıyla Gösterimi" konusunu gözden geçiriniz.
5. a Yanıtımız yanlış ise "Zaman Serilerinin Grafik Yardımıyla Gösterimi" konusunu gözden geçiriniz.
6. d Yanıtımız yanlış ise "Değişken Türleri" konusunu gözden geçiriniz.
7. d Yanıtımız yanlış ise "Birikimli Seriler" konusunu gözden geçiriniz.
8. b Yanıtımız yanlış ise, "Frekans Serilerinin Grafik Yardımıyla Gösterimi" konusunu gözden geçiriniz.
9. a Yanıtımız yanlış ise "Frekans Serilerinin Grafik Yardımıyla Gösterimi" konusunu gözden geçiriniz.
10. c Yanıtımız yanlış ise "Birikimli Serilerin Grafik Yardımıyla Gösterimi" konusunu gözden geçiriniz.

Sıra Sizde Yanıt Anahtarı

Sıra Sizde 1

Öğrencinin ağırlığı	→	sayısal ve sürekli
Göz rengi	→	sözel
Ailedeki kişi sayısı	→	sayısal ve kesikli
Medeni durum	→	sözel

Sıra Sizde 2

Cinsiyet	→	sınıflayıcı ölçek
Ders kodları	→	sınıflayıcı ölçek
Mezuniyet derecesi	→	sıralayıcı ölçek
Sıcaklık	→	aralıklı ölçek
Boy uzunluğu	→	oransal ölçek
Aylık gelir	→	oransal ölçek

Sıra Sizde 3

a. 30
b. (4-0) = (9-5) = ... = (24-20) = 4
c. (15 - 19) sınıf aralığında 15, (20 - 24) sınıf aralığında 10 gün bulunduğundan,
15+10=25

Sıra Sizde 4

40'tan az 20 gözlem olduğundan, toplam gözlem sayısı olan 30'dan 20 eksilterek sonuca ulaşılır.

30-20=10

Sıra Sizde 5

30 öğrencinin istatistik dersinden aldığı başarı notlarının "-den az" ve "-den çok" birikimli serileri aşağıdaki gibidir.

Başarı Notları	f_i	(-den az)	(-den çok)
AA	2	2	30
BB	7	2+7=9	30-2=28
CC	11	9+11=20	28-7=21
DD	3	20+3=23	21-11=10
FF	7	23+7=30	10-3=7
Toplam	**30**		

Sıra Sizde 6

%15'lik dilime 9 adet gözlem değeri düşüyorsa %100'lük dilimin kaç olacağı,

$$\text{Toplam gözlem sayısı} = \frac{9}{0.15} = 60$$

olarak hesaplanır.

Sıra Sizde 7

Histogram, gruplandırılmış seride ilgili sınıf aralığında bulunan frekansları alanlar yardımıyla gösteren ve dikdörtgenlerden oluşan bir grafik olduğundan bu grafiğin çizilmesinde ilgili frekans değerlerini ayarlamak (oranlamak) en önemli noktadır. Daha sonra, yatay eksende sınıf aralıkları ve düşey eksende oranlanmış frekanslar kullanılarak histogram oluşturulur.

Sıra Sizde 8

Histograma bakıldığında 20 - 30 serisi için sınıf aralığı büyüklüğü,

30-20=10

olur. İstenilen gözlem sayısını bulmak için de sınıf aralığı büyüklüğü ile oransal frekans değerini çarparak,

10.(0.5)=5

sonucuna ulaşılır.

Sıra Sizde 9

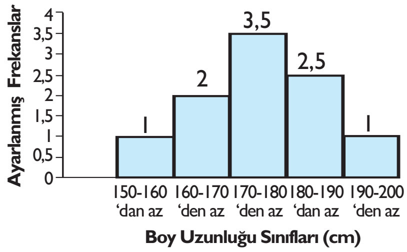
a. Gruplandırılmış Seri
b. Anakütle: AÖF'te okuyan bütün öğrencilerin oluşturduğu topluluk.
Örneklem: Eskişehir Merkez AÖF Bürosuna bağlı 100 adet öğrencinin oluşturduğu topluluk.
İstatistik birimi: AÖF öğrencilerinin her biri.
Değişken: AÖF öğrencilerinin boyu.
c. Sınıf aralığı büyüklüğü;

(160-150)=(170-160)=...=(200-190)=10'dur.

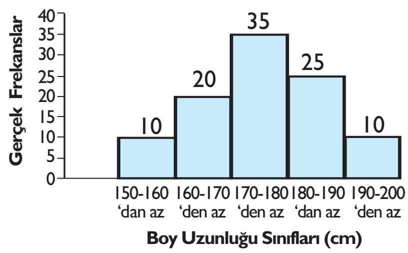
Ayarlanmış frekansların tablosu,

Boy Uzunluğu Sınıfları (cm)	Frekanslar (f_i)	Ayarlanmış Frekanslar (f_i')
150 - 160'dan az	10	10/10=1
160 - 170'den az	20	20/10=2
170 - 180'den az	35	35/10=3.5
180 - 190'dan az	25	25/10=2.5
190 - 200'den az	10	10/10=1

şeklindedir. Buna göre, histogram:



d. Gerçek frekanslar kullanılarak çizilen histogram:



e.

Boy Uzunluğu Sınıfları (cm)	Frekanslar (f_i)	$\sum f_i$ (-den az)	$\sum f_i$ (-den çok)
150 - 160'dan az	10	10	100
160 - 170'den az	20	10+20=30	100-10=90
170 - 180'den az	35	30+35=65	90-20=70
180 - 190'dan az	25	65+25=90	70-35=35
190 - 200'den az	10	90+10=100	35-25=10
Toplam	**100**		

Sıra Sizde 10

6 ve 6'dan çok 15 gözlem; 8 ve 8'den çok 8 gözlem bulunuyor. 6 ile 8 arasında da,

15-8=7

gözlem bulunmaktadır.

Sıra Sizde 11

a. 50+60+85+125=320

b. Bu grafik çubuk grafiktir.

Yararlanılan ve Başvurulabilecek Kaynaklar

Atlas, Mahmut (2010). **İstatistik 1,** Ak Ofset Matbaacılık

Casella, George; Berger Roger L. (2001). **Statistical Inference,** Duxbury, Thomson Learning

Çömlekçi, Necla (1989). **İstatistik,** Bilim Teknik Yayınevi

Huntsberger, David V. (1981). **Elements of Statistical Inference**

Serper, Ö. (2004). **Uygulamalı İstatistik 1,** Ezgi Kitabevi.

2

Amaçlarımız

Bu üniteyi tamamladıktan sonra;

- Merkezi eğilim ve değişkenlik ölçülerini sayabilecek, hangi durumda hangi ölçüyü hesaplamanız gerektiğine karar verebilecek,
- Herhangi bir veri seti için uygun ortalamayı hesaplayıp yorumlayabilecek,
- Herhangi bir veri seti için uygun değişkenlik ölçüsünü hesaplayıp yorumlayabilecek,
- Serilerin birbirlerine göre değişkenliğini kıyaslayabileceksiniz.

Anahtar Kavramlar

- Merkezî Eğilim Ölçüleri
- Değişkenlik Ölçüleri
- Ortalama
- Varyans ve Standart Sapma

İçindekiler

Merkezî Eğilim ve Değişkenlik Ölçüleri

GİRİŞ

Anakütle ya da örneklemde yer alan verilerin grafiklerle sunulması genellikle yeterli olmaz. Gözlemlenen değerler kümesinin özetlenmesi, merkezinin ve verilerin dağılım merkezi etrafında nasıl değiştiğinin betimlemesine ihtiyaç duyulur. Veri setini betimlemek için kullanılan özet ölçülerine merkezî eğilim ölçüleri ya da ortalama, verilerin dağılım merkezi etrafında nasıl değiştiğini gösteren ölçülere ise değişkenlik ölçüleri denir.

Elde edilen sayısal betimleme ölçülerinden anakütle için geliştirilenlerle örneklemler için geliştirilenler farklı tanımlanır. Anakütle için hesaplanan sayısal betimleyici ölçümler parametre, örneklem için hesaplanan sayısal betimleyici ölçümler ise istatistik olarak adlandırılır. İstatistiksel çıkarsama gerektiren problemlerde her zaman parametre değerleri hesaplanamaz. Ancak, bu parametrelere karşı gelen istatistikler hesaplanabilir ve bu nicelikler ilgili anakütle parametrelerini tahmin etmede kullanılır. Ünite kapsamında söz konusu ölçüler önce parametreler, ardından örneklem istatistikleri için verilecektir.

MERKEZÎ EĞİLİM ÖLÇÜLERİ (ORTALAMALAR)

Veri setini betimlemede ilk akla gelen özet ölçüleri merkezî eğilim ölçüleri ya da **ortalamalar**dır. Bu kitap kapsamında verilecek olan bu ölçülerden aritmetik ortalama ve kareli ortalama verilerin tümünün hesaba katıldığı ortalamalarken (duyarlı ortalamalar), mod, medyan, dördebölenler (kartiller) hesaplamalarda verilerin tümünün hesaba katılmadığı ortalamalardır (duyarlı olmayan ortalamalar).

Ortalamalar her zaman serinin en küçük gözlem değeri ile en büyük gözlem değeri arasında yer alır. $x_{min} \leq$ ortalama $\leq x_{max}$

Aritmetik Ortalama

Günlük hayatta en sık kullanılan merkezî eğilim ölçüsü olan aritmetik ortalama, seriyi oluşturan tüm gözlem değerlerinin gözlem sayısına oranı olarak tanımlanır. Aritmetik ortalama konum olarak verilerin en çok hangi değer etrafında toplandığının ya da yoğunlaştığının sayısal bir ölçüsüdür.

Hangi tür ortalama olduğu belirtilmeden sadece ortalamadan söz ediliyorsa kastedilen aritmetik ortalamadır.

DİKKAT

Basit Serilerde Aritmetik Ortalama Hesabı

Bir anakütle üzerinde tanımlanmış X değişkeninin sayısal olarak aldığı değerler $x_1, x_2, \ldots, x_N$ ile gösterilsin. Ünite 1'den de hatırlanacağı gibi bu değerler basit bir seri ile ifade edilebilir. Bu durumda, n_i değişkeninin aldığı değerler

$$x_i : x_1 \quad x_2 \quad \ldots \quad x_N$$

şeklinde gösterilir ve anakütle aritmetik ortalaması,

$$\mu = \frac{x_1 + x_2 + \ldots + x_N}{N} = \frac{\sum_{i=1}^{N} x_i}{N}$$

formülü ile hesaplanır. N toplam frekanstır.

Bir anakütleden çekilen n birimlik örneklem için X değişkeninin sayısal olarak aldığı değerler $x_1, x_2, \ldots, x_n$ ile gösterilsin. Bu durumda X değişkeninin aldığı değerler

$$x_i : x_1 \quad x_2 \quad \ldots \quad x_n$$

şeklinde gösterilir ve örneklem aritmetik ortalaması,

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$

şeklinde hesaplanır. n toplam frekanstır.

DİKKAT

Her zaman $n \leq N$ olduğuna dikkat ediniz.

Örnek 1: *6 öğrencinin kayıtlı olduğu bir seçmeli derste alınan final notları aşağıdaki gibidir. Öğrencilerin final notlarının aritmetik ortalamasını bulunuz.*

x_i: 30 40 50 60 70 80

Çözüm: Seçmeli derse 6 öğrenci kayıtlı olduğundan, bunun bir anakütle olduğu düşünülmelidir. $N = 6$ olmak üzere anakütle aritmetik ortalaması

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} = \frac{\sum_{i=1}^{6} x_i}{6} = \frac{30 + 40 + 50 + 60 + 70 + 80}{6} = \frac{330}{6} = 55$$

olarak bulunur. Buna göre ilgili seçmeli derse kayıtlı öğrencilerin final notu ortalaması 55'tir.

Örnek 2: *Büyük bir mağazada vadesi geçmiş hesaplardan $n = 5$ birimlik bir örneklem alınmış ve veriler ₺ cinsinden aşağıdaki gibi elde edilmiştir:*

x_i: 18 28 41 55 78

5 gözlem değerinden oluşan bu örneklemin aritmetik ortalamasını bulunuz.

Çözüm: $n = 5$ olmak üzere örneklem ortalaması aşağıdaki gibi hesaplanır:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{\sum_{i=1}^{5} x_i}{5} = \frac{18 + 28 + \ldots + 78}{5} = \frac{220}{5} = 44$$

Vadesi geçmiş hesapların ortalaması ₺44 olarak bulunur.

Frekans Serilerinde Aritmetik Ortalama Hesabı

Bir anakütle üzerinde tanımlanmış *X* değişkeninin sayısal olarak aldığı değerler

x_i	f_i	$x_i f_i$
x_1	f_1	$x_1 f_1$
x_2	f_2	$x_2 f_2$
x_3	f_3	$x_3 f_3$
$\vdots$	$\vdots$	$\vdots$
x_k	f_k	$x_k f_k$
	$N = \sum_{i=1}^{k} f_i$	$\sum_{i=1}^{k} x_i f_i$

şeklinde bir frekans serisi olarak verilsin. Burada *N*, frekansların toplamını ifade etmektedir. Bu durumda, anakütle aritmetik ortalaması,

$$\mu = \frac{x_1 f_1 + x_2 f_2 + ... + x_k f_k}{N} = \frac{\sum_{i=1}^{k} x_i f_i}{N}$$

formülü ile hesaplanır ve μ (mü şeklinde okunur) ile gösterilir.

Bir anakütleden çekilmiş *n* birimlik bir örneklem için tanımlanmış *X* değişkenin sayısal olarak aldığı değerler

x_i	f_i	$x_i f_i$
x_1	f_1	$x_1 f_1$
x_2	f_2	$x_2 f_2$
x_3	f_3	$x_3 f_3$
$\vdots$	$\vdots$	$\vdots$
x_k	f_k	$x_k f_k$
	$n = \sum_{i=1}^{k} f_i$	$\sum_{i=1}^{k} x_i f_i$

şeklinde bir frekans serisi olarak verilsin. Örneklem aritmetik ortalaması ise

$n = \sum_{i=1}^{k} f_i$ olmak üzere aşağıdaki gibi bulunur:

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + ... + x_k f_k}{n} = \frac{\sum_{i=1}^{k} x_i f_i}{n}$$

Formüllerden de görüldüğü gibi, frekans serilerinde aritmetik ortalama hesaplanırken, gözlem değerleri ile bu gözlem değerlerine karşı gelen frekanslar çarpılıp toplanır ve elde edilen bu değer toplam frekansa bölünür.

Örnek 3: *Bir ampul fabrikasında rassal olarak seçilen 20 ampulun ömrü ölçülmüştür ve aşağıdaki tabloda verilmiştir.*

Ampul Ömrü (ay)	Ampul Sayısı
10	3
20	5
30	8
40	3
50	1

Bu fabrikanın ürettiği ampullerin ortalama ömrünü bulunuz.

Çözüm: Bir ampul fabrikasında rassal olarak seçilen 20 tane ampulun ortalama ömrü istendiğinden örneklem aritmetik ortalaması formülü kullanılmalıdır.

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + \ldots + x_k f_k}{n} = \frac{\sum_{i=1}^{k} x_i f_i}{n}$$

Aşağıdaki tabloda bu formülde gerekli olan tüm işlemler gösterilmiştir.

x_i	f_i	$x_i f_i$
10	3	30
20	5	100
30	8	240
40	3	120
50	1	50
Toplam	$n = 20$	$\sum_{i=1}^{k} x_i f_i = 540$

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i f_i}{n} = \frac{540}{20} = 27$$

Bu fabrikada üretilen ampullerin ortalama dayanma süresi 27 ay olarak bulunmuştur.

Gruplandırılmış Serilerde Aritmetik Ortalama Hesabı

Gruplandırılmış seriler için aritmetik ortalamanın formülü frekans serileri için verilen formüle benzerdir. Frekans serisinden farkı gözlem değerleri ayrı ayrı bilinmediğinden, diğer bir ifadeyle sadece ilgili gözlemin yer aldığı sınıf aralığı bilindiğinden, elde edilen ortalama, gerçek anakütle ya da örneklem aritmetik ortalamasına bir yaklaşım olacaktır.

k sayıda grup varken, gruplandırılmış seriler için anakütle aritmetik ortalaması

$N = \sum_{i=1}^{k} f_i$ olmak üzere;

$$\mu = \frac{m_1 f_1 + m_2 f_2 + ... + m_k f_k}{N} = \frac{\sum_{i=1}^{k} m_i f_i}{N}$$

şeklinde hesaplanır. Burada;

m_i= i. sınıf aralığının orta noktasını

f_i= i. sınıf aralığının frekansını

k= sınıf (grup) sayısını

gösterir.

Yine k tane grup için, gruplandırılmış serilerde örneklem aritmetik ortalaması $n = \sum_{i=1}^{k} f_i$ olmak üzere;

$$\bar{x} = \frac{m_1 f_1 + m_2 f_2 + ... + m_k f_k}{n} = \frac{\sum_{i=1}^{k} m_i f_i}{n}$$

olur.

Örnek 4: *Yüksek lisans yeni mezunlarından 40'ının aylık gelirleri aşağıda verilmiştir. Bu örneklem değerlerine göre ortalama aylık geliri belirleyiniz.*

Aylık Gelir (₺)	Mezun Sayısı
1000-2000'den az	6
2000-3000'den az	10
3000-4000'den az	15
4000-5000'den az	7
5000-6000'den az	2

Çözüm: Veriler gruplandırılmış seri şeklinde verildiğinden öncelikle sınıf orta noktaları bulunacak ardından ilgili frekanslarla çarpılacaktır. Sınıf orta noktaları bulunurken, sınıf alt sınırı ile üst sınırı toplanarak 2'ye bölünür. Aşağıdaki tabloda 3. sütun sınıf orta noktalarını göstermektedir.

Aylık Gelir (₺)	Mezun Sayısı f_i	Sınıf Orta Noktası m_i	$m_i f_i$
1000-2000'den az	6	1500	9000
2000-3000'den az	10	2500	25000
3000-4000'den az	15	3500	52500
4000-5000'den az	7	4500	31500
5000-6000'den az	2	5500	11000
Toplam	n= 40	-	$\sum_{i=1}^{k} m_i f_i = 129000$

$$\bar{x} = \frac{\sum_{i=1}^{k} m_i f_i}{n} = \frac{129000}{40} = 3225$$

Yüksek lisans mezunları aylık ortalama ₺3225 gelire sahiptirler.

Tartılı (Ağırlıklı) Aritmetik Ortalama

Bazen gözlem değerlerinin tümü aynı öneme sahip olmayıp temsil ettikleri değer bakımından farklılık gösterebilirler. Bu durumda, gözlem değerleri tartılandırılır, diğer bir ifadeyle ağırlık verilir. Gözlem değerleri bu ağırlıklarla çarpıldıktan sonra tartılı aritmetik ortalama hesaplanır.

Basit Serilerde Tartılı Aritmetik Ortalama Hesabı

Basit serilerde tartılı aritmetik ortalama hesaplamak için gözlem değerleri, aşağıdaki tabloda olduğu gibi, doğrudan ağırlık değerleri ile çarpılarak tartılandırılır.

x_i	t_i	$x_i t_i$
x_1	t_1	$x_1 t_1$
x_2	t_2	$x_2 t_2$
x_3	t_3	$x_3 t_3$
$\vdots$	$\vdots$	$\vdots$
x_k	t_k	$x_k t_k$
	$\sum_{i=1}^{k} t_i$	$\sum_{i=1}^{k} x_i t_i$

Anakütle için tartılı aritmetik ortalama aşağıdaki formülle hesaplanır:

$$\mu_t = \frac{\sum_{i=1}^{k} x_i t_i}{\sum_{i=1}^{k} t_i}$$

Burada, t_i ağırlıkları gösterir. Benzer şekilde örneklem için tartılı aritmetik ortalama

$$\bar{x}_t = \frac{\sum_{i=1}^{k} x_i t_i}{\sum_{i=1}^{k} t_i}$$

olarak hesaplanır.

Frekans Serilerinde Tartılı Aritmetik Ortalama Hesabı

Tartılı aritmetik ortalama hesabı için gözlem değerleri tartılarla beraber ayrıca frekanslarla çarpılır ve tartılarla frekansların çarpımlar toplamına oranlanır.

x_i	t_i	f_i	$t_i f_i$	$x_i t_i f_i$
x_1	t_1	f_1	$t_1 f_1$	$x_1 t_1 f_1$
x_2	t_2	f_2	$t_2 f_2$	$x_2 t_2 f_2$
x_3	t_3	f_3	$t_3 f_3$	$x_3 t_3 f_3$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_k	t_k	f_k	$t_k f_k$	$x_k t_k f_k$
		$N=\sum_{i=1}^{k} f_i$	$\sum_{i=1}^{k} t_i f_i$	$\sum_{i=1}^{k} x_i t_i f_i$

Anakütle için;

$$\mu_t = \frac{\sum_{i=1}^{k} x_i t_i f_i}{\sum_{i=1}^{k} t_i f_i}$$

formülü geçerlidir. Daha önce olduğu gibi frekansların toplamının, birim sayısı toplamına eşit olduğuna dikkat edilmelidir.

Örneklem için ise $n=\sum_{i=1}^{k} f_i$ olmak üzere veri düzeni aşağıdaki tabloda olduğu gibidir:

x_i	t_i	f_i	$t_i f_i$	$x_i t_i f_i$
x_1	t_1	f_1	$t_1 f_1$	$x_1 t_1 f_1$
x_2	t_2	f_2	$t_2 f_2$	$x_2 t_2 f_2$
x_3	t_3	f_3	$t_3 f_3$	$x_3 t_3 f_3$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_k	t_k	f_k	$t_k f_k$	$x_k t_k f_k$
		$n=\sum_{i=1}^{k} f_i$	$\sum_{i=1}^{k} t_i f_i$	$\sum_{i=1}^{k} x_i t_i f_i$

Örneklem için tartılı aritmetik ortalama

$$\bar{x}_t = \frac{\sum_{i=1}^{k} x_i t_i f_i}{\sum_{i=1}^{k} t_i f_i}$$

formülü ile hesaplanır.

Örnek 5: *Bir işletmede çalışanlar iş tecrübelerine göre ücretlendirilmektedirler. Aşağıdaki tabloda iş tecrübesine göre ücret verisi yer almaktadır. Bu veri setine göre ücretlerin tartılı aritmetik ortalamasını bulunuz.*

Günlük Ücret (₺)	İşçi sayısı	İş Tecrübesi (Yıl)
10	10	2
20	8	4
30	5	5
40	1	8

Çözüm: Söz konusu işletmede işçiler iş tecrübelerine göre ücretlendirildiklerinden, aldıkları ücretlerin doğrudan ortalaması hesaplanmayacak, bunun yerine iş tecrübelerine göre tartıları dikkate alınacaktır.

Günlük Ücret (₺) x_i	İşçi sayısı f_i	İş Tecrübesi (Yıl) t_i	$t_i f_i$	$x_i t_i f_i$
10	10	2	20	200
20	8	4	32	640
30	5	5	25	750
40	1	8	8	320
Toplam			$\sum_{i=1}^{k} t_i f_i = 85$	$\sum_{i=1}^{k} x_i t_i f_i = 1910$

$$\bar{x}_t = \frac{\sum_{i=1}^{k} x_i t_i f_i}{\sum_{i=1}^{k} t_i f_i} = \frac{1910}{85} = 22.47$$

İşletmede çalışan işçilerin iş tecrübelerine göre aldıkları günlük ücret ortalaması ₺22.47 olur.

Gruplandırılmış Serilerde Tartılı Aritmetik Ortalama Hesabı

Gruplandırılmış seriler için aritmetik ortalamanın hesabında olduğu gibi tartılı aritmetik ortalamanın hesabında da sınıf orta noktaları dikkate alınır. Buna göre; *k* sayıda grup varken, gruplandırılmış seriler için anakütle tartılı aritmetik ortalaması;

$$\mu_t = \frac{\sum_{i=1}^{k} m_i t_i f_i}{\sum_{i=1}^{k} t_i f_i}$$

ve benzer şekilde örneklem tartılı aritmetik ortalaması;

$$\bar{x}_t = \frac{\sum_{i=1}^{k} m_i t_i f_i}{\sum_{i=1}^{k} t_i f_i}$$

formülleri ile hesaplanır.

Aritmetik Ortalamanın Özellikleri

Aritmetik ortalamanın önemli matematiksel özellikleri vardır:

1. Gözlem sayısının (n) aritmetik ortalama ile çarpımı, gözlem değerlerinin toplamına eşittir;

$$\sum x_i = n \,.\, \bar{x}$$

Örnek 6: *Bir işletmede haftanın 5 farklı gününde yapılan kasa işlemlerinin sayısı not edilmiştir. Bu veri seti için aritmetik ortalamayı hesaplayarak gözlem sayısının aritmetik ortalama ile çarpımının gözlem değerleri toplamına eşit olduğunu gösteriniz.*

72 48 55 90 65

Çözüm: Gözlem değerlerinin toplamı:

$$\sum_{i=1}^{5} x_i = 48 + 55 + 65 + 72 + 90 = 330$$

Serinin aritmetik ortalaması $\bar{x}$ = 66 olduğundan;

$$n\bar{x} = (5).(66) = 330$$

bulunur. Gözlem sayısının aritmetik ortalama ile çarpımının gözlemler toplamına eşit olduğu görülmektedir.

2. Gözlem değerlerinin aritmetik ortalamadan sapmalarının (farklarının) cebirsel toplamı sıfıra eşittir;

$$\sum_{i=1}^{n} (x_i - \bar{x}) = 0$$

Örnek 7: *Örnek 6'daki verileri göz önünde bulundurarak gözlemlerin aritmetik ortalamadan sapmalarının toplamının sıfıra eşit olduğunu gösteriniz.*

Çözüm:

x_i	$x_i - \bar{x}$
48	-18
55	-11
65	-1
72	6
90	24
$\sum x_i = 330$	$\sum (x_i - \bar{x}) = 0$

Gözlem değerlerinin aritmetik ortalamadan sapmalarının cebirsel toplamı 0 olduğu görülmektedir.

3. Gözlem değerlerinin aritmetik ortalamadan sapmalarının karelerinin toplamı minimumdur. Bir başka ifade ile $\sum_{i=1}^{n}(x_i - \bar{x})^2$ minimum bir değerdir.

Örnek 8: *Örnek 6'da verilen veri seti için gözlemlerin aritmetik ortalamadan sapmalarının kareleri toplamının minimum olduğunu gösteriniz.*

Çözüm: Öncelikle gözlem değerlerinin aritmetik ortalama olan 66'dan sapmalarının kareleri toplamı hesaplanacak, ardından aritmetik ortalamadan daha küçük bir değer olan 60'dan ve daha büyük bir değer olan 70'ten sapmalarının kareleri toplamı hesaplanarak karşılaştırılacaktır.

	$(x_i - 66)^2$	$(x_i - 60)^2$	$(x_i - 70)^2$
	324	144	484
	121	25	225
	1	25	25
	36	144	4
	576	900	400
Toplam	1058	1238	1138

Gözlemlerin aritmetik ortalamadan olan sapmalarının kareleri toplamı 1058 iken 60'tan sapmalarının karelerinin toplamı 1238, 70'ten sapmalarının karelerinin toplamı 1138'dir. Gözlem değerlerinin aritmetik ortalamadan olan sapmalarının kareleri toplamının minimum bir değer olduğu görülmektedir.

4. Aritmetik ortalama serideki tüm gözlem değerleri dikkate alınarak hesaplandığından, diğer bir ifadeyle duyarlı ortalama olduğundan seride yer alan aşırı değerlerden etkilenir.

Örnek 9: *Aşağıdaki seri 5 farklı futbolcunun aylık kazançlarını vermektedir. Veri seti için aritmetik ortalamayı hesaplayarak yorumlayınız.*

x_i(1000EURO) : 10 11 15 20 100

Çözüm: Tüm gözlem değerlerini hesaba katarak aritmetik ortalamayı hesaplayalım.

$$\bar{x} = \frac{\sum_{i=1}^{5} x_i}{5} = \frac{10+11+15+20+100}{5} = \frac{156}{5} = 31.2$$

Futbolcuların aylık kazançlarının ortalaması 31200 Euro olarak hesaplanmıştır. Veri setinde 100 değerinin olmadığını varsayarak ortalamayı tekrar hesaplayalım:

$$\bar{x} = \frac{\sum_{i=1}^{4} x_i}{4} = \frac{10+11+15+20}{4} = \frac{56}{4} = 14$$

Sadece bir gözlem değeri örneklemden çıkarılmasına rağmen aritmetik ortalama 31.2'den 14'e düşmüştür. 100 değeri seri için bir uç değer olup ortalamayı etkilemiştir. Bu gibi durumlarda, hesaplamalarında tüm gözlem değerlerinin dikkate alınmadığı duyarlı olmayan ortalamalar tercih edilmelidir.

Kareli Ortalama

Seriyi oluşturan gözlem değerlerinin karelerinin toplamının gözlem sayısına oranının karekökü kareli ortalamayı verir. Bir başka ifadeyle gözlem değerlerinin karelerinin aritmetik ortalamasının karekökü kareli ortalama olarak adlandırılır.

Her zaman aritmetik ortalama < kareli ortalama eşitsizliği geçerlidir.

DİKKAT

Basit Serilerde Kareli Ortalama Hesabı

Basit serilerde kareli ortalama anakütle için

$$K = \sqrt{\frac{x_1^2 + x_2^2 + \cdots + x_N^2}{N}} = \sqrt{\frac{\sum_{i=1}^{N} x_i^2}{N}}$$

formülüyle hesaplanırken örneklem için aşağıdaki formül kullanılır:

$$K = \sqrt{\frac{x_1^2 + x_2^2 + \cdots + x_n^2}{n}} = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}}$$

Örnek 10: *Aşağıdaki örneklem serisinin kareli ortalamasını hesaplayınız.*

x_i: 3 6 10 15 22

Çözüm:

x_i	x_i^2
3	9
6	36
10	100
15	225
22	484
	$\sum x_i^2 = 854$

$$K = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}} = \sqrt{\frac{854}{5}} = 13.07$$

Serinin kareli ortalaması 13.07'ye eşittir.

Aynı serinin aritmetik ortalaması $\bar{x}$ =11.2'dir. Görüldüğü gibi kareli ortalamanın değeri aritmetik ortalamadan her zaman daha büyük olur.

Frekans Serilerinde Kareli Ortalama Hesabı

Frekans serilerinde anakütle için kareli ortalama aşağıdaki formülle hesaplanır:

$$K = \sqrt{\frac{\sum x_i^2 f_i}{N}}$$

Örneklem için ise

$$K = \sqrt{\frac{\sum x_i^2 f_i}{n}}$$

formülü geçerlidir.

Örnek 11: *Aşağıdaki tabloda bir büyük şehirde seçilen 100 ailenin yaşadıkları evlerin oda sayıları verilmiştir. Ailelerin yaşadıkları evlerin oda sayılarının kareli ortalamasını bulunuz.*

Oda sayısı	Aile sayısı
1	14
2	24
3	41
4	18
5	3

Çözüm:

x_i	f_i	x_i^2	$x_i^2 f_i$
1	14	1	14
2	24	4	96
3	41	9	369
4	18	16	288
5	3	25	75
Toplam	100		842

$$K = \sqrt{\frac{\sum x_i^2 f_i}{n}} = \sqrt{\frac{842}{100}} = 2.9$$

Ailelerin oturdukları evlerin oda sayılarının kareli ortalaması 2.9’dur.

Gruplandırılmış Serilerde Kareli Ortalama Hesabı

Daha önceki bölümlerde olduğu gibi, gruplandırılmış seriler için sınıf orta noktalarını dikkate almak üzere, anakütle için kareli ortalama;

$$K = \sqrt{\frac{\sum m_i^2 f_i}{N}}$$

formülüyle hesaplanırken örneklem için kareli ortalama;

$$K = \sqrt{\frac{\sum m_i^2 f_i}{n}}$$

formülüyle hesaplanır.

Örnek 12: *Aşağıdaki seriyi örneklem serisi kabul ederek kareli ortalamayı hesaplayınız.*

Sınıflar	f_i
0-10	2
10-20	5
20-30	10
30-40	8
40-50	5
	n= 30

Çözüm:

Sınıflar	f_i	m_i	m_i^2	$m_i^2 f_i$
0-10	2	5	25	50
10-20	5	15	225	1125
20-30	10	25	625	6250
30-40	8	35	1225	9800
40-50	5	45	2025	10125
Toplam	n=30		$\sum_{i=1}^{k} m_i^2 f_i = 27350$	

$$K = \sqrt{\frac{\sum m_i^2 f_i}{n}} = \sqrt{\frac{27350}{30}} = \sqrt{911.67} = 30.19$$

Serinin kareli ortalaması 30.19'dur.

Medyan (Ortanca)

Gözlem değerleri küçükten büyüğe sıralandığında ortada kalan gözlem değeri **medyan**dır ve tanımından da anlaşılacağı gibi bir seride yer alan gözlemlerin tümünün hesaba katılmadığı ortalamalardan biridir.

Basit serilerde seri tek sayıda gözlemden oluşuyorsa serinin gözlem değerleri küçükten büyüğe sıralandığında tam ortada yer alan gözlem değeri medyandır. Seri çift sayıda gözlemden oluşuyorsa ortada kalan iki gözlem değerinin aritmetik ortalaması medyanı verir.

Medyan, ölçümlerin %50'sinin üzerinde, %50'sinin aşağısında yer aldığı merkezi değerdir.

Örnek 13: *Bir büyük alışveriş merkezi hafta sonları kasalarda müşterilerinin bekleme zamanlarını azaltmak için çalışmaktadır. Bu amaçla yoğun gün ve saatlerde rassal olarak seçilen 5 müşterinin kasalarda bekleme zamanı dakika olarak kaydedilmiştir. Müşterilerin bekleme zamanlarının medyanını bulunuz.*

15 16 11 12 6

Çözüm: Öncelikle veriler küçükten büyüğe sıralanmalıdır. Buna göre:

x_i : 6 11 12 15 16

olur. Veri sayısı tek olduğundan (5 gözlem) ortada kalan gözlemin doğrudan medyan olduğu açıktır. Buna göre müşterilerin kasalarda bekleme zamanlarının medyanı 12 dakikadır.

DİKKAT

Seride uç değerler bulunduğunda duyarlı olmayan ortalamalar tercih edilmelidir.

Örnek 14: *Bir özel bankanın son 10 yılda işe aldığı iktisat ve işletme mezunlarının sayısı aşağıdaki gibidir:*

86 78 90 62 73 89 92 84 95 76

Banka tarafından istihdam edilen işletmeci ve iktisatçıların sayılarının medyanını hesaplayınız.

Çözüm: Öncelikle istihdam edilen işletmeci ve iktisatçıların sayılarını büyüklük sırasına göre sıralamalıyız.

62 73 76 78 84 86 89 90 92 95

Çift sayıda gözlem olduğundan, medyan ortadaki iki gözlem değerinin aritmetik ortalaması olacaktır.

$$Medyan = \frac{84+86}{2} = 85$$

Bankanın işe aldığı işletmeci ve iktisatçıların sayılarının medyanı 85'e eşittir.

DİKKAT

Bir veri seti için sadece bir tane medyan değeri vardır.

Frekans serilerinde medyan hesaplanırken, sıralanmış verilerin birikimli frekansları dikkate alınır ve ortada yer alan gözlem değeri belirlenir.

Örnek 15: *Bir işletmede çalışan 100 işçinin kullandıkları yıllık izinler aşağıda verilmiştir. Kullanılan yıllık izinlere ilişkin medyan değerini bulunuz.*

Kullanılan izin (gün)	İşçi Sayısı
10	18
15	32
20	30
25	20

Çözüm: Veriler sıralanmış gözlem değerleri olduğundan, doğrudan birikimli frekanslar hesaplanmalıdır.

Kullanılan izin (gün) x_i	İşçi Sayısı f_i	Birikimli frekans
10	18	18
15	32	50
20	30	80
25	20	100

Veri seti 100 birimden oluştuğundan $n/2$=50. gözlem dikkate alınır. 50. gözlem değeri 15 iken, 51. gözleme karşı gelen değer 20'dir. Medyan değerini hesaplamak için 50. gözlem ile 51. gözlemin aritmetik ortalaması hesaplanmalıdır. Buna göre medyan (15+20)/2=17.5 değeri bulunur.

Gruplandırılmış seriler için medyan hesabı biraz daha zordur. Gözlemlerin gerçek değerleri bilinmediğinden, medyanın belirli bir sınıf aralığında ortaya çıktığı bilinir, ancak medyanın bu aralıkta nerede olduğu bilinmez. Gözlemlerin aralık içinde düzgün bir dağılıma sahip olduğunu varsayarak, medyan aşağıdaki gibi belirlenir.

$$Medyan = L + \frac{h_m}{f_m}(0.5n - f_a)$$

L = medyanı içeren sınıf aralığının alt sınırı
n = toplam frekans
f_a = medyan sınıfından önceki tüm sınıfların frekansları toplamı
f_m = medyanı içeren sınıfın frekansı
h_m = medyan sınıfının büyüklüğü olsun.

Örnek 16: *Aşağıdaki tabloda yer alan veriler için medyanı hesaplayınız.*

Sınıf Aralığı	f_i	Birikimli frekans
10-20'den az	23	23
20-30'dan az	32	55
30-40'dan az	35	90
40-50'den az	10	100
Toplam	n= 100	

Çözüm: Birikimli frekansı $n/2$=50. gözleme karşı gelen sınıf medyanı içeren sınıf olacaktır. Örnekteki veriler için Tablonun 3. sütununda görüldüğü gibi birikimli frekansı 50'ye karşı gelen sınıf, 20-30 aralığıdır. Buna göre bu aralık medyanı içermektedir.

L=20 n=100 f_a=23 f_m=32 h_m=10
ve

$$Medyan = L + \frac{h_m}{f_m}(0.5n - f_a) = 20 + \frac{10}{32}(0.5\,.\,100 - 23) = 28.44$$

Gruplandırılmış serinin medyanı 28.44 olarak hesaplanmıştır.

Medyan aşırı değerlerden etkilenmez. DİKKAT

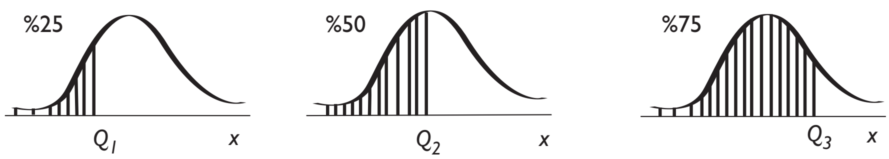
Kartiller (Dördebölenler)

Kartiller medyandan önceki ve sonraki grubun yarısını işaret eder. Diğer bir ifadeyle kartiller, serideki gözlemlerin %25'inin hangi değerden önce geldiğini, yine %75'inin hangi değerden önce geldiğini belirtir. İlk kartil Q_1, üçüncü kartil Q_3 ile gösterilir. Daha önce de tanımlandığı gibi gözlemlerin %50'sinin hangi değerden önce geldiğini gösteren medyan ise ikinci kartil kabul edilir. Şekil 2.1'de birinci kartil, medyan ve üçüncü kartilin gözlemlerin hangi noktasında kaldığı gösterilmektedir.

Q_1 medyana kadar olan verilerin medyanı olarak düşünülebilir. Aynı şekilde Q_3 medyandan sonraki verilerin medyanıdır. Buna göre Q_1 (birinci kartil), Q_2 (medyan), Q_3 (üçüncü kartil) Şekil 2.1'de verildiği gibi sıralanır.

Şekil 2.1

Birinci, ikinci ve üçüncü kartilin grafik üzerinde gösterimi



Örnek 17: *Aşağıdaki 11 gözlemli basit seri için medyanı, birinci ve üçüncü kartili hesaplayınız.*

30 45 18 32 52 98 65 77 25 44 32

Çözüm: Öncelikle veriler küçükten büyüğe sıralanmalı:

18	25	30	32	32	44	45	52	65	77	98
		Q_1			*Medyan*			Q_3		

Sıralanmış gözlem değerleri için 6. gözlem ortada yer alan gözlem olduğundan medyan değeri 44 olarak belirlenir. İlk 5 gözlemin medyanı, 3. gözlem, birinci kartili verecektir. Bu değer 30'a eşittir. Serinin son 5 gözleminin medyanı olan 65 ise 3. kartile karşı gelmektedir.

Gruplandırılmış seriler için birinci ve üçüncü kartilin formülü, medyanın formülüne benzemektedir:

$$Q_1 = L + \frac{h_m}{f_m}(0.25n - f_a)$$

$$Q_3 = L + \frac{h_m}{f_m}(0.75n - f_a)$$

Örnek 18: *Örnek 16'da yer alan veri seti için birinci ve üçüncü kartili hesaplayınız.*

Sınıf Aralığı	f_i	Birikimli frekans
10-20'den az	23	23
20-30'dan az	32	55
30-40'dan az	35	90
40-50'den az	10	100
Toplam	n= 100	

Çözüm:

Veriler küçükten büyüğe sıralandığında 25. gözlem 20-30 sınıfına düşmektedir. Buna göre birinci kartil için formülde yer alacak değerler aşağıdaki gibi olacaktır:

L=20 n=100 f_a=23 f_m=32 h_m=10

$$Q_1 = L + \frac{h_m}{f_m}(0.25n - f_a) = 20 + \frac{10}{32}(0.25\,.\,100 - 23) = 20.63$$

Benzer şekilde 75. gözlem 30-40 aralığında yer almaktadır. Buna göre,

L=30 n=100 f_a=55 f_m=35 h_m=10

$$Q_3 = L + \frac{h_m}{f_m}(0.75\text{n} - f_a) = 30 + \frac{10}{35}(0.75\,.\,100 - 55) = 35.71$$

Sonuçlara göre serinin birinci kartili Q_1=20.63 ve üçüncü kartili Q_3=35.71'dir.

Mod

Mod en sık ortaya çıkan (en yüksek frekanslı) ölçüm olarak tanımlanmaktadır.

Mod aşırı değerlerden etkilenmez.

Örnek 19: *Aşağıda Eskişehir'de okuyan 24 öğrencinin öğrenim yılının ilk ayında yaptıkları harcamalar verilmiştir. Veri setinin modunu belirleyiniz.*

965	1005	1030	1005	1042	975	975	955
975	965	1030	1030	955	965	1005	1015
1000	975	995	1042	1042	1000	975	1015

Çözüm: Verilerde bazı gözlemlerin birkaç kez tekrarlandığı görülmektedir. Bununla birlikte hangi gözlemin kaç kez gözlemlendiğinin daha net incelenebilmesi için frekans serisi oluşturmakta fayda vardır.

x_i	f_i
955	2
965	3
975	6
1000	2
1005	3
1015	2
1030	3
1042	3

Seri basit seriden frekans serisine dönüştürüldüğünde en çok tekrarlanan gözlem değerini belirlemek kolaylaşmıştır. Buna göre mod değeri 975'tir.

Bir veri seti için birden fazla mod olabilir.

DİKKAT

Gruplandırılmış Serilerde Mod

Örnek 19 için modun belirlenmesi oldukça kolaydır, çünkü her bir ölçümün kaç kez ortaya çıktığını sayabildik. Gruplandırılmış veriler üzerinde çalışırken en yüksek frekansa sahip sınıf aralığını belirleyip mod için geçerli olan aşağıdaki formülü kullanırız.

$$Mod = L + \frac{\Delta_1}{\Delta_1 + \Delta_2}.h$$

L : mod sınıfının alt sınırı

Δ_1 : mod sınıfının frekansı ile ondan bir önceki sınıfın frekansları arasındaki fark

Δ_2 : mod sınıfının frekansı ile ondan bir sonraki sınıfın frekansları arasındaki fark
h : sınıf aralığı
olmak üzere gruplandırılmış seriler için mod aşağıdaki gibi hesaplanır:

SIRA SİZDE 1

Merkezî eğilim ölçülerini sayınız.

Örnek 20: *Aşağıdaki seri için modu hesaplayınız.*

Sınıflar	f_i
0-8'den az	12
8-16'dan az	23
16-24'ten az	32
24-32'den az	20
32-+	13
	100

Çözüm: Mod sınıfı en yüksek frekansa sahip olan 16-24 sınıfıdır. Buna göre,
$\Delta_1 = 32 - 23 = 9$
$\Delta_2 = 32 - 20 = 12$

$$Mod = L + \frac{\Delta_1}{\Delta_1 + \Delta_2}.h = 16 + \frac{9}{9+12}.8 = 19.43$$

Serinin modu 19,43'e eşittir.

Mod yaygın olarak merkezî eğilim ya da görüşü belirleyen popülerliğin bir ölçüsü olarak kullanılır. Örneğin, en çok tercih edilen hisse senedi veya en popüler aday hakkında konuşabiliriz. Her bir durumda dağılımın modunu ifade etmiş oluruz.

Dikkat edilmesi gereken bir nokta, bazı dağılımlarda en yüksek frekansla gözlenen birden fazla ölçümün olabileceğidir. Böylece bir modlu, iki modlu, v.b. çok modlu dağılımlarla karşılaşabiliriz.

DİKKAT

Mod hem nitel veriler hem de nicel veriler için hesaplanabilir

Örnek 21: *Aşağıda 13 işçi için saat başına verimliliklerine göre ücret verileri yer almaktadır.*

4.4 4.9 4.2 4.4 4.8 4.9 4.8 4.5 4.3 4.8 4.7 4.4 4.2

Bu veri seti için modu belirleyiniz.

Çözüm: Öncelikle gözlemleri küçükten büyüğe doğru sıralamalıyız:

4.2 4.2 4.3 4.4 4.4 4.4 4.5 4.7 4.8 4.8 4.8 4.9 4.9

Bu veri seti için üçer kez gözlemlenen iki değer vardır. Buna göre 4.4 ve 4.8 olmak üzere iki modlu bir veri seti söz konusudur.

SIRA SİZDE 2

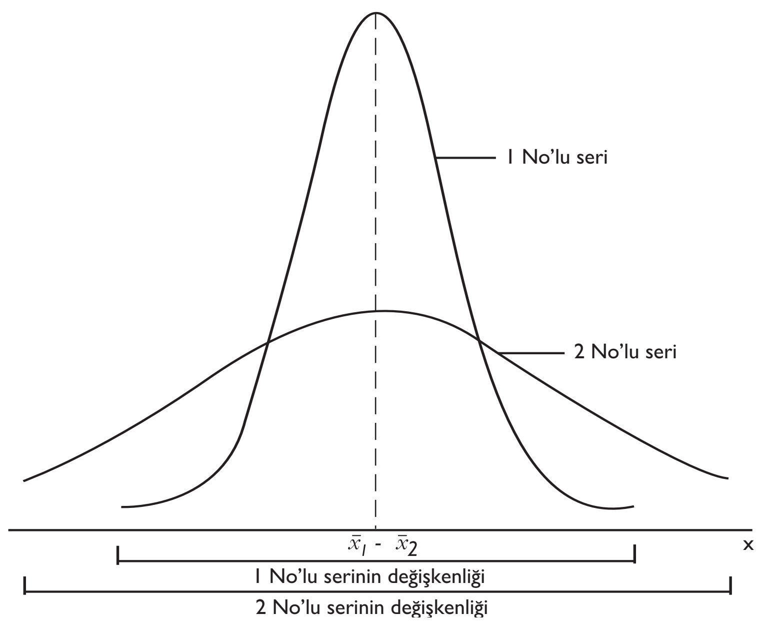
Aşağıda 10 üniversite mezununun aylık kazançları yer almaktadır. Uygun ortalamayı belirleyip hesaplayınız.

x_i: 1050 1200 1225 1250 1300 1500 1900 2400 2500 15000

DEĞİŞKENLİK ÖLÇÜLERİ

Merkezî eğilim ölçüleri bir ölçümler setinin frekans dağılımının tam bir resmini ortaya koymaz. Şekil 2.2'de $\bar{x}_1=\bar{x}_2$ olmakla birlikte, 2 nolu serinin değişkenliği 1 nolu serininkinden çok daha fazladır. Buradan da anlaşılacağı gibi, ortalamaları aynı olan iki dağılımın değişkenliği birbirinden çok farklı olabilir. Bu nedenle dağılımlar sadece merkezî eğilim ölçüleriyle ifade edilmeyip buna ilave olarak verilerin yayılımının bir ölçüsü de geliştirilmelidir.

Şekil 2.2



Ortalamaları aynı, değişkenlikleri farklı iki serinin dağılımının grafiği

Bu bölümde üzerinde durulacak merkezî değişkenlik ölçüleri değişim aralığı, standart sapma, varyans ve değişim katsayısıdır.

Değişim Aralığı

Bir gözlemler setinin değişim aralığı, bu serinin en büyük gözlemi ile en küçük gözlemi arasındaki farkıdır:

$$D.A.= x_{\max} - x_{\min}$$

Örnek 22: *Bir işletme belirli bir hammadde için 5 farklı tedarikçi firma ile görüşmüş ve her birinden fiyat almıştır. Tedarikçi firmaların ₺ cinsinden verdiği fiyatların değişim aralığını belirleyiniz.*

x_i:	18	22	25	26	29

Çözüm:

$D.A.= x_{\max} - x_{\min} = 29 - 18 = 11$

Tedarikçi firmaların verdiği fiyatlar için değişim aralığı ₺11 olarak bulunur.

Gruplandırılmış veriler için, ölçümleri ayrı ayrı bilemediğimizden, değişim aralığı son sınıf aralığının üst limiti ile ilk sınıf aralığının alt limitinin farkı olarak hesaplanır.

Örnek 23: *Aşağıdaki tabloda bir işletmede çalışan işçilerin gün olarak iş tecrübeleri verilmiştir. Bu veri seti için değişim aralığını belirleyiniz.*

Sınıflar	İşçi sayısı
500 – 600'den az	3
600 – 700'den az	8
700 – 800'den az	15
800 – 900'den az	30
900 – 1000'den az	58

Çözüm:

$D.A. = x_{\max} - x_{\min} = 1000 - 500 = 500$

İşçilerin iş tecrübesine ilişkin serinin değişim aralığı 500'dür.

Değişim aralığı hesaplanması kolay olmasına karşın, en uç değerlere bağlı olduğundan, aykırı değerlere karşı hassastır. Ayrıca serinin sadece 2 gözlemine bağlı olarak hesaplanan bu ölçü değişkenliğin şekli hakkında çok fazla bilgi vermez. Bu nedenle diğer değişkenlik ölçüleri kadar sık kullanılmaz.

Standart Sapma ve Varyans

Gözlem değerlerinin aritmetik ortalamadan sapmaları dikkate alınarak farklı değişkenlik ölçüleri geliştirilebilir. Bu noktada ilk akla gelen ortalama sapmadır, ancak gözlemlerin aritmetik ortalamadan sapmalarının her zaman sıfıra eşit olduğu daha önce verilmişti. Bu sorunu ortadan kaldırmak için gözlemlerin aritmetik ortalamadan olan sapmalarının karelerinin toplamının gözlem sayısına oranı değişkenlik ölçüsü olarak yorumlanabilir. Bu ölçü varyans olarak adlandırılır. Varyansın karekökü standart sapmadır.

Basit Serilerde Varyans ve Standart Sapma Hesabı

Anakütle için varyans aşağıdaki formülle hesaplanır:

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}$$

Anakütle standart sapması σ (sigma şeklinde okunur) anakütle varyansının karekökü ile verilir:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}}$$

Örnek 24: *8 öğrenciden oluşan bir grup lise 1 öğrencisi yabancı dil eğitimi için yurt dışına gönderilmiş, döndüklerinde sınava tabi tutulmuşlardır. Aldıkları pu-*

anlar aşağıda verilmiştir. Öğrencilerin aldıkları puanlara ilişkin varyansı ve standart sapmayı hesaplayınız.

x_i : 55 62 68 72 75 80 83 85

Çözüm:

x_i	$(x_i-\mu)$	$(x_i-\mu)^2$
55	-17.5	306.25
62	-10.5	110.25
68	-4.5	20.25
72	-0.5	0.25
75	2.5	6.25
80	7.5	56.25
83	10.5	110.25
85	12.5	156.25
$\sum x_{i=}$ 580		$\sum (x_i - \mu)$= 766

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N} = \frac{766}{8} = 95.75 \qquad (\mu=72.5)$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}} = \sqrt{95.75} = 9.79$$

Örneklem varyansı s^2, standart sapması s ile gösterilir ve aşağıdaki formüllerle hesaplanır:

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

Frekans Serilerinde Varyans ve Standart Sapma Hesabı

Frekans serilerinde anakütle için varyans ve standart sapma formülleri aşağıdaki gibidir:

$$\sigma^2 = \frac{\sum_{i=1}^{N} f_i (m_i - \mu)^2}{N}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N} f_i (m_i - \mu)^2}{N}}$$

Frekans serilerinde örneklem için varyans ve standart sapma formülleri ise aşağıdaki gibi olacaktır:

$$s^2 = \frac{\sum_{i=1}^{n} f_i \left(m_i - \bar{x}\right)^2}{n-1}$$

$$s = \sqrt{\frac{\sum_{i=1}^{n} f_i \left(m_i - \bar{x}\right)^2}{n-1}}$$

Örnek 25: *Aşağıdaki tabloda 5 farklı mühendislik bölümünden rassal olarak seçilen öğrenciler için mezuniyet puanı ortalaması ve mezun olan öğrenci sayıları verilmiştir. Buna göre mühendislikten mezun olan öğrencilerin mezuniyet puanlarının standart sapması nedir?*

Bölümler	Mezuniyet Puanı Ortalaması	Mezun Olan Öğrenci Sayısı
Endüstri	2.88	58
Elektrik-Elektronik	2.76	32
Bilgisayar	2.82	30
Kimya	2.70	46
Çevre	2.80	45

Çözüm:

Bölümler	Mezuniyet Puanı Ortalaması x_i	Mezun Öğrenci Sayısı f_i	$f_i x_i$	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$f_i(x_i - \bar{x})^2$
Endüstri	2.88	58	167.04	0.08	0.0064	0.3712
Elektrik-Elektronik	2.76	32	88.32	-0.04	0.0016	0.0512
Bilgisayar	2.82	30	84.6	0.02	0.0004	0.012
Kimya	2.70	46	124.2	-0.1	0.01	0.46
Çevre	2.80	45	126	0	0	0
Toplam		n= 211	590.16			0.8944

$$\bar{x}_t = \frac{\sum_{i=1}^{k} f_i x_i}{n} = \frac{590.16}{211} = 2.797 \approx 2.8$$

$$s = \sqrt{\frac{\sum_{i=1}^{n} f_i \left(x_i - \bar{x}\right)^2}{n-1}} = \sqrt{\frac{0.8944}{210}} = 0.065$$

Mühendislik Fakültesi mezunlarının mezuniyet puanı standart sapması 0.065'tir.

Gruplandırılmış Serilerde Varyans ve Standart Sapma Hesabı

Gruplandırılmış seriler için işlem yaparken verilen formüller gözlem değeri x_i'nin yerine sınıf orta noktaları yazılacaktır. Buna göre anakütle varyansı ve standart sapması için formüller

$$\sigma^2 = \frac{\sum_{i=1}^{N} f_i \left(m_i - \mu\right)^2}{N}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N} f_i \left(m_i - \mu\right)^2}{N}}$$

şeklinde olurken örneklem için formüller aşağıdaki gibi tanımlanacaktır.

$$s^2 = \frac{\sum_{i=1}^{n} f_i \left(m_i - \bar{x}\right)^2}{n-1}$$

$$s = \sqrt{\frac{\sum_{i=1}^{n} f_i \left(m_i - \bar{x}\right)^2}{n-1}}$$

Örnek 26: *Bir havaalanına gelen uçaklardan 100 birimlik bir örneklem seçilerek hava ulaşımındaki gecikmeler incelenmiştir. Uçakların gecikme süreleri aşağıdaki tabloda verilmiştir. Uçuşlar için gecikme sürelerine ilişkin varyansı ve standart sapmayı hesaplayınız.*

Uçakların gecikme süresi (dk.)	Uçak sayısı
0-10	29
10-20	23
20-30	17
30-40	14
40-50	11
50-60	6

Uçakların Gecikme süresi (dk.)	Uçak sayısı f_i	Sınıf Orta Değeri m_i	$m_i f_i$	$(m_i - \bar{x})$	$(m_i - \bar{x})^2$	$f_i(m_i - \bar{x})^2$
0-10	29	5	145	-17.3	299.29	8679.41
10-20	23	15	345	-7.3	53.29	1225.67
20-30	17	25	425	2.7	7.29	123.93
30-40	14	35	490	12.7	161.29	2258.06
40-50	11	45	495	22.7	515.29	5668.19
50-60	6	55	330	32.7	1069.29	6415.74
Toplam	n= 100		2230			24371

Çözüm: Gözlemlerin aritmetik ortalamadan olan sapmalarının kareleri toplamını hesaplayacağımızdan, öncelikle aritmetik ortalama bulunacaktır.

$$\bar{x} = \frac{\sum_{i=1}^{k} m_i f_i}{\text{n}} = \frac{2230}{100} = 22.30$$

$$s^2 = \frac{\sum_{i=1}^{n} f_i \left(m_i - \bar{x}\right)^2}{n-1} = \frac{24371}{99} = 246.17$$

$$s = \sqrt{\frac{\sum_{i=1}^{n} f_i \left(m_i - \bar{x}\right)^2}{n-1}} = \sqrt{246.17} = 15.69$$

Uçakların gecikme süresinin aritmetik ortalaması 22.30 dk., varyansı 246.17 dk. ve standart sapması 15.69 dk. olarak bulunmuştur.

Çok az değişkenliğe sahip bir veri setinin gözlemlerinin çoğu, dağılımın merkezine yakın olacaktır. Daha değişken bir veri seti için aritmetik ortalamadan olan sapmalar nispeten daha büyük olacaktır.

En yaygın olarak kullanılan değişiklik ölçüsü varyans, asağıdaki formüllerle de hesaplanabilir;

$$\sigma^2 = \frac{\sum x_i^2}{N} - \mu^2$$

Frekans serileri için ise formül aşağıdaki gibi yazılabilir:

$$\sigma^2 = \frac{\sum x_i^2 f_i}{N} - \mu^2$$

Örnek 27: *Kareli ortalaması 12, artitmetik ortalaması 8 olduğu bilinen bir anakütlenin varyansı nedir?*

Çözüm:

$$\sigma^2 = \frac{\sum \text{x}_\text{i}^2}{\text{N}} - \mu^2 = 144 - 64 = 80$$

Anakütlenin varyansı 80'e eşittir.

SIRA SİZDE 3

Kareli ortalaması 11 aritmetik ortalaması 10 olan anakütlenin standart sapması kaçtır?

Değişim Katsayısı

Farklı serilerin değişkenliklerinin karşılaştırılmasında, farklı birimlerle ölçülmüş veri setleri söz konusu olduğundan standart sapma kullanışlı değildir. Bunun yerine ilgili serilerin standart sapmaları serilerin ortalama değerinin yüzdesi olarak ifade edilir ve gözlem değerlerinin büyüklüklerinden kaynaklanan farklılık ortadan kalkmış olur. Elde edilen bu yeni değişkenlik ölçüsü kullanılarak serilerin birbirlerine göre daha değişken ya da daha homojen oldukları konusunda yorum yapılabilir.

Anakütle için: $\text{D.K.} = \frac{\sigma}{\mu}.100$ Örneklem için: $\text{D.K.} = \frac{s}{\bar{x}}.100$

Örnek 28: *Aşağıda verilen 2 serinin değişkenliğini değişim katsayısı ile karşılaştırınız.*

x_i	y_i
4	3
7	7
11	15
15	16
18	20

Çözüm:

$\bar{x} = 11$

$s_x = 5.7$

$\bar{y} = 12.2$

$s_y = 6.98$

$$\text{D.K.}(x) = \frac{s_x}{\bar{x}}.100 = \frac{5.7}{11}.100 = \%51.82$$

$$\text{D.K.}(y) = \frac{s_y}{\bar{y}}.100 = \frac{6.98}{12.2}.100 = \%57.21$$

x serisinin değişkenliği %51.82, y serisinin değişkenliği %57.21 olarak hesaplanmıştır. Buna göre y serisinin değişkenliği daha fazladır.

Değişim katsayısına göre aşağıdaki serilerden hangisinin değişkenliği diğerine göre daha fazladır?

SIRA SİZDE 4

x_i	y_i
10	8
15	15
18	20
22	22
25	26

Özet

AMAÇ 1

Merkezî eğilim ve değişkenlik ölçülerini saymak, hangi durumda hangi ölçüyü hesaplamanız gerektiğine karar vermek.

Bir veri setinde yer alan değerleri tek bir sayı ile özetlemek, temsil etmek ve bu yolla yorumlamak için hesaplanan ölçüler merkezî eğilim ölçüleridir.

Merkezî eğilim ölçüleri arasında en yaygın olanı, serinin gözlem değerlerinin toplanıp birim sayısına oranlanmasıyla hesaplanan aritmetik ortalamadır. Kareli ortalama, aritmetik ortalama gibi hesaplanmasında tüm birimlerin dikkate alındığı bir diğer ortalamadır.

Medyan, kartiller ve mod serinin tüm gözlemlerinin hesaplamaya katılmadığı merkezî eğilim ölçüleridir. Medyan serinin orta noktasını, birinci kartil serinin birinci dördebölenin sınırını, üçüncü kartil serinin üçüncü dördebölenin sınırını gösterirken mod en çok tekrarlanan gözlem değeridir.

Dağılımın şeklinden söz edebilmek için merkezi eğilim öçüleri yetersiz kaldığından, aynı zamanda değişkenlik ölçülerine de başvurulur. Standart sapma, standart sapmanın karesi olan varyans sık kullanılan değişkenlik ölçüleridir. Diğer değişkenlik ölçüleri arasında değişim aralığı ve değişim katsayısı yer alır.

AMAÇ 2

Herhangi bir veri seti için uygun ortalamayı hesaplayıp yorumlamak.

Merkezî eğilim ölçüleri arasında en yaygın olarak kullanılanı aritmetik ortalamadır. Sadece ortalama denildiğinde aritmetik ortalama anlaşılır. Matematiksel özellikleri ve seriyi güzel temsil etmesi yönüyle tercih edilir.

Varyansın pratik hesabında kullanılan, hesaplanmasında gözlem değerlerinin kareleri dikkate alındığından her zaman pozitif sonuç veren ortalama, kareli ortalamadır.

Serinin uç değerler alması durumunda tüm gözlemlerin hesaba katıldığı duyarlı ortalamalar yerine, medyan, kartiller, mod gibi duyarlı olmayan ortalamalar tercih edilmelidir.

AMAÇ 3

Herhangi bir veri seti için uygun değişkenlik ölçüsünü hesaplayıp yorumlamak.

İlk akla gelen değişkenlik ölçüleri standart sapma ve standart sapmanın karesi olan varyanstır. Bu değişkenlik ölçüleri dağılımın değişkenliğiyle ilgili özet bilgiyi verirler.

Duyarlı olmayan ortalamalarda olduğu gibi serinin tüm gözlemlerinin değil sadece minimum ve maksimumunun dikkate alındığı değişkenlik ölçüsü değişim aralığıdır.

AMAÇ 4

Serilerin birbirlerine göre değişkenliğini kıyaslayabilmek

Pratikte incelenen örnek olaylar için birimlerden bağımsız olarak hesaplanan değişim katsayısı yardımıyla serilerin birbirlerine göre daha az değişken (homojen) ya da daha çok değişken (heterojen) oldukları konusunda yorum yapılabilir.

Kendimizi Sınayalım

1. Anadolu Üniversitesi Fen Fakültesini kazanıp hazırlık okuluna devam eden öğrencilerin bölümlerine göre hazırlık okulu yıl sonu başarı puanı ortalamaları aşağıdaki gibidir. Fen Fakültesi öğrencilerinin hazırlık okulu yıl sonu başarı ortalaması nedir?

Bölümler	Bölümlerin Başarı Ortalaması	Bölümlerdeki Öğrenci Sayıları
Biyoloji	62	60
Fizik	60	56
İstatistik	68	85
Kimya	65	66
Matematik	70	88
Toplam		355

a. 50.50
b. 55.25
c. 58.55
d. 60.55
e. 65.66

2. Bir zar 40 kez atılmış ve aşağıdaki sonuçlar elde edilmiştir.

Değer	Frekans
1	9
2	8
3	5
4	5
5	6
6	7

Örneklem aritmetik ortalaması, medyan ve modu hesaplayınız.

a. $\bar{x}$=4, *Medyan*=4.2, *Mod*=1
b. $\bar{x}$=3.3, *Medyan*=3, *Mod*=1
c. $\bar{x}$=4, *Medyan*=4.2, *Mod*=2
d. $\bar{x}$=3, *Medyan*=4, *Mod*=6
e. $\bar{x}$=3.09, *Medyan*=4.1, *Mod*=1

3. Aşağıdakilerden hangisi minimum gözlem değeri 45, maksimum gözlem değeri 100, kareli ortalaması 58 olan serinin aritmetik ortalaması olabilir?

a. 45
b. 55
c. 65
d. 75
e. 85

4. Aşağıdaki tabloda 50 öğrencinin haftalık harcamaları verilmiştir.

Harcama (Lira)	f_i
50-75	3
75-100	6
100-125	14
125-150	11
150-175	2
175-200	5
200-225	2
225-250	5
250-275	1
275-300	1
Toplam	50

Öğrencilerin 1 haftalık harcalamalarına ilişkin bu veri seti için;

a. Aritmetik ortalamayı hesaplayınız.
b. Kareli ortalamayı hesaplayınız.

a. $\bar{x}$=144, *K*=154.15
b. $\bar{x}$=152, *K*=158.15
c. $\bar{x}$=154, *K*=154.15
d. $\bar{x}$=154, *K*=158.15
e. $\bar{x}$=144, *K*=158.15

5. Aşağıdaki veriler için aritmetik ortalama, medyan ve modu hesaplayınız.

11 17 18 10 22 23 15 17 14 13 10 12 18 18 11 14

a. $\bar{x}$=14.62, Mod=10, Medyan=14
b. $\bar{x}$=16.16, Mod=11, Medyan=14
c. $\bar{x}$=17.21, Mod=14, Medyan=14
d. $\bar{x}$=14.13, Mod=14, Medyan=14.5
e. $\bar{x}$=15.19, Mod=18, Medyan=14.5

6. 5. sorudaki verilerde 22 ve 23'ün yerine 42 ve 43'ü gözlemlediğinizi düşünün. Aritmetik ortalama, medyan ve modu yeniden hesaplayın. Üç ayrı merkezî eğilim ölçüsüne bu aşırı değerlerin etkisini belirtin.

a. Aritmetik ortalama değişmez, Mod=18, Medyan=14 olur.
b. Aritmetik ortalama değişmez, Mod=17, Medyan=15 olur.
c. $\bar{x}$=17.69 olarak değişir. Mod ve Medyan değişmez.
d. $\bar{x}$=18.12 olarak değişir. Mod ve Medyan değişmez.
e. $\bar{x}$=19.11 Mod=17 ve Medyan=14.5 olarak değişir.

7. Aşağıdaki ölçümler için mod, medyan ve aritmetik ortalamayı hesaplayınız.

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

a. Mod=1, Medyan=5, $\bar{x}$=6
b. Mod=5, Medyan=5.5, $\bar{x}$=6
c. Mod=8, Medyan=6, $\bar{x}$=6
d. Mod=10, Medyan=6, $\bar{x}$=5.5
e. Mod=8, Medyan=5.5, $\bar{x}$=5.5

8. Aşağıdaki frekans tablosunda yer alan veriler için standart sapmayı hesaplayınız.

Aralık	Frekans
0-2	1
3-5	3
6-8	5
9-11	4
12-14	2

a. 3.44
b. 3.96
c. 5.45
d. 9.45
e. 11.83

9. Aşağıdaki serinin aritmetik ortalamasının 15 olduğu bilindiğine göre, 5. gözlemin değeri nedir?

x_1=10 x_2=12 x_3=14 x_4=18 x_5= ?

a. 18
b. 19
c. 20
d. 21
e. 22

10. Aşağıdakilerden hangisi bir merkezî eğilim ölçüsü değildir?

a. Mod
b. Aritmetik ortalama
c. Kareli ortalama
d. Değişim aralığı
e. Medyan

Kendimizi Sınayalım Yanıt Anahtarı

1. e	Yanıtınız yanlış ise "Frekans Serilerinde Aritmetik Ortalama Hesabı" konusunu yeniden gözden geçiriniz.
2. b	Yanıtınız yanlış ise "Frekans Serilerinde Aritmetik Ortalama, Mod ve Medyan Hesabı" konusunu yeniden gözden geçiriniz.
3. b	Yanıtınız yanlış ise "Varyans ve Standart Sapma Hesabı" konusunu yeniden gözden geçiriniz.
4. a	Yanıtınız yanlış ise "Frekans Serilerinde Aritmetik Ortalama Hesabı ve Frekans Serilerinde Kareli Ortalama Hesabı" konusunu yeniden gözden geçiriniz.
5.e	Yanıtınız yanlış ise "Basit Serilerde Aritmetik Ortalama, Mod ve Medyan Hesabı" konusunu yeniden gözden geçiriniz.
6. c	Yanıtınız yanlış ise "Basit Serilerde Aritmetik Ortalama, Mod ve Medyan Hesabı" konusunu yeniden gözden geçiriniz.
7. b	Yanıtınız yanlış ise "Basit Serilerde Aritmetik Ortalama, Mod ve Medyan Hesabı" konusunu yeniden gözden geçiriniz.
8. a	Yanıtınız yanlış ise "Frekans Serilerinde Varyans ve Standart Sapma Hesabı" konusunu yeniden gözden geçiriniz.
9. d	Yanıtınız yanlış ise "Aritmetik Ortalamanın Özellikleri" konusunu yeniden gözden geçiriniz.
10. d	Yanıtınız yanlış ise "Merkezî Eğilim Ölçüleri" konusunu yeniden gözden geçiriniz.

Sıra Sizde Yanıt Anahtarı

Sıra Sizde 1

Aritmetik ortalama, kareli ortalama, medyan, kartiller ve mod merkezî eğilim ölçüleri arasında yer almaktadır.

Sıra Sizde 2

Veri seti içinde bir uç değer yer aldığından, serideki tüm gözlem değerlerinin hesaplamaya katıldığı aritmetik ortalama değil, duyarlı olmayan ve dolayısıyla uçdeğerlerden etkilenmeyen bir ortalama tercih edilmelidir. Tekrarlanan gözlem değeri bulunmadığından modun hesabı da uygun değildir. En uygun merkezi eğilim ölçüsü medyandır. 5. gözlem değeriyle 6. gözlem değerinin aritmetik ortalaması medyanı verecektir.

$$Medyan = \frac{1300 + 1500}{2} = 1400$$

10 üniversite mezununun aylık gelirlerinin medyanı ₺1400 olarak bulunur.

Sıra Sizde 3

$$\sigma^2 = K^2 - \mu^2 = 121 - 100 = 21$$

$$\sigma = \sqrt{21} = 4.58$$

Anakütlenin standart sapması 4,58'dir.

Sıra Sizde 4

Değişkenliğin araştırılması için her bir seri için ayrı ayrı değişim katsayılarının hesaplanması gerekmektedir.

$$D.K.(x) = \frac{s_x}{\overline{x}}.100 = \frac{5.87}{18}.100 = \%\ 32.61$$

$$D.K.(y) = \frac{s_y}{\overline{y}}.100 = \frac{6.94}{18,2}.100 = \%\ 38.13$$

y serisinin değişim katsayısı % 38.13 olarak hesaplandığından x serisine göre daha heterojen bir seri olduğu söylenebilir.

Yararlanılan ve Başvurulabilecek Kaynaklar

Akdeniz, F: **Olasılık ve İstatistik,** 13. Baskı, Adana Nobel Kitabevi, Adana, 2007.

Anderson, D.R., Sweeney, D.J., Williams, T.A.: Essentials of Statistics for Business and Economics, Publishing Company. USA, 1997.

Çömlekçi, N: **Temel İstatistik İlke ve Teknikleri,** 2. Baskı, Bilim Teknik Yayınevi, Eskişehir, 1994.

Newbold, Paul: (Çeviren: Ümit Şenesen), **İşletme ve İktisat İçin İstatistik,** Literatür Yayınları, İstanbul, 2000.

Ott, Lymann: **An Introduction to Statistical Methods and Data Analysis,** 3rd Edition, Duxbury Series in Statistics and Decision Sciences, 1988

İSTATİSTİK-I

3

Amaçlarımız

Bu üniteyi tamamladıktan sonra;

- Bir deneyin örnek noktalarını belirleyebilecek,
- Bir deneyin örnek uzayını oluşturabilecek,
- Örnek uzayda basit ve bileşik olaylar tanımlayabilecek,
- Olaylar üzerinde kesişim ve birleşim işlemlerini uygulayabilecek,
- Kombinasyon ve sayma kuralı yardımıyla örnek uzayın eleman sayısını hesaplayabilecek,
- Olasılık ölçüsünün özelliklerini tanımlayabilecek,
- Verilen tanımları kullanarak, bir olayın olasılığını hesaplayabileceksiniz.

Anahtar Kavramlar

- Deney
- Örnek Nokta
- Örnek Uzay
- Olay
- Basit Olay
- Bileşik Olay
- Ayrık Olaylar
- Olasılık Hesaplama

İçindekiler

Olasılık I

GİRİŞ

Olasılık kuramı, çeşitli sonuçlara ilişkin belirsizlikleri ölçmek için metotlar sağlamasının yanı sıra, belirsiz bir durum altında doğru ve sağlıklı kararlar verebilmede yardımcı olur.

Rastgelelik ve belirsizliği çağrıştıran "olasılık", terim olarak hemen hemen herkes tarafından bilinen ve sıkça kullanılan bir kavramdır. Örneğin; "Gelecek yıl işsizlik oranı *muhtemelen* artacak", "En az 900 biletin satılması *beklenmektedir*." şeklindeki ifadelere günlük hayatımızda sıklıkla karşılaşılır. Öte yandan, "Suya atılan bir demir parçasının suya batması olasılığı yüksektir." şeklindeki bir ifade doğru olmayacaktır. Çünkü fizik kanunları doğrultusunda, demirin yoğunluğunun suyun yoğunluğundan fazla olması nedeniyle demirin suya battığı bilinmektedir. Bu, bir fizik kanunudur ve demirin suya batması kesindir. Bu tür olaylara *kesin olaylar* denir. Kesin olayın aksine, "bir demir parçasının suya batmaması" şeklinde bir olay ise *imkansız (olanaksız) olay* olarak adlandırmaktadır.

Bir firma yetkilisi, piyasadaki bir ürünü hakkında tüketicilerinin beğenisini araştırmak istemektedir. Bu amaçla, ürünü kullanan bir gruba ürünü beğenisiyle ilgili bir soru sorar. Böyle bir deneyin sonuçları, "*Tüketici ürünü beğenmektedir.*", "*Tüketici ürünü beğenmemektedir.*" ve "*Tüketicinin ürün hakkında hiç bir fikri yoktur.*" şeklinde olabilir. Tüketicinin hangi cevabı vereceği önceden bilinmez. Başka bir örnek olarak, bir paranın atılması deneyi ele alınsın. Bu deneyin sonuçları, "*yazı*" ya da "*tura*" şeklindedir. Ancak, hangi sonucun geleceği önceden bilinmemektedir. Bir başka ifade ile deney sonuçlarının ortaya çıkması rastgeledir ve sonucu önceden bilinmez. Bu deneyde "*paranın yazı gelmesi* " şeklinde bir olay tanımlandığında, bu bir rassal (rastgele) olaya örnektir. Böylece gerçekleşmesi rastgele olan ve sonucu önceden bilinmeyen olaylara *rassal (rastgele) olay* denmektedir.

Bu örneklerden yola çıkarak, karşılaşılan olayları, kesin, imkansız ve rassal olmak üzere üç grupta değerlendirmek uygun olacaktır. Burada vurgulamak gerekir ki, olasılık kuramı, rassal olaylarla ilgilenen bilim dalıdır.

Bu bölümde, olasılıkla ilgili temel kavramlar ve olasılık hesaplama kuralları verilerek, verilen konulara ilişkin çeşitli örnekler çözülecektir.

DENEY, ÖRNEK NOKTA VE ÖRNEK UZAY

Bir firma yetkilisi, piyasadaki bir ürünü hakkında tüketicisinin beğenisini araştırmak istemektedir. Bunun için ürünü kullanan rastgele bir tüketici seçer ve ürünü beğenip beğenmediğini sorar. Yetkilinin bu eylemi rassal deneye örnek iken, *tüketici ürünü beğenmektedir.*", "*Tüketici ürünü beğenmemektedir.*" ve "*Tüketicinin ürün hakkında hiçbir fikri yoktur.*" şeklindeki sonuçlar ise deneyin sonuçlarıdır. Bu gözlem bilgisi sonuçlarına deneyin en temel sonuçları, bu sonuçların tamamına ise deneyin *örnek uzayı* adı verilmektedir.

Böylece, iyi tanımlanmış çeşitli gözlemler üreten eylem veya sürece *rassal deney* denmektedir. Rassal deneyin en temel sonucuna *örnek nokta* ve rassal deney sonucunda karşılaşılması muhtemel tüm örnek noktaların oluşturduğu kümeye *örnek (örneklem) uzay* denir ve örnek uzay S sembolü ile gösterilir.

Örnek 1:

a) *Hilesiz (Dengeli ve düzgün) bir para atılması deneyi ele alınsın. Y=Yazı ve T=Tura olmak üzere, bu deneyin en temel sonuçları, Y ve T dır. Bir başka deyişle, bu deneydeki örnek noktalar Y ve T dır. Deneyin örnek uzayı, örnek noktaların tümünün oluşturduğu küme olduğundan* $S=\{Y, T\}$ *olur.*

b) *Hilesiz bir zar atılması deneyinin ise altı muhtemel sonucu vardır. Bunlar; 1, 2, 3, 4, 5 ve 6 dır. 1, 2, 3, 4, 5 ve 6 aynı zamanda bu deneydeki örnek noktalar olup, deneyin örnek uzayı* $S=\{1,2,3,4,5,6\}$ *şeklindedir.*

Örnek uzayın her alt kümesine *olay* denir. Bu durumda, olaylar *örnek noktalardan* oluşmaktadır. Olaylar genellikle A, B, C gibi büyük harflerle gösterilmektedir.

Örnek 2: *Örnek 1 (b)'de verilen zar deneyi için iki olay*

A: Üste gelen sayının çift olması ve

B: Üste gelen sayının en az 5 olması

şeklinde tanımlansın. Bu durumda, A olayı 2, 4 ve 6 örnek noktalarından; B olayı ise 5 ve 6 örnek noktalarından oluşur. Bir başka ifade ile

$A=\{2,4,6\}$ *ve* $B=\{5,6\}$ *dır.*

Eğer olay tek bir örnek noktadan oluşuyorsa *basit (elemanter) olay,* birden çok örnek noktadan oluşuyorsa *bileşik olay* adını almaktadır.

Örnek 3: *Aşağıda bazı deneyler, bu deneylerin örnek uzayları ve tanımlanmış çeşitli olaylar verilmiştir.*

a)	*Deney*	:	*Hilesiz paranın bir kez atılması*
	Örnek noktalar	:	*Yazı (Y) ve Tura (T)*
	Örnek Uzay	:	$S=\{Y,T\}$
	A olayı	:	*Yazı gelmesi*

Bu deneyde, $A=\{Y\}$ olduğu açıktır. A olayı, tek bir örnek noktadan oluştuğundan basit olaya örnektir.

b)	*Deney*	:	*Hilesiz paranın iki kez atılması*
	Örnek noktalar	:	*YY, YT, TY, TT*
	Örnek Uzay	:	$S=\{YY, YT, TY, TT\}$
	B olayı	:	*En az bir yazı gelmesi*

Paranın iki kez atılması deneyinde Y= *Yazı* ve T= *Tura* olmak üzere, deneyin dört sonucu

YY= Birinci ve ikinci atışta yazı gelmesi,
YT= Birinci atışta yazı ve ikinci atışta tura gelmesi,
TY= Birinci atışta tura ve ikinci atışta yazı,
TT= Birinci ve ikinci atışta tura gelmesi

şeklinde gösterilirse, örnek uzay

$S = \{YY, YT, TY, TT\}$

olur. *B* olayı, "*en az bir yazı gelmesi*" olarak tanımlandığından, $B = \{YY, YT, TY\}$ biçiminde olur. Dolayısıyla, *B* olayı, üç örnek noktadan oluştuğu için bileşik olaya örnektir.

c) *Deney* : *Bir kutuda 1'den 50'ye kadar numaralanmış toplardan bir top seçimi*
Örnek noktalar :1,2,...,50
Örnek Uzay : $S = \{1,2,...,50\}$
C olayı :*Seçilen topun çift sayı ile numaralanmış top olması*

Rastgele bir top seçimi deneyinde olası sonuçlar yani örnek noktaları 1,2,...,50 şeklindedir.

C olayı, "*Seçilen topun çift sayı ile numaralanmış top olması*" olarak tanımlandığı için, $C = \{2,4,...,50\}$ olur. *C* olayı, 25 örnek noktadan oluştuğu için bileşik olaya örnektir.

DİKKAT

Örnek uzay, sonlu ya da sayılabilir sonsuz sayıda örnek noktadan oluşuyorsa bu örnek uzaya kesikli örnek uzay denir. Yukarıda verilen örnek uzaylar kesikli örnek uzay örnekleridir. Eğer örnek uzay (a,b) aralığı ya da bu aralıkların birleşimi gibi ise sürekli örnek uzay adını almaktadır. Sürekli örnek uzaylar, genellikle uzunluk, ağırlık, hız gibi ölçüm deneylerinde ortaya çıkarlar.

Bir deneyin örnek uzayı *venn şeması* ile gösterilip, *ağaç diyagramı* ile de oluşturulabilmektedir. *Venn şemasında*, örnek uzayın örnek noktaları kapalı bir şekil içinde birer nokta olarak gösterilir. *Ağaç diyagramı* ise bir deneyin mümkün olan tüm sonuçlarını göstermede kullanılan bir grafiktir.

Örnek 4: *Hilesiz bir paranın iki kez atılması deneyi tekrar ele alınsın.*
a) Venn şeması ile deneyin örnek uzayını gösterip, ağaç diyagramı yardımıyla deneyin örnek uzayını oluşturunuz.
b) Bu deneyde tüm basit olayları belirtiniz. Ayrıca bu deneye ilişkin iki olay
A : En çok bir paranın yazı gelmesi,
B : Her iki paranın da yazı gelmesi
olarak tanımlanmaktadır. Bu olaylar hangi örnek noktalarından oluşur? Bu olayları basit ya da bileşik olarak sınıflandırınız.

Çözüm:
a) *Y=Yazı* ve *T=Tura* olmak üzere, bu deneyin dört sonucu yani örnek noktaları

YY= Birinci ve ikinci atışta yazı gelmesi,
YT= Birinci atışta yazı ve ikinci atışta tura gelmesi,

TY= Birinci atışta tura ve ikinci atışta yazı,
TT= Birinci ve ikinci atışta tura gelmesi

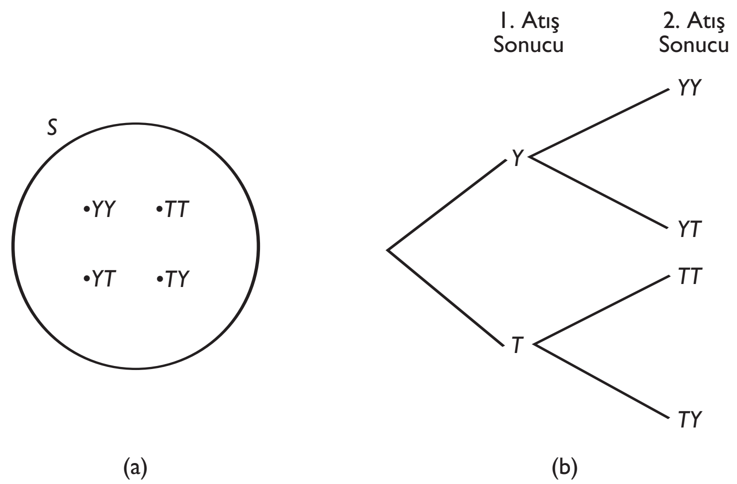
şeklinde gösterilirse, örnek uzayı

$S = \{YY, YT, TY, TT\}$

şeklinde olacaktır. Bu örnek uzay, venn şeması yardımıyla Şekil 3.1. (a)'da ve ağaç diyagramı yardımıyla Şekil 3.1 (b)'de gösterilmektedir. Şekil 3.1. (a)'da venn şemasında kapalı bir eğri içinde sonuçlar işaretlenir. Şekil 3.1. (b)'deki ağaç diyagramında ise aynı noktadan başlayan iki dal ve dalların sonuna sonuçlar, deneyin ikinci aşamasında iki dal daha ve dalların sonuna sonuçlar eklenir.

Şekil 3.1

Bir paranın iki kez atılması deneyinin örnek uzayı: (a) Venn şeması, (b) Ağaç diyagramı.



b) Bu deneyde ortaya çıkan dört örnek nokta, bu deneyin basit olaylarıdır. Bu örnek noktalar sırasıyla E_1, E_2 , E_3 ve E_4 şeklinde gösterilirse

$E_1 = YY$, $E_2 = YT$, $E_3 = TY$ ve $E_4 = TT$

olur. Bu bilgiler ışığında; A ve B olayları aşağıda gösterildiği gibi elde edilir.
A olayı, "*en çok bir paranın yazı gelmesi*" olarak tanımlandığından A olayı, hiç yazı gelmemesi ve bir yazı gelmesi sonuçlarını içerir. Dolayısıyla, $A = \{YT, TY, TT\}$ olur.

B olayı "*her iki paranın da yazı gelmesi*" olarak tanımlandığından $B = \{YY\}$ olur.

Birden fazla örnek noktadan oluşan A bileşik olay iken B basit olaydır.

Örnek 5: *Bir öğrenci eğitim-öğretim yılının ilk döneminde Sanat Tarihi (T) ya da Beden Eğitimi (B) derslerinden yalnızca birini, ikinci döneminde ise Halk dansları (H) ya da Salon Dansları (S) derslerinden yalnızca birini seçmeli ders olarak seçmek zorundadır. Bu durumda,*

a) Öğrenci kaç farklı ders seçimi yapabilir?
b) Venn şeması ile deneyin örnek uzayını gösterip, ağaç diyagramı yardımıyla deneyin örnek uzayını oluşturunuz.
c) Bu deneye ilişkin

A : Öğrencinin ilk dönem Sanat Tarihi dersini seçmesi olayı hangi örnek noktalardan oluşur, listeleyiniz.

Çözüm:

a) Bu deney, iki kez para atılması deneyine benzer olarak düşünülebilir. Ancak deneyin ilk aşamasında Sanat Tarihi (T) ve Beden Eğitimi (B) biçiminde iki sonuç, ikinci aşamasında Halk Dansları (H) ve Salon Dansları (S) biçiminde iki sonuç vardır. Bu deneyin sonuçları;

TH = İlk dönem Sanat Tarihi ve ikinci dönem Halk Dansları seçilmesi,
TS = İlk dönem Sanat Tarihi ve ikinci dönem Salon Dansları seçilmesi,
BH= İlk dönem Beden Eğitimi ve ikinci dönem Halk Dansları seçilmesi,
BS = İlk dönem Beden Eğitimi ve ikinci dönem Salon Dansları seçilmesi

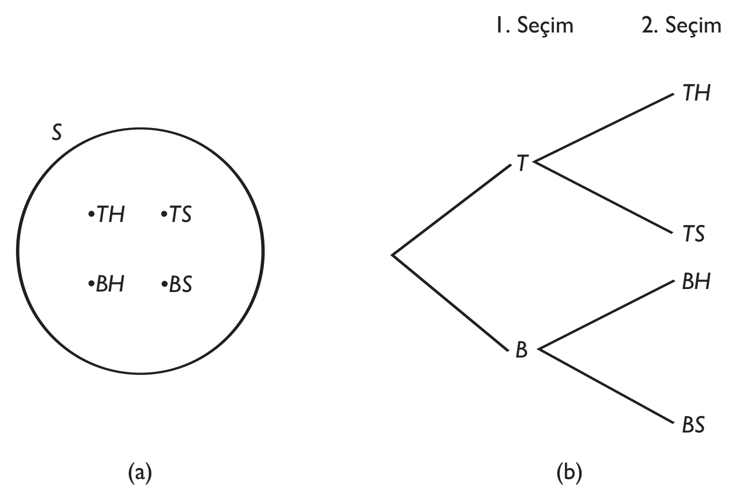
şeklinde gösterilirse, örnek uzayı

$S=\{TH, TS, BH, BS\}$

biçiminde olur. Dolayısıyla 4 farklı seçim yapabilir.

b) Bu deneyin örnek uzayı Şekil 3.2 (a) Venn şeması ve (b) Ağaç diyagramı şeklinde gösterilmiştir. Sonuç olarak, öğrenci 4 farklı seçim yapabilir.

Şekil 3.2



Öğrencinin ders seçimi deneyinin örnek uzayı: (a) Venn şeması, (b) Ağaç diyagramı.

c) A olayı, "*öğrencinin ilk dönem Sanat Tarihi dersini seçmesi*" olarak tanımlandığından

$A=\{TH, TS\}$

olur.

Örnek 6: *Dört farklı banka (X, Y, Z, U) müşterilerine, sabit (S) oranlı veya değişken (D) oranlı olmak üzere ev kredisi teklif etmektedir. Bu durumda,*

a) Bir müşteri kaç farklı ev kredisi seçimi yapabilir?

b) A olayı, "bir müşterinin sabit oranlı kredi seçmesi" olarak tanımlanırsa A olayı hangi örnek noktalarından oluşur, listeleyiniz.

Çözüm:

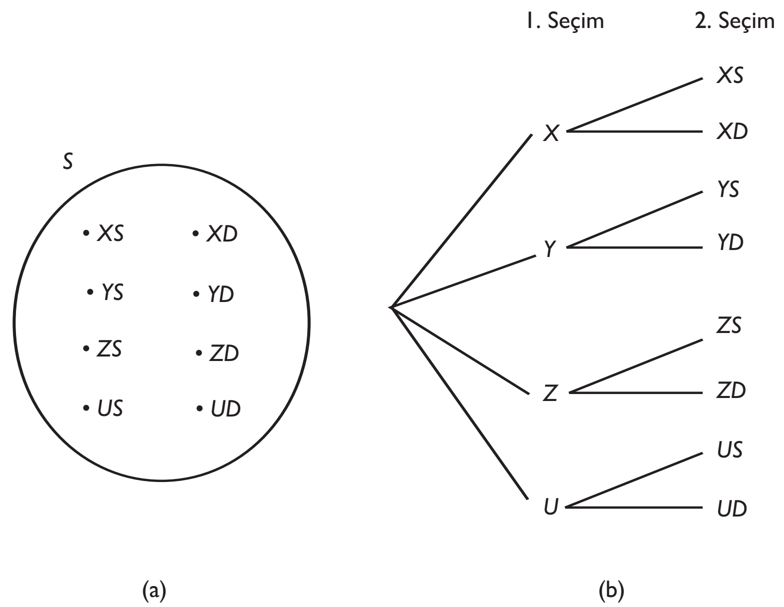
a) Müşterinin yapacağı seçimler "*XS*=*X* bankasını ve sabit oranlı (*S*) kredi seçmesi" şeklinde listelenirse ortaya çıkacak olası tüm sonuçlar

$S=\{XS, XD, YS, YD, ZS, ZD, US, UD\}$

şeklinde olur ve müşteri 8 farklı seçim yapabilir. Şekil 3.3'te deney sonuçları venn şeması ve ağaç diyagramı ile gösterilmektedir.

Şekil 3.3

Müşterinin banka ve kredi seçimi deneyinin örnek uzayı (a) Venn şeması (b) Ağaç diyagramı



b) Bir müşterinin sabit oranlı kredi seçmesi *A* olayı ile gösterilirse;

$A=\{XS, YS, ZS, US\}$

biçimindedir.

SIRA SİZDE 1

1. Hilesiz bir para ve hilesiz bir zarın birlikte atılması deneyi ele alınsın.
a) Bu deneye ilişkin örnek uzayı venn şeması ile gösteriniz ve ağaç diyagramı yardımıyla oluşturunuz.
b) Bu deneyde basit olay ve bileşik olay örneği veriniz.
2. Bir firma, çalışanları arasından 3 kişiye, gelecek yıl yönetimde yapacakları belli bir değişikliği onaylayıp onaylamadığını soracak olursa toplam kaç olası sonuçla karşılaşabilir. Bu deneyin örnek uzayını ağaç diyagramı yardımıyla oluşturunuz.
3. Art arda iki zar atılması deneyinde *A* olayı, "*en az bir zarın* 6 *gelmesi*" ise *A* olayı hangi örnek noktalardan oluşur listeleyiniz.

OLAYLAR ÜZERİNDE İŞLEMLER

Olaylar birer küme olduğu için olaylarla çalışırken kümelerle ilgili bazı işlemleri kullanmak yararlı olacaktır. Bu nedenle, bu bölümde, olaylar üzerinde çeşitli işlemler tanıtılacaktır.

S örnek uzayı, *A* ve *B* iki olay olsun.

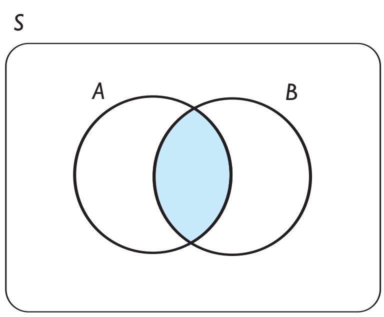
- *A* ve *B* olaylarının her ikisine birden ait olan örnek noktalarının oluşturduğu kümeye *A* ile *B*'nin *kesişim (arakesit) kümesi* denir. Bu küme $A \cap B$ ile gösterilir ve "*A* kesişim *B*" diye okunur.
- *A* veya *B* olaylarından en az birine ait olan örnek noktalarının oluşturduğu kümeye *A* ile *B*'nin *birleşim kümesi* denir. Bu küme $A \cup B$ ile gösterilir ve "*A* birleşim *B*" diye okunur. *A* birleşim *B* kümesinde *A*'nın, *B*'nin ya da her ikisinin örnek noktaları bulunur.
- *A* ve *B* olaylarının ortak örnek noktası yok ise, bunlara **ayrık olaylar** denir ve $A \cap B$ kümesinin boş küme olduğu söylenir. Boş küme $\varnothing$ sembolü ile gösterilir. Daha genel olarak *S* örnek uzayında $A_1, A_2, \ldots, A_n$ gibi *n* tane olayın her bir çifti ayrık ise bu olaylara *karşılıklı ayrık olaylar* denir. Kısaca $i \neq j$ için $A_i \cap A_j = \varnothing$ şeklinde gösterilir.
- *S* örnek uzayında $A_1, A_2, \ldots, A_n$ karşılıklı ayrık olaylar ($i \neq j$ için $A_i \cap A_j = \varnothing$) olsun. $A_1 \cup A_2 \cup \ldots \cup A_n = S$ ise bu olaylara **bütüne tamamlayıcı (bütüne tamamlanan)** olaylar denmektedir.
- *S* örnek uzayında *A* bir olay olsun. *A*'da bulunmayan *S* içinde diğer tüm örnek noktalarının oluşturduğu olaya *A* olayının *tümleyeni (tamamlayanı)* denir. *A*'nın tümleyeni $\overline{A}$ ile gösterilir. Ayrıca *A* ve $\overline{A}$ olayları ayrık olaylardır. Bir başka ifade ile $A \cap \overline{A} = \varnothing$ dir.

A ve *B* olaylarının ortak örnek noktası yok ise bu olaylara **ayrık olaylar** denir.

Bir örnek uzayda tüm örnek noktaları hem ayrık hem de **bütüne tamamlayan**dır.

Grafiksel Gösterimler

Olaylarla işlem yapılırken venn şeması çizimi işlemleri kolaylaştırmaktadır. Şekil 3.4'te *S* dikdörtgeni örnek uzayı, *A* ile *B* olayları göstermektedir. *A* ve *B* olaylarının her ikisine ait olan örnek noktalarının oluşturduğu küme kesişim (arakesit) yani $A \cap B$ dir. Bu arakesit, taralı alan şeklinde ifade edilmektedir.



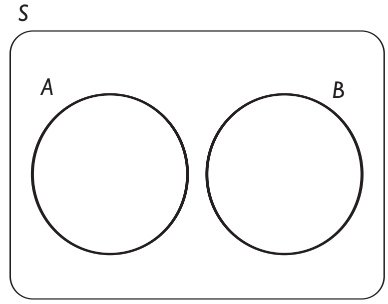
Şekil 3.4

A ve B kesişimi: taralı alan

A ve *B* olaylarının ortak örnek noktası olmayabilir. Böyle olaylar ayrık olaylardır. Şekil 3.5'de *A* ve *B* ayrık olaylara örnektir.

Şekil 3.5

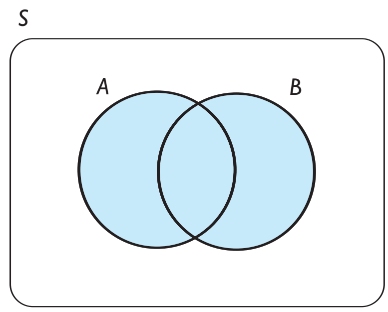
A ve B ayrık olaylar



Şekil 3.6'da *A* ve *B* olaylarından en az birine ait olan örnek noktalarının oluşturduğu küme, yani birleşim kümesi ($A \cup B$) gösterilmektedir. Şekil 3.6'dan görüldüğü gibi, bir örnek nokta *A*'nın ya da *B*'nin ya da her ikisinin örnek noktası ise $A \cup B$'nin örnek noktası olacaktır.

Şekil 3.6

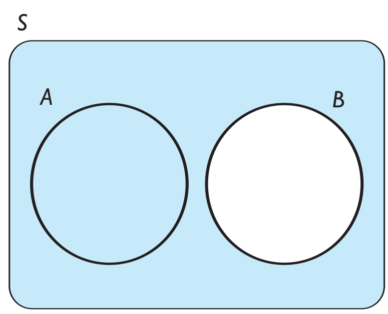
A ve B birleşimi: taralı alan



Şekil 3.7'de ise *B* olayının tümleyeni ($\bar{B}$), *S*'nin içinde olan ancak *B*'nin içinde olmayan örnek noktalarının oluşturduğu olaydır. Şekil 3.7'den görüldüğü gibi $\bar{B}$ olayı taralı alandır.

Şekil 3.7

B olayının tümleyeni: taralı alan



DİKKAT

A ve $\bar{A}$ olayları ayrık olaylardır. ($A \cap \bar{A} = \varnothing$)

Örnek 7: *Hilesiz bir zar atılsın A ve B ile olayları aşağıdaki gibi tanımlansın:*
A: Üste gelen sayının 4'ten küçük olması (4 dahil değil),
B: Üste gelen sayının çift olması.
Buna göre,
a)A ve B olayları hangi örnek noktalarından oluşmaktadır?
b) $\overline{A}$ olayı hangi örnek noktalarından oluşmaktadır?
c) $A \cap B$ ve $A \cup B$ olayları hangi örnek noktalarından oluşmaktadır?
d) A ve B olayları ayrık olaylar mıdır?
e) Bu deneyde bütüne tamamlayıcı iki olay örneği veriniz.

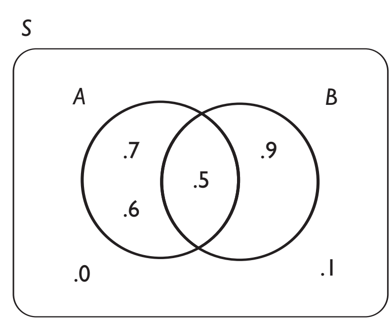
Çözüm: Bu deneyin örnek uzayının

S={1, 2, 3, 4, 5, 6}

olduğu açıktır. Buna göre,

a) *A* olayı, "*üste gelen sayının 4'ten küçük olması*" ise *A* olayının örnek noktaları 1, 2 ve 3' tür ve bunun bir sonucu olarak *A*={1, 2, 3} olur. *B* olayı, "*üste gelen sayının çift olması*" ise *B* olayının örnek noktaları 2, 4 ve 6 dır ve bunun bir sonucu olarak *B*={2,4, 6} olur.
b) *A* olayının tümleyeni *A*'yı içermeyen diğer tüm örnek noktaları içeren olay olduğuna göre $\overline{A}$ = {4, 5, 6} olur.
c) $A \cap B$ olayı, *A* ve *B* olaylarının ortak örnek noktalarından oluştuğundan $A \cap B = \{2\}$ olur. $A \cup B$ olayı, *A* ve *B* olaylarından en az birine ait olan örnek noktalarının oluşturduğu olay olduğundan $A \cup B = \{1, 2, 3, 4, 6\}$ olur.
d) *A* ve *B*'nin arakesiti boş küme olmadığından, bu olaylar ayrık olaylar değildir.
e) *C* ve *D* olayları $C \cap D = \varnothing$ ve $C \cup D = S$ ise bu olaylar bütüne tamamlayıcı olaylardır. Bu koşulları sağlayan pek çok iki olay tanımlanabilir. Örneğin, $C = \{1, 3, 5\}$ ve $D = \{2, 4, 6\}$ şeklinde tanımlandığında $C \cap D = \varnothing$ ve $C \cup D = S$ olur. Böylece tanımlanan *C* ve *D* bütüne tamamlayıcı olaylara örnektir.

Örnek 8: *Bir deneyin örnek uzayı S ve A ile B olayları venn şeması şeklinde verilmiştir.*



Şekil 3.8
S örnek uzayı, A ve B olayları.

a) A ve B ayrık olaylar mıdır?
b) A'nın tümleyeni ($\overline{A}$) hangi örnek noktalarından oluşur?
c) $A \cup B$ ve $A \cap \overline{B}$ olayları hangi örnek noktalardan oluşmaktadır?

Çözüm: A= {5, 6, 7} ve B= {5, 9} olduğu Şekil 3.8'den görülmektedir.

a) A ve B olaylarının kesişimi ($A \cap B$={5}) boş küme olmadığından ayrık olaylar değillerdir.

b) A'nın tümleyeni, A'da olmayan diğer tüm örnek noktaları içeren küme olduğuna göre, $\overline{A}$ = {0, 1, 9} şeklindedir.

c) $A \cup B$ = {5, 6, 7, 9} ve $A \cap \overline{B}$ = {6, 7} şeklindedir.

SIRA SİZDE 2

1. 1,2,...,15 sayıları küçük kâğıtlara yazılıp kapatılıyor ve bir kutunun içine konuluyor. Kutudan rastgele bir kâğıt çekildiğinde ve
***A*: *Asal sayı olması*,**
B: 2 *ile bölünen sayı olması*
olayları tanımlandığında
a) *A*∩*B* olayları hangi örnek noktalarından oluşur?
b) *A* ve *B* ayrık olaylar mıdır?

2. Hilesiz bir paranın üç kez atılması deneyinde, *A*, *B*, *C* ve *D* olayları
***A*: *En az bir tura gelmesi*,**
***B* : *En az iki yazı gelmesi*,**
***C* : *En çok iki yazı gelmesi*,**
D* : *Üç paranın da aynı yüzünün gelmesi
şeklinde tanımlanıyor. Bu olaylar hangi örnek noktalardan oluşur? *A*∩*B* ve *C*∪*D* hangi örnek noktalardan oluşur?

3. Hilesiz bir zar atılması deneyinde,
A: Çift sayı gelmesi,
B: 3'ten büyük sayı (3 dahil) gelmesi
olayları tanımlanıyor. Buna göre,
a) *A* ve *B* olayları ayrık mıdır?
b) *A* ve *B* olaylarının tümleyenlerini bulunuz.
c) Bu örnek uzayda bütüne tamamlayıcı iki olay ifade ediniz.

FAKTÖRİYEL, KOMBİNASYON VE SAYMA KURALI

Bir olayın olasılığı hesaplanırken, örnek uzayın eleman sayısının bulunması temel zorluklardan birisidir. Bazı durumlarda, kombinasyon ve sayma kuralı örnek uzayın eleman sayısını bulmada kolaylık sağlamaktadır.

Faktöriyel

1'den n'ye kadar olan pozitif tamsayıların çarpımına "n faktöriyel" denir ve $n!$ şeklinde gösterilir.

$n! = n.(n-1)...1$
$n! = n.(n-1)!$

dir. Ayrıca 0!=1 ve 1!=1 olarak tanımlanır. Örneğin 6! değerini bulmak için 6'dan 1'e kadar tüm tamsayıların çarpılması gerekir. 6!= 6. 5. 4. 3. 2. 1=720

Kombinasyon

"n" sayıda farklı eleman arasından "k" elemanlı farklı grup oluşturma işlemine kombinasyon denir. Böylece "n" elemanın "k" elemanlı farklı grup sayısı

$$\binom{n}{k} = \frac{n!}{(n-k)!k!}$$

şeklinde hesaplanır.

Kombinasyon, sembol olarak $\binom{n}{k}$ ile gösterilmekle birlikte, C_k^n sembolü de kullanılmaktadır.

Örneğin A, B, C elemanlarından oluşabilecek iki elemanlı farklı gruplar AB, AC, BC şeklindedir. Yani 3 elemanlı bir kümeden seçilen 2 elemanlı birbirinden farklı grupların sayısı 3'tür. Bu sonuç kombinasyon yardımıyla

$\binom{3}{2} = \frac{3!}{(3-2)!2!} = \frac{3.2.1}{(1).(2.1)} = 3$ şeklinde elde edilir.

Örnek 9: $\binom{5}{1}$ *ve* $\binom{5}{3}$ *kombinasyonlarını hesaplayınız.*

Çözüm: Kombinasyonun tanımından

$$\binom{5}{1} = \frac{5!}{(5-1)!1!} = \frac{5.4.3.2.1}{(4.3.2.1)} = 5$$

şeklinde bulunur. Bu sonuç daha basit şekilde aşağıdaki gibi hesaplanır.

$$\binom{5}{1} = \frac{5.4!}{4!} = 5$$

$$\binom{5}{3} = \frac{5!}{(5-3)!3!} = \frac{5!}{2!3!} = \frac{5.4.3.2.1}{(2.1)(3.2.1)} = 10$$

Örnek 10: *Bir firma yetkilisi 4 kişi arasından 2 kişiyi üretimi denetlemesi için seçmek istemektedir. Firma yetkilisi 2 kişilik kaç farklı grup oluşturabilir?*

Çözüm: 4 kişi arasından 2 kişilik farklı grup sayısı, kombinasyon yardımıyla bulunabilir. $n = 4$, $k = 2$ olduğundan

$$\binom{4}{2} = \frac{4!}{(4-2)!2!} = \frac{4.3.2.1}{(2.1)(2.1)} = 6$$

bulunur. Böylece, firma yetkilisi 4 kişi arasından 2 kişilik 6 farklı grup oluşturabilir.

Sayma Kuralı

Sayma kuralı "*Eğer bir deneyde birden fazla aşama (adım) varsa bu deneydeki toplam sonuç sayısı, her bir aşamadaki sonuç sayılarının çarpımına eşittir.*" şeklinde ifade edilir. Örneğin, bir deneyde ilk aşamada n_1, ikinci aşamada n_2 ve üçüncü aşamada n_3 tane sonuç olmak üzere üç aşama bulunuyorsa bu deneyin toplam sonuç sayısı, $n_1.n_2.n_3$ tür.

Örnek 11: *A şehrinden B şehrine 2, B şehrinden de C şehrine 3 farklı yol vardır. Bu durumda A'dan C'ye, B'ye uğramak koşuluyla kaç farklı yolla gidilebilir?*

Çözüm: Deney iki aşamadan oluşmaktadır. Deneyin ilk aşamasında A şehrinden B şehrine 2 yol seçeneği (sonuç sayısı 2), benzer olarak ikinci aşamasında B'den C'ye 3 yol seçeneği (sonuç sayısı 3) olduğuna göre toplam gidilebilecek farklı yol sayısı, sayma kuralı yardımıyla 2.3=6'dır.

Örnek 12: *Bir hentbol liginde her takım, kendi grubundaki takımlarla bir sezonda 6 maç yapmaktadır. Her maçın sonucu: galibiyet, mağlubiyet ve beraberlik şeklinde sınıflandırılmaktadır. Buna göre, bir takım için bir sezon boyunca kaç farklı sonuç söz konusudur?*

Çözüm: Bir sezonda 6 maç, her maçta üç sonuç olduğuna göre, deneyin aşama sayısı 6, her bir aşamadaki sonuç sayısı 3'tür. Bu durumda, bu deneyin toplam sonuç sayısı, $3.3.3.3.3.3=3^6$ olur.

SIRA SİZDE 3

1. Hilesiz iki para ve hilesiz iki zarın art arda atılması deneyinin örnek uzayının eleman sayısı kaçtır?
2. Belli bir üniversitede İstatistik kulübünün 5 üyesi vardır. Bu üyeler arasından 2 kişi Matematik kulübüne seçilecektir. Bu iki kişi kaç farklı şekilde seçilebilir?
3. Bir şirket yöneticisinin, bir projeyi yönetmek üzere görevlendirebileceği 3 çalışanı ve proje yöneticisine yardımcı olabilecek 4 çalışanı vardır. Şirket yöneticisi rastgele olarak bir proje yöneticisi ve bir yardımcı seçecektir. Şirket yöneticisi kaç farklı seçim yapabilir?

OLASILIK HESAPLAMA

Olasılık, bir olayın gerçekleşebilirliğinin sayısal bir ölçüsüdür. Olasılık değeri, 0 ve 1 arasında değişir. Eğer olay imkânsız ise olasılığı 0, kesin olay ise 1'dir.

Olasılık Ölçüsünün Özellikleri

Bir rassal deneyin örnek uzayı S, örnek noktaları E_1, E_2, ..., E_n olmak üzere $S=\{E_1, ..., E_n\}$ ve A da bir olay olsun. E_i örnek noktasının olasılığı $P(E_i)$ ve A olayının olasılığı $P(A)$ ile gösterilsin. Olasılık ölçüsü aşağıda (i), (ii) ve (iii) de verilen özellikleri sağlar.

Bir deneyde tüm örnek noktaların olasılıkları toplamı 1'dir.

$E_i \cap E_j = \emptyset$, $i \neq j$,

$E_1 \cup E_2 \cup ... \cup E_n = S$ dir.

(i) Bir olayın olasılığı her zaman 0 ile 1 arasındadır:

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

(ii) Bir deneyde tüm örnek noktaların olasılıkları toplamı 1'dir:

$P(E_1)+...+P(E_n)=1$

(iii) $P(S)=1$

Klasik Olasılık Tanımı

Klasik olasılık tanımı, tüm örnek noktaları (sonuçları) eşit olasılıklı olan deneylerde olasılık hesaplamalarında kullanır. Eğer bir deneyde n sayıda örnek nokta varsa, her örnek noktanın olasılığı $1/n$ dir. Örneğin, hilesiz bir para deneyinde yazı gelmesi ve tura gelmesi örnek noktaları eşit olasılığa sahip ve 1/2'dir. Hilesiz bir zar deneyinde tüm sonuçların ortaya çıkması eşit olasılığa sahip ve 1/6'dır. Benzer olarak, 100 kişilik bir gruptan rassal (rastgele) olarak bir kişinin seçilmesi olasılığı 1/100'dür.

S örnek uzayı *n* tane örnek noktadan oluşsun ve tüm örnek noktalar eşit olasılığa sahip olsun. E_1, ..., E_n örnek noktaları göstermek üzere

$$P(E_i) = \frac{1}{n} \text{ dir.}$$

Örnek 13: *Bir kutuda 1'den 10'a kadar numaralandırılmış 10 kart vardır. Kutudan rastgele seçilen bir kartın 6 numaralı kart olması olasılığı kaçtır?*

Çözüm: Bu deneyde 10 tane sonuç, bir başka deyişle, 10 örnek nokta vardır. Bu durumda, n=10'dur. Tüm örnek noktalar eşit olasılıklıdır; çünkü her bir örnek noktanın ortaya çıkma olasılığı 1/10'dur.

E={5} olup, klasik olasılık tanımına göre, E olayının olasılığı

$$P(E) = \frac{1}{n} = \frac{1}{10} = 0.1 \text{ dir.}$$

Örnek 14: *Hilesiz bir para iki kez atıldığında, her iki paranın da yazı gelmesi olasılığı nedir?*

Çözüm: Bu deney sonucunda, dört örnek nokta E_1, E_2, E_3 ve E_4

E_1=*YY*, E_2=*YT*, E_3=*TY* ve E_4=*TT*

şeklinde olur ve deneyin örnek uzayı S= {*YY*, *YT*, *TY*, *TT*} biçimindedir. Bir paranın iki kez atılması deneyinde tüm örnek noktaların ortaya çıkması eşit olasılığa sahip ve deneyin sonuç sayısı n=4'tür. Dolayısıyla her bir örnek noktanın olasılığı 1/4'tür.

$$P(E_1) = P(E_2) = P(E_3) = P(E_4) = \frac{1}{4}$$

Böylece $P(E_1) = \frac{1}{4}$ tür.

Olasılığın Göreli Sıklık Tanımı

Bu olasılık yaklaşımı, bir olayın olasılığını, olayın gerçekleşme sıklığının toplam deney sayısına oranı olarak tanımlanmaktır.

Aşağıdaki örnekler ele alınsın.

- Hileli (kusurlu veya dengeli olmayan) bir paranın atılması sonucunda yazı gelmesi olasılığı,
- Hileli bir zarın atılması sonucunda 6 gelmesi olasılığı,
- Bir meteoroloji uzmanının yarın yağmur yağma olasılığının %90 olduğunu iddia etmesi,
- Bir ampulün dayanma süresinin en az 1000 saat olması.

Yukarıda ifade edilen deneylere ilişkin sonuçlar eşit olasılığa sahip olmadığı için bu olayların olasılıkları, klasik olasılık tanımı ile hesaplanamaz. Böyle durumlarda, deney birçok kez tekrar edilerek gözlem üretilmekte ve bu gözlemlerden yararlanılarak bir olaya ilişkin, frekanslar (sıklıklar) yardımıyla yaklaşık bir olasılık değeri hesaplanmaktadır. Bu olasılık tanımına **olasılığın göreli sıklık** tanımı denmektedir.

Bir rassal deney, aynı koşullar altında *N* kez tekrarlanmış ve bu *N* deneyde herhangi bir *A* olayı *f* kez gözlemlenmiş ise göreli sıklık tanımına göre *A* olayının ortaya çıkması olasılığı, deney sayısı *N* artarken *f/N* oranıdır. Böylece, *A* olayının olasılığı

Göreli sıklık tanımına göre, *A* olayının olasılığı $P(A) = \frac{f}{N}$, dir. Burada, *f*, *A* olayının gerçekleşme sıklığı, *N* gerçekleşen deney sayısıdır.

$P\left(A\right)=\dfrac{f}{N}$ şeklinde bulunur.

Örnek 15: *Bir firma piyasadaki bir ürünü hakkında tüketicisinin beğenisini araştırmak istemektedir. Bunun için bir soru 500 kişiye sorulmuş ve elde edilen sonuçlar aşağıdaki tabloda verilmiştir.*

Tablo 3.1
Tüketicilerin beğenisi ile ilgili sonuçlar.

Çok beğeniyorum	Beğeniyorum	Fikrim Yok	Beğenmiyorum	Hiç beğenmiyorum
80	200	60	120	40

Buna göre,

a) *Örnek noktaları listeleyiniz ve olasılıklarını belirleyiniz.*

b) Rastgele seçilen yeni bir tüketicinin "ürünü beğeniyor" olması olasılığı nedir?

Çözüm:

a) Deneyin örnek noktaları

E_1=Çok beğeniyorum, E_2=Beğeniyorum, E_3=Fikrim Yok, E_4=Beğenmiyorum, E_5=Hiç beğenmiyorum şeklindedir.

N=500 toplam tüketici sayısı, E_1, E_2, E_3, E_4 ve E_5 olaylarının görülme sayıları (f) sırasıyla 80, 200, 60, 120 ve 40 olmak üzere, olasılığın göreli sıklık tanımına göre (f/N)

$$P(E_1)=\frac{80}{500},\ P(E_2)=\frac{200}{500},\ P(E_3)=\frac{60}{500},\ P(E_4)=\frac{120}{500},\ P(E_5)=\frac{40}{500}\ \text{dır.}$$

b) Rastgele seçilen bir tüketicinin "ürünü beğeniyor" olması

$$P(E_2)=\frac{200}{500}=0.4\ \text{dür.}$$

Örnek 16: *Bir fabrikada üretim bandındaki ürünler, rastgele olarak dört farklı renkte ambalajlarla paketlenmektedir. Bu üretim bandından rassal olarak 1000 tane ürün seçilmiş ve paket renkleri ile ilgili olarak aşağıdaki sonuçlar elde edilmiştir.*

Tablo 3.2
Paket renkleri ve görülme sıklığı

Renk	Görülme Sıklığı
Beyaz	300
Kırmızı	200
Siyah	230
Mavi	270
Toplam	1000

Buna göre, bunlardan sonra üretilecek ilk ürünün ambalajının kırmızı renkli olması olasılığı nedir?

Çözüm: E olayı, "*üretilecek ilk ürünün ambalajının kırmızı renkli olması*" şeklinde tanımlanırsa; N=1000 toplam ürün sayısı, kırmızı ambalajlı ürün sayısı f=200 olmak üzere, olasılığın göreli sıklık tanımına göre E olayının olasılığı

$$P\left(E\right)=\frac{f}{N}=\frac{200}{1000}=0.2$$

biçiminde elde edilir. İlk ürünün ambalajının kırmızı renkli olması olasılığı yaklaşık olarak 0.2 dir. Bu olasılığın yaklaşık olasılık olduğu unutulmamalıdır.

Örnek 17. *Hileli bir zar 1000 kez atılmış ve aşağıdaki sonuçlar elde edilmiştir.*

Sonuç	Görülme Sıklığı
1	50
2	200
3	400
4	250
5	30
6	70

Tablo 3.3
Hileli zar atma deneyi sonuçları ve görülme sıklığı

Buna göre, yeniden atılan zarın 3 gelmesi olasılığı nedir?

Çözüm: E olayı "*atılan zarın 3 gelmesi*" şeklinde tanımlanırsa, N=1000 toplam deney sayısı, E olayının gözlenme sayısı f=400 olmak üzere, olasılığın göreli sıklık tanımına göre E olayının olasılığı

$$P\left(E\right) = \frac{400}{1000} = 0.4$$

olarak elde edilir. Gerçekleştirilen deney sayısı arttıkça E olayının olasılık değerinin daha iyi tahmin edileceği açıktır.

Olaylar ve Olasılıkları

Bir A olayının olasılığı, A içindeki bütün örnek noktaların olasılıkları toplamına eşittir:

$$P(A) = \sum_{E_i \in A} P\left(E_i\right)$$

Böylece aşağıdaki sonuçlar çıkartılabilir.

1. S örnek uzayında A olay ve A'nın tümleyeni de $\bar{A}$ olsun. Bu durumda,

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

eşitliği geçerlidir.

2. $P(S) = 1$ olduğundan $P(\varnothing) = 0$
3. S örnek uzayı n tane örnek noktadan oluşsun ve tüm örnek noktalar eşit olasılığa sahip olsun. A olayı içinde n_A sayıda örnek nokta var ise, A olayının olasılığı (A olayı için klasik olasılık tanımı)

$$P(A) = \frac{n_A}{n}$$

şeklindedir.

4. A ve B iki olay $B \subseteq A$ ise $P(B) \leq P(A)$

Böylece B olayı, A'nın alt kümesi ise B'nin olasılığı, A'dan büyük olamaz.

DİKKAT

Bir A olayının olasılığı hesaplanırken izlenecek adımlar aşağıdaki gibi sıralanabilir:

1. **Deneyi tanımla.**
2. **Deneyin örnek noktalarını listele.**
3. **Örnek noktaların olasılıklarını belirle.**
4. **A olayının içindeki bütün örnek noktaları belirle.**
5. **A olayı içindeki bütün örnek noktaların olasılıklarını topla.**

Örnek 18: *Hilesiz bir zar atışında iki olay*

A : Gelen sayının tek olması

B : Gelen sayının bir ***olmaması***

şeklinde tanımlansın. A ve B olaylarının olasılığı kaçtır?

Çözüm: Bu deneyde 6 tane sonuç, bir başka deyişle, 6 örnek nokta vardır. Bu durumda, n=6'dır. Tüm örnek noktalar eşit olasılıklıdır; çünkü her bir örnek noktanın ortaya çıkma olasılığı 1/6'dır.

A={1,3,5} olup, A olayı 3 örnek noktadan oluşmaktadır. A olayının olasılığı

$P(A) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = 0.5$ olarak bulunur.

B={2,3,4,5,6} olup, B olayı 5 örnek noktadan oluşmaktadır. Böylece B olayının olasılığı

$P(B) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{5}{6} = 0.833$ 'tür.

Öte yandan, olaylar için klasik olasılık tanımına göre ($\frac{n_A}{n}$) göre istenen olasılıklar: A olayı içindeki örnek noktası sayısı n_A=3, örnek uzaydaki tüm örnek noktaları sayısı n=6 olmak üzere

$P(A) = \frac{n_A}{n} = \frac{3}{6} = 0.5$

ve benzer olarak B olayı içindeki örnek noktası sayısı n_B=5, örnek uzaydaki tüm örnek noktaları sayısı n=6 olmak üzere

$P(B) = \frac{n_B}{n} = \frac{5}{6} = 0.833$ olarak bulunur.

Örnek 19: *Hilesiz bir zar atışında,*

A : Gelen sayının tek olması

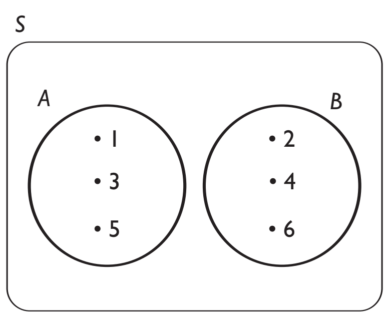
B : Gelen sayının çift olması

şeklinde iki olay tanımlansın. A veya B olayının olasılığı kaçtır?

Çözüm: Zar deneyinin örnek uzayı S={1,2,3,4,5,6} olduğundan 6 örnek noktaya sahiptir. Bu durumda, n=6'dır. Tüm örnek noktalar eşit olasılığa sahiptir ve her bir örnek noktanın ortaya çıkma olasılığı 1/6'dır. Açıktır ki, A={1,3,5} ve B={2,4,6} şeklindedir. Bu durumda, $A \cup B = \{1, 2, 3, 4, 5, 6\}$ olup $P(A \cup B) = \frac{1}{6} + \ldots + \frac{1}{6} = 1$ dir. Öte yandan, $A \cup B = S$ olduğundan $P(S)$=1'dir.

Şekil 3.9

S örnek uzayı, A ve B olaylarının örnek noktaları



Örnek 20: *Hilesiz bir para iki kez atıldığında,*

A : En az bir paranın yazı gelmesi,

B : En çok iki paranın da tura gelmesi,

C : Her iki paranın da yazı gelmesi

olayları tanımlansın. Bu olayların olasılıklarını hesaplayınız.

Çözüm: Bu deneyde ortaya çıkan dört örnek nokta bu deneyin basit olaylarıdır. Bu örnek noktalar sırasıyla E_1, E_2, E_3 ve E_4 şeklinde gösterilirse

$E_1=YY$, $E_2=YT$, $E_3=TY$ ve $E_4=TT$

olur ve deneyin örnek uzayı $S=\{YY, YT, TY, TT\}$ biçimindedir. Bir paranın iki kez atılması deneyinde tüm örnek noktalar eşit olasılığa sahiptir. Deneyin sonuç sayısı n=4 olduğundan dolayısıyla her bir örnek noktanın olasılığı 1/4'tür.

$$P(E_1) = P(E_2) = P(E_3) = P(E_4) = \frac{1}{4}$$

A olayı, "*en az bir paranın yazı olması*" olarak tanımlandığından A olayı, bir yazı ve iki yazı gelmesi sonuçlarını içerir. Dolayısıyla, $A=\{YY, YT, TY\}$ ve $A=\{E_1, E_2, E_3\}$ olur. Böylece

$$P(\mathrm{A}) = P(E_1) + P(E_2) + P(E_3) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4} \text{ tür.}$$

Diğer bir çözüm, A olayının klasik olasılık tanımı: deneyin toplam sonuç sayısı n, A olayının sonuç sayısı n_A ise A olayının olasılığı

$$P(A) = \frac{n_A}{n} = \frac{3}{4} \text{ tür.}$$

B olayı, "*en çok iki paranın da tura olması*" olayı ise hiç tura gelmemesi, bir tura gelmesi ve iki tura gelmesi olaylarını içerir. Dolayısıyla, $B=\{YY, YT, TY, TT\}$ veya

$B=\{E_1, E_2, E_3, E_4\}=S$

Böylece

$$P(\mathrm{B}) = P(E_1) + P(E_2) + P(E_3) + P(E_4) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = 1 \text{ olur.}$$

Öte yandan B olayının içinde n_B=4 sayıda örnek nokta vardır ve B olayının klasik olasılık tanımına göre olasılığı:

$$P(B) = \frac{n_B}{n} = \frac{4}{4} = 1$$

Ayrıca B olayı S örnek uzayı olduğuna göre $P(S)$=1'dir.

C olayı "*her iki paranın da yazı gelmesi*" olarak tanımlandığından $C=\{YY\}$ olur.

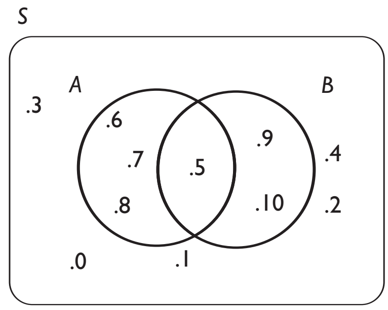
$$P(\mathrm{C}) = P(E_1) = \frac{1}{4} \text{ 'tür.}$$

Örnek 21: *Bir deneyin örnek uzayı A ile B olayları venn şeması şeklinde Şekil 3.10'da verilmiştir.*

$P(0)=P(1)= P(2)= P(3)=$ 1/30; $P(4)= P(5)=$ 1/30; $P(6)= P(7)= P(8)= P(9)=$4/30 ve $P(10)$=8/30 *şeklinde verilsin. Bu durumda, aşağıdaki olasılıkları hesaplayınız.*

Şekil 3.10

S örnek uzay, A ve B olaylarının örnek noktaları



a) $P(A)$
b) $P(\overline{B})$
c) $P(A \cup \overline{A})$
d) $P(A \cap B)$
e) $P(\overline{A} \cap B)$
f) *Tüm örnek noktalar eşit olasılığa sahip olsa idi $A \cap B$ olayının olasılığı ne olurdu?*

Çözüm: Bu örnekte, tüm örnek noktaların *eşit olasılığa sahip olmadığına* dikkat edilmelidir.

A olayının olasılığı, *A* içindeki bütün örnek noktalarının olasılıkları toplamına eşittir.

a) *A* olayının örnek noktalarının {5,6,7,8} olduğu Şekil 3.9'dan görülmektedir.

$$P(A) = P(5) + P(6) + P(7) + P(8) = \frac{1}{30} + \frac{4}{30} + \frac{4}{30} + \frac{4}{30} = \frac{13}{30}$$

olarak elde edilir.

b) *B* olayının örnek noktaları {5,9,10} dur. $\overline{B}$ olayı, *B* olayının tümleyeni olduğuna göre

$$\begin{aligned} P(\overline{B}) &= 1 - P(B) \\ &= 1 - \{P(5) + P(9) + P(10)\} \\ &= 1 - \left\{\frac{1}{30} + \frac{4}{30} + \frac{8}{30}\right\} = \frac{17}{30} \end{aligned}$$ şeklindedir.

c) A={5,6,7,8} ve $\overline{A}$ ={0,1,2,3,4,9,10} dir. Açıktır ki, $A \cup \overline{A}$ = {0,1,2,3,4,5,6,7,8,9,10} dır. Buradan, $A \cup \overline{A} = S$ olduğu görülür. Bu durumda, $P(A \cup \overline{A})=P(S)=1$ elde edilir.

d) Hem *A*'da hem de *B*'de olan örnek noktalardan oluşan olay $A \cap B$'dir. Bu olay, Şekil 3.10'daki arakesite karşılık gelen bölümde ifade edilmektedir. Dolayısıyla,

$A \cap B$={5} olduğundan $P(5) = \frac{1}{30}$ bulunur.

e) $P(\overline{A} \cap B)$ hesaplayabilmek için önce $\overline{A} \cap B$'nin örnek noktalarının belirlenmesi gerekir. $\overline{A}$ ={0,1,2,3,4,9,10} ve B={5,9,10} ise hem $\overline{A}$ 'da hem de *B*'de olan örnek noktalar oluşan küme $\overline{A} \cap B$={9,10} dur. Bu durumda,

$$P(\overline{A} \cap B) = P(9) + P(10) = \frac{4}{30} + \frac{8}{30} = \frac{12}{30}$$ olarak hesaplanır.

f) Tüm örnek noktaları eşit olasılığa sahip olsa idi 11 örnek nokta olduğu için

$P(0) = P(1) = ... = P(10) = \frac{1}{11}$ olurdu ve

$P(A \cap B) = P(5) = \frac{1}{11}$ olarak hesaplanırdı.

Örnek 22: *Bir deneyin örnek uzayı, dört örnek noktadan oluşmaktadır. Aşağıdaki tabloda örnek noktalar ve bunların olasılıkları verilmektedir.*

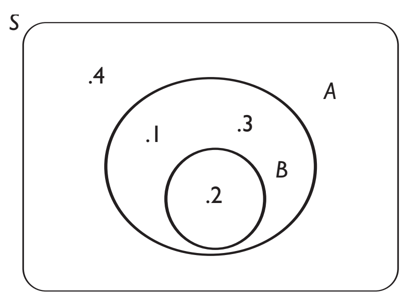
$P(S)=1$, $P(\varnothing)=0$'dır.

Örnek noktalar	Olasılıkları
1	0.15
2	0.25
3	0.20
4	0.40

Olaylar:
$A=\{1,2,3\}$; $B=\{2\}$ şeklinde tanımlansın.
a) *A ve B olaylarını venn şeması içerisinde gösteriniz.*
b) *$P(A\cup B)$ ve $P(A\cap B)$ olasılıklarını hesaplayınız.*
c) *$B\subseteq A$ ise $P(B) \leq P(A)$ olduğunu gösteriniz.*

Çözüm:
a) Deneyin örnek uzayı (S), A ve B olayları aşağıdaki şekildeki gibidir.

Şekil 3.11



S örnek uzay ve A ve B olayları

b) $A\cup B=\{1,2,3\}=A$'dır.
$P(A\cup B)= P(1)+ P(2)+P(3)= 0.15+0.25+0.20= 0.60$
$A\cap B=\{2\}$'dir.
$P(A\cap B)= 0.25$
c) $B\subseteq A$ ($\{2\} \subseteq \{1,2,3\}$), $P(B)=0.25$ ve $P(A)=0.60$ olduğundan $P(B)\leq P(A)$ sağlandığı görülür.

Örnek 23: *Bir okulda yıl sonu balosunda düzenlenen piyango çekilişinde bir kişinin herhangi bir ödül kazanma olasılığı 1/50'dir. Buna göre, bu kişinin çekilişte ödül* ***kazanamama*** *olasılığı nedir?*

Çözüm Bu deney için bütüne tamamlayıcı iki olay

A : *Bir kişinin ödül kazanması,*
$\bar{A}$: *Bir kişinin ödül kazanamaması*

şeklinde tanımlandığında $S=\{A, \bar{A}\}$ dir. Bu durumda, soruda verilenlere göre,

$$P(A) = \frac{1}{50}$$

dir. Ayrıca, $P(\bar{A})+P(A)=1$ olduğundan

$$P(\bar{A}) = 1 - P(A) = 1 - \frac{1}{50} = \frac{49}{50}$$

bulunur. Bir başka deyişle, bu piyangoda bir kişinin ödül kazanamama olasılığı 49/50'dir.

DİKKAT

Ayrık olaylar birlikte ortaya çıkmazlar.

Örnek 24: *Belli bir üniversitede Ayşe ve Umut'unda içinde olduğu beş kişiden oluşan bir İstatistik kulübü vardır. Bu üyeler arasından rastgele 2 kişi Matematik kulübüne seçilecektir. Bu iki kişinin Ayşe ve Umut olması olasılığı kaçtır?*

Çözüm: Bu deneyin örnek uzayının eleman sayısı, "5" farklı eleman arasından "2" elemanlı grup sayısıdır. Dolayısıyla kombinasyon yardımıyla

$$\binom{5}{2} = \frac{5!}{(5-2)!2!} = \frac{5.4.3.2.1}{3.2.2} = 10$$

olarak bulunur. 10 deneyin toplam sonuç sayısıdır

10 sayıda 2 elemanlı grup arasında "*Ayşe ve Umut*" bir örnek noktadır. Buna göre

$$P(\text{Ayşe ve Umut seçilmesi}) = \frac{1}{10} \text{ dur.}$$

SIRA SİZDE 4

1. Hileli bir zar 1000 kez atılmış, 10 kez 6 geldiği gözlenmiştir. Buna göre, ilk atılan zarın altı gelmemesi olasılığı nedir?
2. Üretim yapan bir fabrikada belli bir üründen rastgele 500 tanesi seçilmiş ve üretime ilişkin aşağıdaki sonuçlar elde edilmiştir.

Tablo 3.4
Kusurlu ve kusursuz üretimin görülme sıklığı

Sonuç	Görülme Sıklığı
Kusursuz	450
Kusurlu	50

Buna göre, rastgele seçilen ilk ürünün
a) Kusurlu olması olasılığı nedir?
b) Kusurlu veya kusursuz olması olasılığı nedir?
3. İki farklı hilesiz para atılsın ve *A* ile *B* olayları aşağıdaki gibi tanımlansın:
A*: *Her iki paranın üste gelen yüzünün yazı olması,
B*: *En az bir paranın üste gelen yüzünün yazı olması.
Buna göre, $A \cup B$ ve $A \cap B$ olaylarının olasılıklarını bulunuz.

Özet

AMAÇ 1

Bir deneyin örnek noktalarını belirlemek.

Bir deneyin en temel sonucuna **örnek nokta** denmektedir. Örnek noktaların her biri basit olaydır.

AMAÇ 2

Bir deneyin örnek uzayını oluşturmak.

Rassal deney sonucu karşılaşılması muhtemel tüm örnek noktaların oluşturduğu kümeye **örnek (örneklem) uzay** denir.

AMAÇ 3

Örnek uzayda basit ve bileşik olayları tanımlamak.

Örnek uzayının her alt kümesine **olay** denmektedir. Yani olaylar örnek noktalardan oluşmaktadır. Eğer olay tek bir örnek noktadan oluşuyorsa **basit olay**, birden çok örnek noktadan oluşuyorsa **bileşik olay** adını almaktadır.

AMAÇ 4

Olaylar üzerinde kesişim ve birleşim işlemlerini uygulamak.

Olayların kesişimi ve birleşimi olasılık hesaplamalarında oldukça önemlidir.

A ve B olaylarının her ikisine birden ait olan örnek noktalarının oluşturduğu kümeye A ile B'nin **kesişim (arakesit) kümesi** denir. Bu küme $A \cap B$ ile gösterilir ve "A kesişim B" diye okunur.

A veya B olaylarından en az birine ait olan örnek noktalarının oluşturduğu olaya A ile B'nin **birleşim kümesi** denir. Bu küme $A \cup B$ ile gösterilir ve "A birleşim B" diye okunur. $A \cup B$ birleşim kümesinde A'nın, B'nin ya da her ikisinin örnek noktaları bulunur.

AMAÇ 5

Kombinasyon ve sayma kuralı yardımıyla örnek uzayın eleman sayısını hesaplamak.

Kombinasyon

"n" farklı elemanın *"k"* elemanlı farklı grup sayısı

$\binom{n}{k} = \frac{n!}{(n-k)!k!}$ şeklinde hesaplanır.

Sayma Kuralı

Sayma kuralı *"Eğer bir deneyde birden fazla aşama varsa bu deneydeki toplam sonuç sayısı, her bir aşamadaki sonuç sayılarının çarpımına eşittir."* şeklinde ifade edilir. Örneğin, bir deneyde ilk aşamada n_1, ikinci aşamada n_2 ve üçüncü aşamada n_3 tane sonuç olmak üzere üç aşama bulunuyorsa bu deneyin toplam sonuç sayısı, $n_1.n_2.n_3$ tür.

AMAÇ 6

Olasılık ölçüsünün özelliklerini tanımlamak.

Bir rassal (rastgele) deneyin örnek uzayı S, örnek noktaları $E_1, E_2, ..., E_n$ olmak üzere $S = \{E_1, ..., E_n\}$ ve A da bir olay olsun. E_i örnek noktasının olasılığı $P(E_i)$ ve A olayının olasılığı $P(A)$ ile gösterilsin. Olasılık ölçüsü aşağıda (i), (ii) ve (iii) de verilen özellikleri sağlar.

(i) Bir olayın olasılığı her zaman 0 ile 1 arasındadır:
$0 \leq P(E_i) \leq 1$ ve $0 \leq P(A) \leq 1$

(ii) Bir deneyde tüm örnek noktaların olasılıkları toplamı 1'dir:
$P(E_1) + ... + P(E_n) = 1$

(iii) $P(S) = 1$

AMAÇ 7

Verilen tanımları kullanarak bir olayın olasılığını hesaplamak.

Bir A olayının olasılığı, A içindeki bütün örnek noktaların olasılıkları toplamına eşittir:

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

Genel olarak bir A olayının olasılığı hesaplanırken izlenecek adımlar aşağıdaki gibi sıralanabilir:

1. Deneyi tanımla.
2. Deneyin örnek noktalarını listele.
3. Örnek noktaların olasılıklarını belirle.
4. A olayının içindeki bütün örnek noktaları belirle.
5. A olayı içindeki bütün örnek noktaların olasılıklarını topla.

Kendimizi Sınayalım

1. Hilesiz bir paranın üç kez atılması deneyinde *A* olayı “en az bir tura gelmesi” şeklinde tanımlandığında, *A* olayının tümleyeni hangi örnek noktalarından oluşur?

a. YYY
b. YYT, TYY, YTY
c. YTT, TYT, TTY
d. YTT, TYT, TTY,YYY
e. YYT, TYY, YTY, YYY

2. Belli bir marka araba satın almak isteyen bir müşterinin, beş farklı renk seçeneklerinden birini, iki vites seçeneklerinden birini tercih etmesi gerekmektedir. Bu durumda müşteri kaç farklı seçim yapabilir?

a. 2
b. 3
c. 5
d. 7
e. 10

3. Bir girişimci, 10 yeni sektörden 3 tanesine yatırım yapmak istemektedir. Girişimci kaç farklı seçim yapabilecektir?

a. 3
b. 13
c. 30
d. 120
e. 720

4. I. 1.5
II. 0.3
III. -0.7

Yukarıdakilerden hangisi ya da hangileri olasılık değeri **olamaz?**

a. Yalnız I
b. Yalnız II
c. Yalnız III
d. I-II
e. I-III

5. Hilesiz bir paranın üç kez atılması deneyinde *A* olayı “en çok bir tura gelmesi” şeklinde tanımlandığında, *A* olayının olasılığı aşağıdakilerden hangisidir?

a. 1/8
b. 4/8
c. 5/8
d. 6/8
e. 1

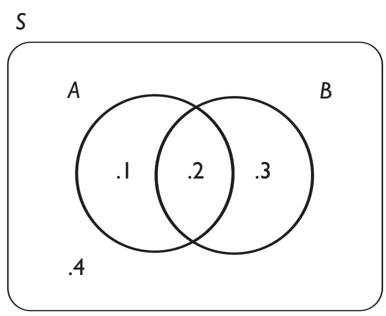
6. Asitli içecek üreten bir firma, dört farklı şişe tasarımı üzerinde çalışmış ve 500 tüketicisine tercih ettiği tasarımı sormuştur. Elde ettiği sonuçlar aşağıdaki tabloda gösterilmiştir.

Şişe tasarımı	Görülme Sıklığı
A	50
B	75
C	350
D	25
Toplam	500

Yeni bir müşterinin D tasarımına sahip şişeyi seçmesi olasılığı nedir?

a. 0.05
b. 0.01
c. 0.15
d. 0.70
e. 0.80

7. *S* bir deneyin örnek uzayı ve *A* ile *B* olayları venn şeması şeklinde aşağıda verilmiştir.
$P(1)= P(2)= 2/10$; $P(3)= 1/10$
$P(4)= 5/10$ şeklinde verildiğinde,
$P(\bar{A} \cap B)$ olasılığı aşağıdakilerden hangidir?



a. 1/10
b. 2/10
c. 3/10
d. 4/10
e. 5/10

8. Bir fabrikanın bir günlük kusurlu üretim sayıları ve onlara ilişkin göreli sıklık olasılıkları aşağıdaki tabloda verilmiştir. Buna göre, bir günde ikiden (2'i hariç) az kusurlu ürün çıkması olasılığı nedir?

Kusurlu Ürün Sayısı	0	1	2	2 den çok
Olasılığı	0.50	0.30	0.15	0.05

a. 0.30
b. 0.50
c. 0.45
d. 0.75
e. 0.80

9. Bir öğrencinin Muhasebe dersinden geçme olasılığı 0.75 ise bu öğrencinin Muhasebe dersinden kalma olasılığı nedir?

a. 0.05
b. 0.25
c. 0.50
d. 0.75
e. 1

10. Çoktan seçmeli bir test sınavında sorular için dört seçenek bulunmaktadır. Herhangi bir sorunun cevabı rassal olarak işaretlenecek olursa, cevabın yanlış olması olasılığı nedir?

a. 0.05
b. 0.25
c. 0.50
d. 0.75
e. 1

Yaşamın İçinden

"Bir otomobil sigorta şirketi, belli bir il için otomobil sigorta primi hesaplanırken sürücünün yaş faktörünün de etkisi olduğunu öne sürmektedir. Bu amaçla, bu ilde son on yılda gerçekleşen otomobil kaza sayıları bilgisine ulaşır. Bu bilgiler yardımıyla kazaların yaşlara göre dağılımını ele alır. Bu bilgilere dayalı olarak toplam kazalar içinde genç sürücü olarak nitelendirdiği 18-28 yaş grubunun oranının yüksek olduğunu bulur. Dolayısıyla genç bir sürücünün kaza yapma olasılığının yüksek olabileceği sonucuna ulaşılır. Bu sonuç, sigorta şirketi tarafından değerlendirilir. Böylece, şirket, otomobil sigorta primi hesaplanırken yaş faktörünün etkisi de hesaba katılabileceği sonucuna ulaşabilir."

Kendimizi Sınayalım Yanıt Anahtarı

1. a	Yanıtınız yanlış ise "Deney, Örnek Nokta ve Örnek Uzay" ve "Olaylar Üzerinde İşlemler" konularını yeniden gözden geçiriniz.
2. e	Yanıtınız yanlış ise "Deney, Örnek Nokta ve Örnek Uzay" ve "Faktöriyel, Kombinasyon ve Sayma Kuralı" konularını yeniden gözden geçiriniz.
3. d	Yanıtınız yanlış ise "Faktöriyel, Kombinasyon ve Sayma Kuralı" konusunu yeniden gözden geçiriniz.
4. e	Yanıtınız yanlış ise "Olasılık Hesaplama" konusunu yeniden gözden geçiriniz.
5. b	Yanıtınız yanlış ise "Olasılık Hesaplama" konusunu yeniden gözden geçiriniz.
6. a	Yanıtınız yanlış ise "Olasılık Hesaplama" konusunu yeniden gözden geçiriniz.
7. a	Yanıtınız yanlış ise "Olasılık Hesaplama" konusunu yeniden gözden geçiriniz.
8. e	Yanıtınız yanlış ise "Olasılık Hesaplama" konusunu yeniden gözden geçiriniz.
9. b	Yanıtınız yanlış ise "Olasılık Hesaplama" konusunu yeniden gözden geçiriniz.
10. d	Yanıtınız yanlış ise "Olasılık Hesaplama" konusunu yeniden gözden geçiriniz.

Sıra Sizde Yanıt Anahtarı

Sıra Sizde 1

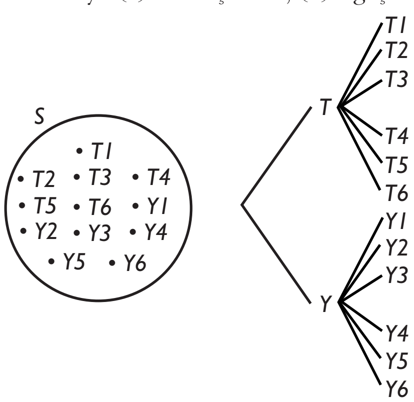
1. Hilesiz bir para ve hilesiz bir zarın birlikte atılması deneyinin sonuçları, paranın yazı (*Y*) veya tura (*T*) gelmesi, zarın 1, 2, 3, 4, 5, 6 gelmesi durumuna göre şekillenir. Örneğin, para tura ve zar 1 gelebilir, bu sonuç *T* 1, şeklinde ve diğer sonuçlar benzer şekilde gösterildiğinde, para ve zar atma deneyinin örnek uzayı

S= {*T1, T2, T3, T4, T5, T6, Y1, Y2, Y3, Y4, Y5, Y6*}

biçiminde olur.

a) Bu deneye ilişkin Venn şeması ve ağaç diyagramı aşağıdaki şekildeki gibidir.

Şekil 3.12. Bir para ve bir zarın birlikte atılması deneyinin örnek uzayı: (a) Venn şeması, (b) Ağaç diyagramı



(a) (b)

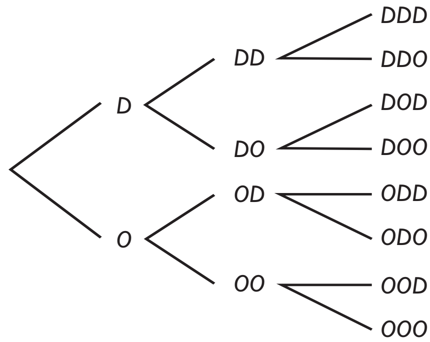
b) *S*= {*T1, T2, T3, T4, T5, T6, Y1, Y2, Y3, Y4, Y5, Y6*} ve *S*'nin her bir sonucu basit olaydır. Örneğin *A*= {*Paranın tura, zarın 3 gelmesi*} = {*T3*} basit olaya örnektir. Birden çok sonuçtan oluşan bileşik olaya ise

B= {*Zarın çift gelmesi*}= {*T2, T4, T6, Y2, Y4, Y6*} olayı örnek olarak verilebilir.

2. Bu deney üç kez para atılması deneyine benzer olarak düşünülebilir. Deneyin üç aşamasında Değişikliği onaylaması *(D)* ve Değişikliği onaylamaması *(O)* biçiminde iki sonuç olacaktır. Üç kişinin vereceği tüm olası cevapların oluşturduğu küme

S= {*DDD, DDO, DOD, ODD, DOO, ODO, OOD, OOO*}

biçiminde olur. Örnek uzayın eleman sayısı 8 olarak elde edilir. Bir sonraki bölümde göreceğimiz, sayma kuralı yardımıyla örnek uzayın eleman sayısı 2.2.2= 8 şeklinde de hesaplanabilir. Bu deneyin örnek uzayı ağaç diyagramı yardımıyla aşağıdaki şekilde elde edilmiştir.



3. Hilesiz iki zar atma deneyinin örnek uzayı S= {(1,1), (1,2),..., (6,6)} olacak şekilde 36 örnek noktadan oluşur. *A* olayı, "en az bir zarın 1 gelmesi" olayı "bir zarın 1 gelmesi" ve "iki zarın 1 gelmesi" olaylarını içerir.

A= {(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,1), (3,1), (4,1), (5,1), (6,1)}

şeklindedir.

Sıra Sizde 2

1. Yalnız kendisine ve 1'e bölünen sayıya asal sayı denmektedir. Ayrıca en küçük asal sayı 2 kabul edilir. Böylece

A= {2, 3, 5, 7, 11, 13} biçiminde, *B* olayı ise

B= {2, 4, 6, 8, 10, 12, 14} biçiminde olur.

a) *A* ve *B* olaylarının ortak örnek noktası 2 olduğundan dolayı, $A \cap B$= {2} dir.

b) *A* ve *B* olaylarının ortak örnek noktası olduğundan dolayı ayrık değillerdir.

2. Hilesiz bir paranın üç kez atılması deneyinin örnek uzayı

S= {*YYY, YYT, YTY, TYY, YTT, TYT, TTY, TTT*} biçiminde olur.

A olayı "en az bir tura gelmesi" ise

A olayı "bir tura gelmesi", "iki tura gelmesi" ve "üç tura gelmesi" olaylarını içerir. Böylece

A= {*YYT, YTY, TYY, YTT, TYT, TTY, TTT*}

B "en az iki yazı gelmesi" olayı, "iki yazı gelmesi", "üç yazı gelmesi" olaylarını içerir.

B= {*YYT, YTY, TYY, YYY*}

C, "en çok iki yazı gelmesi" olayı, "hiç yazı gelmemesi", "bir yazı gelmesi" ve "iki yazı gelmesi" olaylarını içerir. Dolayısıyla

C= {*TTT, YTT, TYT, TTY, YYT, YTY, TYY*}

D, "üç paranın da aynı yüzünün gelmesi"

D= {*YYY, TTT*}

şeklinde olur.

$A \cap B = \{YYT, YTY, TYY\}$ ve $C \cup D = \{TTT, YYT, YTY, TYY, YTT, TYT, TTY, YYY\} = S$

3. Hilesiz bir zar atılması deneyinde, A olayı çift sayı gelmesi ve B olayı 3'ten (3 dahil) büyük sayı gelmesi olayları ise $A = \{2, 4, 6\}$ ve $B = \{3, 4, 5, 6\}$ biçimindedir.

a) $A \cap B = \{4, 6\}$ yani $A \cap B \neq \varnothing$ olduğundan A ve B olayları ayrık değillerdir.

b) A'nın tümleyeni $\bar{A} = \{1, 3, 5\}$ ve B'nin tümleyeni $\bar{B} = \{1, 2\}$ şeklinde olur.

c) Bu örnek uzayda bütüne tamamlayıcı iki olay C ve D olsun.

$C \cap D = \varnothing$ ve $C \cup D = S$ olmalıdır. Bu koşulları sağlayan pek çok C ve D olayları tanımlanabilir. Örneğin, $C = \{1, 2, 3, 4, 5\}$ ve $D = \{6\}$ şeklinde tanımlanan C ve D bütüne tamamlayıcı iki olaydır.

Sıra Sizde 3

1. Hilesiz iki para ve hilesiz iki zar atışı deneyinin örnek uzayının eleman sayısı, sayma kuralı yardımıyla 2. 2. 6. 6=144' tür.

2. "5" sayıda farklı eleman arasından "2" elemanlı farklı grup sayısı kombinasyon yardımıyla

$$\binom{5}{2} = \frac{5!}{(5-2)!2!} = \frac{5.4.3.2.1}{(3.2.1).(2.1)} = 10$$ olarak bulunur.

İstatistik kulübünün 5 üyesi arasından matematik kulübü için 2 kişi 10 farklı şekilde seçilebilir.

3. Projeyi yönetmek için 3 çalışan, yardımcı olabilecek 4 çalışan olduğuna göre, şirket yöneticisi bir proje yöneticisi ve yardımcısını, çarpma kuralı yardımıyla 3.4= 12 farklı şekilde seçebilir.

Sıra Sizde 4

1. Hileli bir zar 1000 kez atılmış 10 kez 6 geldiği gözlenmiş ise göreli sıklık tanımına göre

$$P(\{6\ gelmesi\ olayı\}) = \frac{10}{1000}$$ ve

$$P(\{6\ gelmemesi\ olayı\}) = 1 - \frac{10}{1000} = \frac{990}{1000}$$ 'dir

2. Olaylar,

A: Rastgele seçilen ilk ürünün kusurlu olması

$\bar{A}$:Rastgele seçilen ilk ürünün kusursuz olması şeklinde tanımlanır.

a) Kusurlu olması olayının olasılığı

$$P(A) = \frac{50}{500} = 0.1$$ 'dir.

b) A ve $\bar{A}$ olayları bu deneyin örnek noktaları yani basit olaylarıdır. Dolayısıyla bu olaylar ayrık ve bütüne tamamlanan olaylar olduğundan kusurlu veya kusursuz olması olayı, Tablo 3.4'teki veriler kullanılarak

$$P(A \cup \bar{A}) = P(A) + P(\bar{A}) = \frac{450}{500} + \frac{50}{500} = 1$$ 'dir.

3. Hilesiz iki para deneyinin örnek uzayı $S = \{YY, YT, TY, TT\}$ olacak şekildedir. Tüm örnek noktalar eşit olasılığa sahip ve olasılığı $\frac{1}{4}$ 'tür.

$A = \{YY\}$ ve $B = \{YY, YT, TY\}$

şeklindedir. $A \cup B = \{YY, YT, TY\}$ $A \cap B = \{YY\}$ dir.

$$P(A \cup B) = \frac{3}{4} = 0.75 \text{ ve } P(A \cap B) = \frac{1}{4} = 0.25$$ 'tir.

Yararlanılan ve Başvurulabilecek Kaynaklar

Anderson D.R., Sweeney D.J., Williams T.A. (2005). **Statistics for Business and Economics,** China: Thomson-South-Western.

Akdeniz F. (2007). **Olasılık ve İstatistik,** Adana: Nobel Kitapevi.

Ben M., Levy H. (1983). **Business Statistics Fundamentals and Applications,** New York, USA: Random House Inc.

McClave J.T., Benson P.G., Sincich T. (2001). **Statistics for Business and Economics,** New Jersey, USA:Prentice-Hall Inc.

Yüzer A. (1996). **Olasılık ve İstatistik,** Eskişehir: Anadolu Üniversitesi Yayınları.

Yüzer A.,Ağaoğlu E., Tatlıdil H., Özmen A., Şıklar E. (2006). **İstatistik,** Eskişehir: Anadolu Üniversitesi, Açıköğretim Fakültesi Yayınları.

İSTATİSTİK-I

4

Amaçlarımız

Bu üniteyi tamamladıktan sonra;

- Koşullu olasılıkları hesaplayabilecek,
- Bileşik olasılıkları hesaplayabilecek,
- Bağımsız olaylara ilişkin olasılıkları hesaplayabilecek,
- Bağımsız ve ayrık olaylar arasındaki farkı ayırt edebilecek,
- Olayların birleşiminin olasılığını hesaplayabileceksiniz.

Anahtar Kavramlar

- Olasılık
- Koşullu Olasılık
- Bileşik Olasılık
- Bağımsız Olaylar
- Bağımlı Olaylar
- İki Olayın Birleşiminin Olasılığı

İçindekiler

Olasılık II

GİRİŞ

Bir önceki bölümde olasılık kavramı ve olasılık hesaplama kuralları anlatıldı. Hatırlanacağı gibi bir deneyin en temel sonucu olan örnek noktanın olasılığı ve bir A olayının olasılığı tanımlandı. Bu bölümde, ilk olarak bir B olayının bilinmesi durumunda A olayının olasılığı, yani koşullu olasılık üzerinde durulacak ve konunun izleyen alt bölümlerinde bağımsız ve bağımlı olaylar ele alınacaktır. Daha sonraki bölümde ise olayların birleşiminin olasılığını bulmada kullanılan toplama kuralından söz edilecek ve tüm verilen konulara ilişkin çeşitli örnekler çözülecektir.

SIRA SİZDE 1

1. Bir hastanede çalışan 5 evli doktor çiftten (çiftlerin her ikisi de doktor olmak üzere) rastgele (rassal olarak) ikisi hastane başhekimi ve yardımcısı görevlerine seçilecektir. Seçilen iki doktorun evli bir çift olması olasılığı nedir?

2. Bir sınıfta 22'si kız olmak üzere 50 öğrenci bulunmaktadır. Bu sınıftan rassal olarak bir sınıf başkanı seçilecektir. Bu öğrencinin erkek öğrenci olması olasılığı nedir?

3. Bir kutudaki bilyelerin %40'ı siyah, %60'ı kırmızıdır. Buna göre, bu kutudan rassal olarak çekilen 50 bilyeden kaç tanesinin kırmızı olması beklenir?

KOŞULLU OLASILIK VE ÇARPMA KURALI

Gelecek yıl faizlerin düşeceği bilindiğine göre, ev fiyatlarının artması olasılığı nedir? Çekilen ilk topun beyaz olduğu görüldüğüne göre ikinci topun siyah olması olasılığı nedir? Bir sınıfta soru soran öğrencinin erkek olduğu bilindiğine göre onun Ahmet olması olasılığı nedir?

Yukarıda ifade edilen sorulara, koşullu olasılık ile cevaplar aranır. Koşullu olasılıktaki temel mantık şöyle ifade edilebilir: *Bir olayın gerçekleşme olasılığı, başka bir olayın gerçekleşmesine bağlıdır.* Diğer bir deyişle, koşullu olasılık, bir olayın gerçekleştiğinin bilinmesi durumunda diğer bir olayın gerçekleşme olasılığıdır.

Bir B olayının gerçekleştiği bilindiğine göre A olayının olasılığı, koşullu olasılık olarak ifade edilir ve $P(A|B)$ şeklinde gösterilir. Koşullu olasılık:

$$P(A|B)=\frac{P(A\cap B)}{P(B)},\quad (P(B)\neq 0)$$

şeklinde hesaplanır.

Bir B olayının gerçekleştiği bilindiğine göre A olayının olasılığı, koşullu olasılık olarak ifade edilir ve $P(A|B)$ şeklinde gösterilir. Koşullu olasılık:

$$P(A|B)=\frac{P(A\cap B)}{P(B)},$$

$$(P(B)\neq 0)$$

şeklinde hesaplanır.

Benzer olarak, bir A olayının gerçekleştiği bilindiğine göre B olayının olasılığı, koşullu olasılık olarak ifade edilir ve $P(B|A)$ şeklinde gösterilir. Koşullu olasılık

$$P(B \mid A) = \frac{P(A \cap B)}{P(B)}, \quad (P(A) \neq 0)$$

şeklinde hesaplanır.

DİKKAT

Olasılık sorularında "görüldüğüne göre", "bilinmesi durumunda", gibi şart ifadesi ile karşılaşıldığında bu soruların çözümünde koşullu olasılık formülünü kullanmak uygun olacaktır.

Örnek 1: *1'den 10'a kadar (10 dahil) olan tam sayılar arasından rassal olarak seçilen bir sayının 3 ile bölündüğü bilindiğine göre bu sayının 2 ile bölünme olasılığı nedir?*

Çözüm: İlgili iki olay aşağıdaki gibi tanımlanır:

A: Rassal olarak seçilen bir sayının 2 ile bölünmesi,
B: Rassal olarak seçilen bir sayının 3 ile bölünmesi.

Bu durumda,

$A \cap B$: Rassal olarak seçilen bir sayının 2 ve 3 ile bölünmesi

olur. Açıktır ki, $A=\{2,4,6,8,10\}$, $B=\{3,6,9\}$ ve $A \cap B=\{6\}$'dır. A olayı içinde $n_A=5$ sayıda, B olayı içinde $n_B=3$ sayıda, $A \cap B$ olayı içinde $n_{A \cap B}=1$ sayıda örnek nokta vardır ve $n=10$ olduğuna göre, bu olaylara ilişkin olasılıklar (Olaylar için klasik olasılık kuralı ile) :

$$P(A) = \frac{n_A}{n} = \frac{5}{10} = 0.50$$

ve benzer olarak

$$P(B) = \frac{n_B}{n} = \frac{3}{10} = 0.30 \text{ ve } P(A \cap B) = \frac{n_{A \cap B}}{n} = \frac{1}{10} = 0.10$$

olarak bulunur.

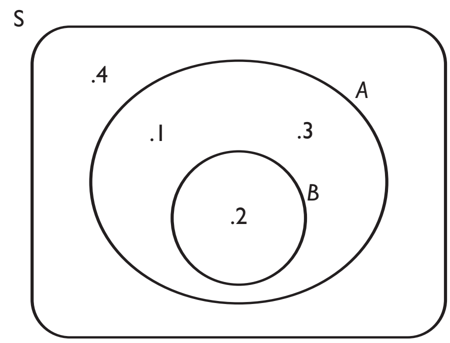
Rassal olarak seçilen bir sayının 3 ile bölündüğü bilindiğine göre bu sayının 2 ile bölünme olasılığı, koşullu olasılık yardımıyla

$$P(A \backslash B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} = \frac{0.10}{0.30} = \frac{1}{3} = 0.333$$

olarak hesaplanır. Rassal olarak seçilen bir sayının 3 ile bölündüğü bilindiğine göre bu sayının 2 ile bölünme olasılığı %33.3'tür.

Örnek 2: *S bir deneyin örnek uzayı, A ve B olayları da aşağıdaki şekilde verildiği gibidir. Tüm örnek noktalar eşit olasılığa sahip ise $P(A|B)$ ve $P(B|A)$ olasılıklarını hesaplayınız.*

Şekil 4.1



S örnek uzayı ve A ile B olayları

Çözüm: $A=\{1,2,3\}$, $B=\{2\}$ ve $A\cap B=\{2\}$ şeklindedir. Dikkat edilirse $B\subseteq A$ ve $A\cap B=B$ dir. $S=\{1,2,3,4\}$ ve $n=4$'tür. Tüm örnek noktalar eşit olasılığa sahip olduğu için

$$P(A)=\frac{n_A}{n}=\frac{3}{4}=0.75$$

$$P(B)=\frac{n_B}{n}=\frac{1}{4}=0.25$$

$$P(A\cap B)=\frac{n_{A\cap B}}{n}=\frac{1}{4}=0.25$$

Böylece

$$P(A|B)=\frac{P(A\cap B)}{P(B)}=\frac{0.25}{0.25}=1 \text{ ve } P(B|A)=\frac{P(A\cap B)}{P(A)}=\frac{0.25}{0.75}=0.333$$

şeklinde bulunur.

Örnek 3: *A ve B olayları ayrık ve olasılıkları $P(A)=0.20$ ve $P(B)=0.30$ olarak verilsin. Buna göre $P(A|B)$ olasılığını hesaplayınız.*

Çözüm: A ve B olayları ayrık ise $A\cap B=\varnothing$ dir. Böylece, $P(A\cap B)=P(\varnothing)=0$'dır.

$$P(A|B)=\frac{P(A\cap B)}{P(B)}=\frac{0}{0.30}=0$$

Örnek 4: *Bir fabrikada bir günde üretilen 1000 ürünün 400'ü A ve 600'ü B makinasında yapılmaktadır. A'da üretilenlerin %1'i ve B'de üretilenlerin %2'si kusurlu üründür. Bu fabrikada üretilen ürünlerden rassal olarak biri çekilmiş ve kusurlu olduğu görülmüştür. Buna göre bu kusurlu ürünün A makinasında üretilen bir ürün olması olasılığı nedir?*

Çözüm: Öncelikle olaylar, rassal olarak seçilen bir ürünün

A: *A makinasında üretilen ürün olması,*
B: *B makinasında üretilen ürün olması,*
C: *Kusurlu ürün olması*

şeklinde tanımlansın. Bu durumda,

$A \cap C$: *A makinasında ve kusurlu üretilen ürün olması*

olayı olur.

$$P(A) = \frac{400}{1000} = 0.40,\ P(B) = \frac{600}{1000} = 0.60$$

şeklindedir. C olayının olasılığı ise toplam üretim içinde kusurlu oranı bulunarak elde edilebilir. Bu durumda; A makinasında kusurlu üretim 400(1/100)=4, B makinasında kusurlu üretim 600(2/100)=12 olduğuna göre, toplam üretimde yani 1000 ürün içinde kusurlu ürün sayısı 4+12=16'dır. Buradan,

$$P(C) = \frac{16}{1000} = 0.016$$

ve

$$P(A \cap C) = \frac{4}{1000} = 0.004$$

olarak elde edilir. Rassal olarak çekilen bir ürünün kusurlu olduğu görüldüğüne göre ürünün A makinasında üretilen bir ürün olması olasılığı $P(A|C)$ olup,

$$P(A|C) = \frac{P(A \cap C)}{P(C)} = \frac{0.004}{0.016} = 0.25$$

olarak bulunur.

Örnek 4'te verilenler, iki yönlü bir tabloya dönüştürülerek de istenen olasılıklar hesaplanabilir. 1000 ürün makine türlerine (A ve B) ve üretimin sonucuna (kusurlu ve kusursuz) göre sınıflandırılır. Buna göre 1000 ürünün 400'ü A ve 600'ü B makinasında yapılmaktadır. A'da üretilenlerin %1'i ve B'de üretilenlerin %2 'si kusurlu ürün olduğunu göre 400.(1/100)=4 ve 600.(2/100) =12 olmak üzere A'da 4 ve B'de 12 ürün kusurludur. Buna göre 400-4=396 ve 600-12=588 olmak üzere A'da 396 ve B'de 588 ürün kusursuzdur. Böylece örnekte verilen 1000 ürünün dağılımı Tablo 4.1 şeklindedir.

Tablo 4.1
Makineler ve Üretimin durumuna ilişkin sonuçlar

Durum / Makine Cinsi	Kusurlu	Kusursuz	Toplam
A	4	396	400
B	12	588	600
Toplam	16	984	1000

Bu tür tablolara *çapraz tablo* (contingeny table) denmektedir. Çapraz tablo, değişkenlerin frekans (sıklık) dağılımını gösteren matris biçiminde tablodur. Çapraz tabloda sayıların bulunduğu kutulara göze ya da hücre adı verilmektedir. Örneğin

A makinesinde üretilen ve kusurlu ürün sayısını gösteren göze 4 sayısının bulunduğu gözedir. Benzer olarak, B makinesinde üretilen ve kusursuz ürün sayısını gösteren göze 588 sayısının bulunduğu gözedir. Dolayısıyla rassal olarak seçilen bir ürünün A makinesinde üretilen ve kusurlu ürün olması olasılığı 4/1000'dir. Aynı şekilde B makinesinde üretilen ve kusursuz ürün olması olasılığı 588/1000'dir. Burada ifade edilen olaylarda, iki olayın arakesiti söz konusudur. Böyle olayların olasılığına **bileşik olasılık** adı verilmektedir. Böylece, A ve B gibi iki olayın arakesitinin (kesişiminin) olasılığı *bileşik olasılık* olarak ifade edilmektedir. Bileşik olasılık $P(A \cap B)$ biçiminde gösterilir.

Rassal olarak seçilen bir ürünün A makinesinde üretilmesi ve kusurlu olması olasılığı, Tablo 4.1 yardımıyla

$$P(A \cap C) = \frac{4}{1000} = 0.004$$

şeklinde, rassal olarak seçilen bir ürünün kusurlu olması olasılığı ise yine Tablo 4.1 yardımıyla

$$P(C) = \frac{16}{1000} = 0.016$$

şeklinde hesaplanır. Böylece;

$$P(A \mid C) = \frac{P(A \cap C)}{P(C)} = \frac{0.004}{0.016} = 0.25$$

olarak bulunur.

Marjinal (bileşen) olasılık ise herhangi başka olay dikkate alınmaksızın sadece bir olaya ilişkin olasılıktır. Yukarıdaki örnek dikkate alındığında, rassal olarak seçilen bir ürünün kusurlu olması olasılığı, marjinal olasılığa örnektir. Benzer olarak, yalnızca makine türü dikkate alındığında, rassal olarak seçilen bir ürünün A makinasında üretilen ürün olması olasılığı, yine marjinal olasılığa örnektir. Çapraz tabloda marjinal olasılıklar satır veya sütun toplamlarının toplam örneklem büyüklüğüne bölünmesi ile bulunur.

Marjinal (bileşen) olasılık, herhangi başka olay dikkate alınmaksızın sadece bir olaya ilişkin olasılıktır.

Tablo 4.1, ürünün kusurlu ya da kusursuz olması türüne göre yazıldığında Tablo 4.2 elde edilmektedir. Böylece Tablo 4.1'de satırlar toplandığında Tablo 4.2 elde edilir.

Tablo 4.2
Ürünün kusurlu veya kusursuz olması durumuna göre sınıflama

Kusurlu	Kusursuz	Toplam
16	984	1000

Tablo 4.1, ürünün A ya da B makinesinde üretilmesine göre yazıldığında Tablo 4.3 elde edilmektedir. Böylece Tablo 4.1'de sütunlar toplandığında Tablo 4.3 elde edilir.

Tablo 4.3
Ürünün A veya B makinasında üretilmiş olması durumuna göre sınıflama

A makinası	B makinası	Toplam
400	600	1000

Rassal olarak bir ürün seçildiğinde

A: *A makinasında üretilen ürün olması,*
C: *Kusurlu ürün olması,*

olayları tanımlandığında Tablo 4.2 ve Tablo 4.3'ten yararlanarak A ve C olaylarının olasılıkları kolayca hesaplanabilir.

E: A makinasında üretilen ve kusursuz ürün olması
F: B makinasında üretilen ve kusurlu ürün olması

olayları tanımlandığında, Tablo 4.1'den yararlanarak E ve F olaylarının olasılıkları kolayca hesaplanabilir.

Burada A ve C olayının olasılığı, *marjinal olasılığa* örnek iken E ve F olaylarının olasılığı *bileşik olasılığa* örnektir. Yukarıda belirtildiği gibi, A makinasında üretilen ürün sayısının (Tablo 4.1'de sütun toplamlarının yani Tablo 4.3) toplam ürün sayısına bölünmesiyle $P(A)$ aşağıdaki gibi

$$P(A)=400/1000=0.40$$

elde edilir. C olayının olasılığı, kusurlu ürün sayısının (Tablo 4.1'de satır toplamlarının yani Tablo 4.2) toplam ürün sayısına bölünmesi ile elde edilmektedir. Bu durumda,

$$P(C)=16/1000=0.016$$

olarak hesaplanır. E ve F olaylarının olasılığı ise, Tablo 4.1'de ilgili hücredeki sayının toplam ürün sayısına bölünmesi ile elde edilir. Bu durumda, A makinasında üretilen kusursuz ürün sayısını gösteren gözede 396 sayısının toplam ürün sayısı olan 1000'e bölünmesi ile E olayının olasılığı,

$$P(E)=396/1000=0.396$$

olarak bulunur. Benzer olarak F olayının olasılığı, B makinasında ve kusurlu ürün sayısını gösteren gözedeki 12 sayısının toplam ürün sayısı olan 1000'e bölünmesi ile elde edilir. Bu durumda,

$$P(F)=12/1000=0.012$$

dir.

Örnek 5: *Bir teknik servis, tamir için gelen 500 adet elektrikli bir cihazı incelemiş ve elde ettiği sonuçları aşağıdaki tabloda özetlemiştir.*

Tablo 4.4
Cihaza ilişkin sonuçlar

	Sorunun Kaynağı		
	Elektrik Aksamı	**Mekanik Aksamı**	**Toplam**
Garanti kapsamı dışında	190	125	315
Garanti kapsamı içinde	70	115	185
Toplam	260	240	500

Tamire gelen bir cihaz rassal olarak seçildiğinde

A: Sorununun mekanik aksamından olması,
B: Sorununun garanti kapsamı dışında olması,
C: Sorununun mekanik aksamından ve garanti kapsamı dışında olması,
D: Garanti kapsamı dışında olduğu bilindiğine göre sorununun mekanik aksamından olması

olayları tanımlanmaktadır. Buna göre, burada tanımlanan her bir olayın olasılığını bulunuz.

Çözüm: Tablo 4.4'te verilen sayılardan yararlanılarak istenen olasılıklar kolaylıkla bulunabilir. Cihazların 240 tanesi mekanik aksamından dolayı sorunludur. Dolayısıyla

$P(A)=240/500=0.48$

dir.

Garanti kapsamı dışında 315 cihaz sorunludur. Böylece

$P(B)=315/500=0.63$

olur.

İlgili gözeye bakıldığında, mekanik aksamından sorunlu ve garanti kapsamı dışında olan cihaz sayısının 125 olduğu görülmektedir. Böylece,

$P(C)= P(A\cap B)=125/500=0.25$

olarak hesaplanır.

Cihazın sorununun garanti kapsamı dışında olduğu bilindiğine göre mekanik aksamından olması olasılığı, koşullu olasılık yardımıyla:

$$P(D) = P(A\,|\,B) = \frac{P(A\cap B)}{P(B)} = \frac{0.25}{0.63} = 0.397$$

şeklinde elde edilir.

Örnek 6: *Bir kişinin bir şirkete iş başvurusuna ilişkin olaylar, aşağıdaki gibi tanımlansın.*

K: İşe alınması
R: İşe alınmaması

buna göre $P(K\cap R)=?$

Çözüm: Soruda tanımlanan K ve R olaylarının birlikte ortaya çıkmayacağı açıktır. Dolayısıyla bu olaylar **ayrık olaylar**dır. $K\cap R = \varnothing$ ve $P(K\cap R)=0$ dır. Böylece K ve R olaylarının bileşik olasılığı $P(K\cap R) = 0$'dır.

A ve *B* **ayrık olaylar** ise bu olayların bileşik olasılığı 0 dır. Bu durum $P(A\cap B) = 0$ biçiminde gösterilir.

Çarpma Kuralı

Koşullu olasılık yardımıyla iki olayın arakesitinin (kesişiminin) olasılığı hesaplanabilir. B olayı verildiğinde, A olayının olasılığı,

$$P(A\,|\,B) = \frac{P(A\cap B)}{P(B)}, (P(B) \neq 0)$$

dır. Benzer şekilde, A olayı verildiğinde, B olayının olasılığı,

$$P(B\,|\,A) = \frac{P(A\cap B)}{P(A)}, (P(A) \neq 0)$$

şeklindedir. Yukarıdaki eşitliklerden içler dışlar çarpımı yapılırsa A ve B olaylarının arakesitinin olasılığı, bir başka ifade ile A ve B'nin bileşik olasılığı

$P(A\cap B)=P(B)\; P(A\,|\,B)$ veya $P(A\cap B)=P(A)\; P(B\,|\,A)$

dir. Buna **çarpma kuralı** adı verilir.

Çarpma kuralı, *A* ve *B* olaylarının ara kesitinin olasılığı $P(A\cap B)=P(B)\,P(A\,|\,B)$ veya $P(A\cap B)=P(A)\,P(B\,|\,A)$ dir.

Örnek 7: *P (B)=0.30 ve P (A|B)=0.60 değerleri veriliyor ise A ve B olaylarının bileşik olasılığı kaçtır?*

Çözüm:

$P(A|B)=\dfrac{P(A\cap B)}{P(B)}$ olduğu için $P(A\cap B)=P(B)\ P(A|B)$ dir. Soruda verilen değerler en son eşitlikte yerine yazılırsa $P(A\cap B)$=(0.30) (0.60)= 0.18 olarak bulunur. Böylece A ve B olaylarının bileşik olasılığı $P(A\cap B)$ = 0.18'dir.

Örnek 8: *Bir ikinci el mağazasındaki 12 televizyondan 5 tanesi bozuktur. Bu televizyonlardan* ***seçilen yerine koyulmaksızın*** *arka arkaya iki tanesi rastgele seçilmiştir. Seçilen televizyonlardan birincisinin bozuk ve ikincisinin sağlam olması olasılığını bulunuz.*

Çözüm: Olaylar

A: Rastgele seçilen birinci televizyonun bozuk olması,
B: Rastgele seçilen ikinci televizyonun sağlam olması

olarak tanımlanırsa

$A\cap B$: *Rastgele seçilen birinci televizyonun bozuk ve ikincinin sağlam olması*

şeklinde olur.

$$P(A)=\frac{5}{12}$$

olduğu açıktır. Seçilen birinci televizyonun bozuk olduğu bilindiğine göre geriye 11 televizyon kalmış, bunlardan 4 tanesi bozuk ve 7 tanesi sağlamdır. İlk seçilen televizyonun bozuk olduğu bilindiğine göre ikinci televizyonun sağlam olması olasılığı

$$P(B|A)=\frac{7}{11}$$

tür. Böylece çarpma kuralından

$$P(A\cap B)=P(A)P(B|A)=\left(\frac{5}{12}\right)\left(\frac{7}{11}\right)=\frac{35}{132}=0.265$$

şeklindedir.

Örnek 9: *Bir kutuda 3 mavi ve 3 sarı top vardır. Bu kutudan* ***yerine konmaksızın*** *iki top çekilmiştir. Birincinin mavi ve ikincinin sarı olması olasılığı nedir?*

Çözüm: Olaylar,

A: Rastgele seçilen birinci topun mavi olması,
B: Rastgele seçilen ikinci topun sarı olması

olarak tanımlanırsa

$A\cap B$= *Rastgele seçilen birinci topun mavi, ikinci topun sarı olması*

şeklinde olur. Bu durumda,

$$P(A)=\frac{3}{6}$$

olduğu açıktır. Seçilen birinci topun mavi olduğu bilindiğine göre geriye 5 top kalmıştır ve bunlardan 2 tanesi mavi, 3 tanesi sarıdır. Bu durumda, seçilen birinci topun mavi olduğu bilindiğine göre ikinci topun sarı olması olasılığı

$$P(B\mid A)=\frac{3}{5}$$

tir. Böylece çarpma kuralından

$$P(A\cap B)=P(A)P(B\mid A)=\left(\frac{3}{6}\right)\left(\frac{3}{5}\right)=\frac{9}{30}=0.3$$

elde edilir.

Örnek 10: *Bir kutuda 4 kırmızı ve 6 siyah bilye vardır. Bu kutudan çekilen* ***yerine konarak*** *iki bilye çekilmiştir. Birincinin siyah, ikincinin kırmızı olması olasılığı nedir?*

Çözüm: Olaylar,

A: Rastgele seçilen birinci bilyenin siyah olması,
B: Rastgele seçilen ikinci bilyenin kırmızı olması

olarak tanımlanırsa
A∩B= Rastgele seçilen birinci bilyenin siyah ve ikincisinin kırmızı olması
şeklinde olur. Bu durumda,

$$P(A)=\frac{4}{10}$$

olduğu açıktır. Birinci seçilen siyah bilye tekrar kutuya geri konmuştur. Kutudaki bilyelerde değişiklik olmamıştır. Dolayısıyla seçilen birinci bilyenin siyah olduğu bilindiğine göre ikinci bilyenin kırmızı olması olasılığı

$$P(B\mid A)=\frac{6}{10}$$

tir. Böylece çarpma kuralından

$$P(A\cap B)=P(A)P(B\mid A)=\left(\frac{4}{10}\right)\left(\frac{6}{10}\right)=\frac{24}{100}=0.24$$

elde edilir.

Bağımsız Olaylar

Hilesiz bir zarın art arda atılması deneyi ele alınsın ve

A: Birinci zarın 6 gelmesi,
B: İkinci zarın 6 gelmesi

şeklinde tanımlansın. Birinci zarın 6 geldiği görüldüğüne göre ikinci zarın 6 gelmesi olasılığı nedir? *A* olayının gerçekleşmesi, *B* olayının olasılığını etkiler mi? Böyle bir soruda, sezgisel olarak bilinir ki bu iki olay bağımsızdır. Çünkü *A*'nın gerçekleşmesi, B'nin olasılığını etkilemez. Bir başka deyişle, A'nın gerçekleşmesi B'nin olasılığını ne arttırır ne de azaltır. B olayının olasılığı, birinci zarın sonucu ne olursa olsun her zaman 1/6'dır.

Bir başka örnek olarak

A : Bugün İMKB 100 endeksinin yükselmesi,
B : Madrid'de yağmur yağması

şeklinde tanımlansın. Bu olayların bağımsız olaylar olduğu da açıktır. Çünkü A'nın gerçekleşmesinin B'nin gerçekleşmesini etkilemeyeceği, B'nin gerçekleşmesinin de A'nın gerçekleşmesini etkilemeyeceği açıktır.

A ve B bağımsız olaylar ise $P(A|B) = P(A)$ ve $P(B|A) = P(B)$ dir.

Yukarıda her iki örnekte tanımlanan A ve B olayları bağımsız olaylara örnektir. A ve B bağımsız olaylar ise

$$P(A|B) = P(A) \text{ ve } P(B|A) = P(B)$$

A ve B bağımsız olaylar ise $P(A \cap B) = P(A)\,P(B)$ dir.

dir. Dolayısıyla, A'nın ortaya çıkması B'nin ortaya çıkmasını etkilemez. B'nin ortaya çıkması da, A'nın ortaya çıkmasını etkilemez. Bu durumda, bağımsız olaylar için çarpma kuralı ise

$P(A \cap B) = P(A)\,P(B)$ ise A ve B olayları bağımsızdır.

$$P(A \cap B) = P(A)\,P(B) \text{ veya } P(A \cap B) = P(B)\,P(A)$$

şeklindedir.

Ayrıca $P(A \cap B) = P(A)P(B)$ eşitliği geçerli ise A ve B olayları *bağımsızdır* denir.

Çarpma kuralı, ikiden fazla bağımsız olaylar için genelleştirilebilir. A, B ve C üç bağımsız olay olsun. Çarpma kuralı:

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

şeklinde olur.

Örnek 11: *Hilesiz (Dengeli ve düzgün) bir para iki kez atılmıştır.*
A: Birinci atışta yazı gelmesi
B: İkinci atışta yazı gelmesi
şeklinde tanımlanırsa A ve B olaylarının bağımsız olduğunu gösteriniz.

Çözüm: Bu deneyin örnek uzayı $S=\{YY, YT, TY, TT\}$ şeklindedir.

$A=\{YY, YT\}$ ve $B=\{TY, YY\}$, $P(A) = \frac{2}{4}$ ve $P(B) = \frac{2}{4}$

dir. $A \cap B = \{YY\}$ ve $P(A \cap B) = \frac{1}{4}$ tür.

A ve B olayları bağımsız ise $P(A \cap B) = P(A)\,P(B)$ olmalıdır.

$$P(A \cap B) = \frac{1}{4} = P(A)P(B) = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{4} \text{ tür.}$$

A ve B olayları bağımsızdır.

Örnek 12: *Hilesiz bir para iki kez atılmıştır. Birinci atışın yazı geldiği bilindiğine göre ikinci atışın yazı gelmesi olasılığı nedir?*

Çözüm: Bir önceki sorudaki olaylar ele alındığında, birinci atışın yazı geldiği bilindiğine göre ikinci atışın yazı gelmesi olasılığı

$$P(B|A) = P(B)$$

şeklinde olur. Çünkü olaylar bağımsızdır. Böylece

$$P(B \mid A) = P(B) = \frac{1}{2}$$

dir.

Örnek 13: *İki farklı öğrencinin İktisat-İşletme Fakültelerinde okutulmakta olan Genel Muhasebe dersinden başarılı olma olasılıkları, sırasıyla 0.8 ve 0.7'dir. Bu iki öğrencinin her ikisinin de Genel Muhasebe dersinden başarılı olma olasılığını bulunuz.*

Çözüm: Bu soru için olaylar

A: Birinci öğrencinin başarılı olması
B: İkinci öğrencinin başarılı olması
A∩B: Her iki öğrencinin de başarılı olması

şeklinde tanımlansın. Açıktır ki A ve B olayları bağımsızdır. Çünkü bir öğrencinin başarısı, diğer öğrencinin başarısını ya da başarısızlığını etkilememektedir.

A ve B olayları bağımsız olduklarından, A ve B'nin birlikte gerçekleşmesi olasılığı,

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

formülü ile hesaplanır. Buradan istenen olasılık,

$$P(A \cap B) = (0.8)(0.7) = 0.56$$

olarak bulunur.

Örnek 14: *Belli bir antibiyotik ilacın çocuklarda alerji yapması olasılığı 1/100'dür. Bu ilacın üç çocuğa verildiği durumda üçünün de alerji olması olasılığı nedir?*

Çözüm: Çocuklardan birinin antibiyotik ilaçtan alerji olması diğerini etkilemeyeceği için A, B, C sırasıyla birinci, ikinci ve üçüncü çocuğun alerji olması olaylarını göstermek üzere A, B, C olayları bağımsızdır. Buna göre,

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

olduğundan

$$P(A \cap B \cap C) = (1/100)(1/100)(1/100) = 1/100^3$$

dir.

Örnek 15: *Hilesiz bir paranın 4 kez atılması deneyinde, dört atışta da yazı gelmesi olasılığı kaçtır?*

Çözüm: Dört atışın her birinde yazı gelmesi olayının olasılığı 1/2 dir. Olaylar birbirinden bağımsız olduğu için dört atışta da yazı gelmesi olasılığı (1/2) (1/2) (1/2) (1/2)=1/16'dır.

Bağımlı Olaylar

Eğer

$$P(B|A) \neq P(A) \text{ ve } P(B|A) \neq P(B)$$

ise A ve B *bağımlı* olaylardır. Başka bir ifade ile,

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

ise A ve B *bağımlı* olaylardır.

DİKKAT

$P(A \cap B) = P(A)\ P(B)$ eşitliği A ve B olaylarının bağımsız olup olmadığını araştırmada kullanılmasının yanı sıra bağımsız olduğu bilinen A ve B olaylarının arakesit olasılığını yani $P(A \cap B)$'ını, tekil olasılıkların yani $P(A)$ ve $P(B)$'nin çarpımı olarak hesaplanmasında da kullanılmaktadır.

Örnek 16: *$P(A) = \frac{1}{2}$, $P(B) = \frac{1}{2}$ ve $P(A \cap B) = \frac{1}{2}$ dir. Buna göre A ve B olayları bağımsız mıdır ?*

Çözüm: $P(A \cap B) \neq P(A)\ P(B)$ olduğundan bağımsız değildir. Bu olaylar bağımlıdır.

Örnek 17: *Bir şirkette çalışanların %70'i erkek, %30'u kadındır. Erkeklerin %50'si lise, %40'ı üniversite, %10'u yüksek lisans mezunu, kadınların %60'ı lise, %30'u üniversite, %10'u yüksek lisans mezunudur. Rastgele seçilen bir çalışanın*

a) Erkek ve lise mezunu olması olasılığı nedir?
b) Lise mezunu olması olasılığı nedir?
c) Kadın olduğu bilindiğine göre, yüksek lisans mezunu olması olasılığı nedir?
d) Kadın olması ve üniversite mezunu olması olayları bağımsız mıdır?

Çözüm: Örnekte verilen bilgiler kullanılarak aşağıdaki tablo oluşturulabilir ve aranan olasılıklar bu tablodan yararlanılarak kolayca hesaplanabilir. Şirkette çalışanların sayısının 100 olduğu kabul edilirse şirkette çalışanların (70/100)=%70'i erkek, erkeklerin (50/100)=%50'si lise mezunu ise lise mezunu erkeklerin toplam çalışan içinde oranı ise (70/100)(50/100)=0.35'tir. Yani 100 çalışanın bulunduğu bir şirkette erkek ve lise mezunu çalışan sayısı 35'tir. 100 çalışanın olduğu şirketin %30'u kadın yani 30 çalışan kadın, kadınların %60'ı lise mezunu, (30/100)(60/100)=0.18 yani %18 lise mezunu ve kadın oranıdır. Böylece, 100 çalışanın olduğu şirkette 18 çalışan, kadın ve lise mezunudur. Benzer şekilde, tablonun tüm hücreleri doldurulur.

Tablo 4.5
Çalışanların cinsiyet ve eğitim durumlarına ilişkin sonuçlar

	Lise (*L*)	**Üniversite (*Ü*)**	**Yüksek lisans (*Y*)**	**Toplam**
Erkek (*E*)	35	28	7	70
Kadın (*K*)	18	9	3	30
Toplam	53	37	10	100

a) Rastgele seçilen bir çalışanın erkek (E) ve lise (L) mezunu olması olasılığı,

$$P(E \cap L) = \frac{35}{100} = 0.35$$

olarak bulunur.

b) Tablo 4.5 dikkate alındığında, toplam lise mezunu sayısı 53, toplam şirket çalışanı sayısı 100 olduğuna göre, rastgele seçilen bir çalışanın lise mezunu olması olasılığı,

$$P(L) = \frac{53}{100} = 0.53$$

olarak bulunur.

c) Rastgele seçilen bir çalışanın kadın (*K*) olması olasılığı

$$P(K) = \frac{30}{100} = 0.30 \text{ ve}$$

rastgele seçilen birinin kadın (*K*) ve yüksek lisans mezunu (*Y*) olması olasılığı

$$P(K \cap Y) = \frac{3}{100} = 0.03 \text{ biçiminde ve dolayısıyla}$$

$$P(Y \mid K) = \frac{P(K \cap Y)}{P(K)} = \frac{0.03}{0.30} = 0.10$$

olarak bulunur. Sonuç olarak, rastgele seçilen bir çalışanın kadın olduğu bilindiğine göre, yüksek lisans mezunu olması olasılığı 0.10'dur.

d) Tablo 4.5'ten görüleceği gibi, rastgele seçilen bir çalışanın "*üniversite mezunu olması*", "*kadın olması*" ve "*üniversite mezunu ve kadın olması*" olasılıkları sırası ile

$$P(Ü) = \frac{37}{100} = 0.37,\ P(K) = \frac{30}{100} = 0.30 \text{ ve } P(K \cap Ü) = \frac{9}{100} = 0.09 \text{ dur.}$$

K ve $Ü$ olaylarının bağımsız olup olmadıklarını belirlemek için

$$P(K \cap Ü) = P(K)\, P(Ü)$$

eşitliğinin sağlanıp sağlanmadığının kontrol edilmesi gerekir.

$$P(K)P(Ü) = \left(\frac{30}{100}\right)\left(\frac{37}{100}\right) = \frac{1110}{10000} = 0.111$$

olduğu için

$$P(K \cap Ü) \neq P(K)\, P(Ü) \text{ dir.}$$

K ve $Ü$ olayları bağımsız değildir. Dolayısıyla K ve $Ü$ olayları bağımlı olaylardır.

İki olayın bağımsız olup olmadığını araştırmanın diğer bir yolu bağımsız olaylar için verilen tanımı kullanmaktır. Diğer bir deyişle K ve $Ü$ olayları bağımsız ise

$$P(K \mid Ü) = P(K) \text{ veya } P(Ü \mid K) = P(Ü)$$

şeklinde olmalıdır.

$$P(K \mid Ü) = \frac{P(K \cap Ü)}{P(Ü)} = \frac{0.09}{0.37} = 0.243 \text{ ve } P(K) = \frac{30}{100} = 0.30$$

ve dolayısıyla $P(K \mid Ü) \neq P(K)$ olduğu için K ve $Ü$ olayları bağımsız değildir. Dolayısıyla K ve $Ü$ olayları bağımlı olaylardır.

Benzer olarak

$$P(Ü \mid K) = \frac{P(K \cap Ü)}{P(K)} = \frac{0.09}{0.30} = 0.30 \text{ ve } P(Ü) = \frac{37}{100} = 0.37$$

dolayısıyla $P(Ü \mid K) \neq P(Ü)$ olduğu için K ve $Ü$ olayları bağımsız değildir.

Örnek 18: *Bir deterjan firmasının, bulaşık makineleri için ürettiği özel bir deterjan ve bununla birlikte kullanılmasını önerdiği makine parlatıcısı ürünleri piyasada satılmaktadır. Ürünlerin satıldığı bir satış mağazasında 250 müşteri üzerinde yapılan bir değerlendirmede aşağıdaki sonuçlar elde edilmiştir.*

Tablo 4.6
Bulaşık deterjanı ve parlatıcı ürünlerine ilişkin sonuçlar

		Bulaşık deterjanı		
		Alan	**Almayan**	**Toplam**
Parlatıcı	**Alan**	120	40	160
	Almayan	20	70	90
	Toplam	140	110	250

Rastgele seçilen bir müşterinin

a) Bulaşık deterjanı alma olasılığını bulunuz.
b) Bulaşık deterjanı ve parlatıcı alma olasılığını bulunuz.
c) Bulaşık deterjanı ve parlatıcı ***almama*** *olasılığını bulunuz.*
d) Bulaşık deterjanı aldığı bilindiğine göre parlatıcı alma olasılığını bulunuz.
e) Parlatıcı aldığı bilindiğine göre bulaşık deterjanı ***almama*** *olasılığını bulunuz.*
f) Bulaşık deterjanı alma olayı ile parlatıcı alma olayları bağımsız mıdır?

Çözüm:

B: *Rastgele seçilen bir müşterinin bulaşık deterjanı alması*
P: *Rastgele seçilen bir müşterinin parlatıcı alması*
$B \cap P$: *Rastgele seçilen bir müşterinin bulaşık deterjanı ve parlatıcı alması*
olayları tanımlansın.

a) 250 müşteriden 140'ı bulaşık deterjanı almıştır. Dolayısıyla rastgele seçilen bir müşterinin bulaşık deterjanı alma olasılığı $P(B)$=140/250=0.56'dır.

b) 250 müşteriden 120'si bulaşık deterjanı ve parlatıcı almıştır. Dolayısıyla rastgele seçilen bir müşterinin bulaşık deterjanı ve parlatıcı alma olasılığı

$P(B \cap P)$=120/250=0.48'dir.

c) 250 müşteriden 70'i ne bulaşık deterjanı ne de parlatıcı almıştır. Dolayısıyla rastgele seçilen bir müşterinin bulaşık deterjanı ve parlatıcı *almaması* olasılığı 70/250=0.28'dir.

d) Rastgele seçilen bir müşterinin bulaşık deterjanı aldığı bilindiğine göre parlatıcı alma olasılığı:

$$P(P \mid B) = \frac{P(P \cap B)}{P(B)} = \frac{120/250}{140/250} = 0.857$$

e) Yukarıda tanımlanan B ve P olaylarının yanısıra

$\overline{B}$: *Rastgele seçilen bir müşterinin bulaşık deterjanı* ***almaması***
$P \cap \overline{B}$: *Rastgele seçilen bir müşterinin parlatıcı alması ve bulaşık deterjanı* ***almaması***

olayları tanımlandığında, bu olaylara ilişkin olasılıklar:

$P(P)$=160/250 ve $P(P \cap \overline{B})$ =40/250'dir.

Rastgele seçilen bir müşterinin parlatıcı aldığı bilindiğine göre bulaşık deterjanı almaması olasılığı

$$P(\bar{B} \mid P) = \frac{P(P \cap \bar{B})}{P(P)} = \frac{40 / 250}{160 / 250} = 0.25$$

dir.

f) $P(B)$=140/250, $P(P)$=160/250 ve $P(B \cap P)$=120/250 ve
$P(B)\ P(P)$=0.358 ve $P(B \cap P)$=0.48 olduğu dikkate alınırsa
$P(P \cap B) \neq P(P)\ P(B)$

olduğu görülür. P ve B olayları bağımsız değildir. Yani bulaşık deterjanı alma olayı ile parlatıcı alma olayları bağımsız değildir.

Örnek 19: *Aşağıdaki tablo belli bir bölgedeki halkın kan gruplarına ilişkin olasılıklarını göstermektedir.*

	A	B	AB	0	Toplam
Rh+	0.30	0.10	0.05	0.40	0.85
Rh-	0.05	0.04	0.01	0.05	0.15
Toplam	0.35	0.14	0.06	0.45	1

Tablo 4.7
Kan grupları ve olasılıkları

Bu bölgeden, rastgele bir kişi seçildiğinde
a) A grubunda olması
b) AB grubunda ve Rh- olması
c) Seçilenin 0 grubu ve Rh+ olması
d) Seçilenin 0 grubu olduğu bilindiğine göre Rh+ olması
e) Seçilenin Rh- olması
f) Seçilenin Rh- olduğu bilindiğine göre B grubu olması
olasılıklarını hesaplayınız.

Çözüm: Yukarıdaki tabloda iki olayın arakesitine ilişkin olasılıklar yani bileşik olasılıklar verilmiştir. Örneğin ilk gözede verilen 0.30 sayısı bu bölgeden rastgele bir kişi seçildiğinde A grubunda ve Rh+ olması olasılığını göstermektedir.

a) Rastgele bir kişi seçildiğinde A grubunda olması olayı A ve Rh+ olması ve A ve Rh- olması gibi iki ayrık olayın birleşimidir. Tablodan görüldüğü gibi rastgele bir kişi seçildiğinde A grubunda olması olasılığı 0.35'tir. Aşağıdaki tablo Tablo 4.7' de satırlar toplanarak elde edilir.

A	B	AB	0	Toplam
0.35	0.14	0.06	0.45	1

Tablo 4.8
Kan grupları ve olasılıkları

Bu tablodan rastgele bir kişi seçildiğinde A, B, AB ve 0 gruplarında olması olasılıkları sırasıyla 0.35, 0.14, 0.06 ve 0.45 olduğu kolayca görülür.

b) Rastgele bir kişi seçildiğinde AB grubunda ve Rh- olması, Tablo 4.7'den görüldüğü üzere 0.01'dir.

c) Rastgele bir kişi seçildiğinde, seçilenin 0 grubu ve Rh + olması olasılığı Tablo 4.7'deki ilgili hücrede 0.40 olduğu görülür.

d) Rastgele bir kişi seçildiğinde, seçilenin 0 grubu olduğu bilindiğine göre Rh+ olması

$$P(\text{Rh}+\ \text{olması}\,|\,0\ \text{grubu olması}) = \frac{P(\text{Rh}+\ \text{ve } 0\ \text{olması})}{P(0\ \text{grubu olması})} = \frac{0.40}{0.45} = 0.889$$

olarak bulunur.

e) Aşağıdaki Tablo 4.9, Tablo 4.7 de sütunlar toplanarak elde edilir.

Tablo 4.9
Kan grupları ve olasılıkları

Rh+	Rh-	Toplam
0.85	0.15	1

Bu tablo yardımıyla rastgele bir kişi seçildiğinde Rh+ ve Rh- olması olasılıkları sırasıyla 0.85 ve 0.15 olduğu kolayca görülür.

f) Rastgele bir kişi seçildiğinde, seçilenin Rh- olduğu bilindiğine göre B grubu olması olasılığı

$$P(\text{B grubu olması}\,|\,\text{Rh - olması}) = \frac{P(\text{B ve Rh - olması})}{P(\text{Rh - olması})} = \frac{0.04}{0.15} = 0.267$$

olarak bulunur

SIRA SİZDE 2

1. Bir Kozmetik firması, bayanlar için piyasaya sürdüğü özel bir yüz kremi ürününün yanı sıra bu ürünü tamamlayıcı nitelikte olan göz kremi ürününü de piyasaya sürmüştür. Ürünlerin satıldığı bir satış mağazasında 100 müşteri üzerinde yapılan bir değerlendirmede aşağıdaki sonuçlar elde edilmiştir.

Tablo 4.10
Yüz kremi ve göz kremi ürününe ilişkin tablo

		Yüz kremi		
		Alan	**Almayan**	**Toplam**
Göz kremi	**Alan**	32	33	65
	Almayan	8	27	35
	Toplam	40	60	100

a) Rastgele seçilen bir müşterinin göz kremi aldığı bilindiğine göre yüz kremi alması olasılığını bulunuz.
b) Rastgele seçilen bir müşterinin yüz kremi aldığı bilindiğine göre göz kremi alması olasılığını bulunuz.
c) Yüz kremi alma olayı ile göz kremi alma olayları bağımsız mıdır?
2. Bir televizyon satıcısı LCD ve Plazma olmak üzere iki tip televizyon satmaktadır. Satılan televizyonların %70'i LCD ve %30'u Plazmadır. Ayrıca, LCD satın alanların %20'si, Plazma satın alanların %50'si uluslararası garanti de satın almıştır. Eğer rastgele seçilen bir kişinin uluslararası garanti satın aldığı bilgisine ulaşıldı ise bu kişinin LCD televizyon satın alan bir kişi olması olasılığı nedir? (Bir kişinin iki tip televizyonu birlikte almadığı varsayımı altında)
3. Büyük bir şirket, çalışanlarına iki farklı sağlık sigortası planı ve iki farklı hayat sigortası planı sunmaktadır. Aşağıdaki tablo, çeşitli planları seçen çalışanların oranlarının bilgisini içermektedir.

Tablo 4.11
Çalışanların tercih ettikleri sigorta oranları

	Hayat Sigortası	
Sağlık Sigortası	1	2
1	%20	%15
2	%25	%40

Rastgele seçilen bir çalışanın hayat sigortasında 2. planı seçtiği bilindiğine göre, sağlık sigortasında 1. planı seçmesi olasılığı nedir?

OLAYLARIN BİRLEŞİMİNİN OLASILIĞI

İki olayın birleşiminin olasılığı hesaplanırken kullanılan kurala **toplama kuralı** denir. Toplama kuralı şu şekilde ifade edilebilir: A ile B olaylarının birleşiminin olasılığı, A'nın olasılığı ve B'nin olasılığı toplamından A ve B'nin kesişiminin olasılığı çıkartılarak elde edilir. Toplama kuralı

Toplama kuralı $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ şeklinde ifade edilir.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

şeklinde ifade edilir.

$A \cup B$ olayının olasılığı başka bir ifade ile A veya B olaylarından en az birinin ortaya çıkması olasılığıdır.

A ile B olaylarının birleşimi, A'da, B'de ve A kesişim B'de olan örnek noktaları içerir.

DİKKAT

Örnek 20: *Bir deneyin örnek uzayı aşağıdaki tabloda verilen dört örnek noktadan oluşmaktadır.*

Örnek noktalar	Olasılıkları
1	0.15
2	0.25
3	0.20
4	0.40

Tablo 4.12 *Bir deneyin örnek noktaları ve olasılıkları*

$A=\{1,2\}$ ve $B=\{1,4\}$ şeklinde tanımlanmış ise $P(A \cup B)$ olasılığını hesaplayınız.

Çözüm: Olayların olasılığı:
$P(A)=P(1)+P(2)=0.15+0.25=0.40$,
$P(B)=P(1)+P(4)=0.15+0.40=0.55$ ve
$A \cap B=\{1\}$ ve $P(A \cap B)=0.15$ tir.
$P(A \cup B)= P(A)+P(B) - P(A \cap B)=0.40+0.55 - 0.15=0.80$

olarak hesaplanır.

Örnek 21: *50 kişilik bir sınıfta öğrencilerin 25'i basketbol, 20'si voleybol, 10'u ise hem basketbol hem de voleybol oynamaktadır. Bu sınıftan rassal olarak seçilen bir öğrencinin basketbol veya voleybol oynuyor olması olasılığı kaçtır?*

Çözüm: İstenen olasılık değerini bulmak için rassal olarak seçilen bir öğrencinin

B: Basketbol oynaması
V: Voleybol oynaması
$B \cap V$: Hem basketbol hem de voleybol oynaması

olayları tanımlansın. Bu durumda, bu olaylara ilişkin olasılıklar

$$P(B)= 25/50=0.50,\ P(V) = 20/50=0.40 \text{ ve } P(B \cap V)=10/50= 0.20$$

şeklinde bulunur.

Toplama kuralı yardımıyla, sınıftan rassal olarak seçilen bir öğrencinin basketbol veya voleybol oynuyor olması olasılığı

$$P(B \cup V)=0.50+0.40 - 0.20=0.70$$

olarak elde edilir.

Örnek 22: *Bir bankaya, müşterilerin %40'ı gişe işlemleri,%30'u kredi işlemleri ve %10'u ise her iki işlem için gelmektedir. Rastgele bir müşteri bankaya girdiğinde, müşterinin gişe işlemleri veya kredi işlemleri için gelmesi olasılığı nedir?*

Çözüm: Öncelikle aşağıdaki olaylar tanımlanmalıdır.

A: Müşterinin gişe işlemleri için gelmesi,
B: Müşterinin kredi işlemleri için gelmesi.

Bu durumda,

$A \cap B$: *Müşterinin gişe ve kredi işlemleri için gelmesi*

olur.

Bu durumda, $P(A)=0.40$ ve $P(B)=0.30$ ve $P(A \cap B)=0.10$'dur. Böylece müşterinin gişe işlemleri veya kredi işlemleri için gelmesi olasılığı yani $P(A \cup B)$ toplama kuralı yardımıyla aşağıdaki gibi bulunur:

$$P(A \cup B)=P(A)+P(B) - P(A \cap B)=0.40+0.30 - 0.10=0.60.$$

Böylece, rastgele bankaya gelen bir müşterinin gişe işlemleri veya kredi işlemleri için gelmesi olasılığı %60'tır.

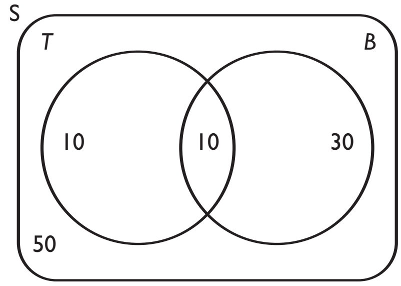
Örnek 23: *Bir kumaş fabrikasında toplam personel sayısı 100'dür. Bu personelin 20'si tasarım bölümünde, 40'ı baskı bölümünde, 10'u ise hem tasarım hem de baskı bölümünde çalışabilecek durumda personeldir. Bu fabrikadan rastgele seçilen bir personelin*

a) Yalnızca tasarım bölümünde çalışabilecek,
b) Tasarım veya baskı bölümünde çalışabilecek,
c) Tasarım veya baskı bölümünde ***çalışamayacak***
personel olması olasılığı nedir?

Çözüm: 100 personelin 20'si tasarım bölümünde çalışabilecek personel, bu personel arasında 10'u aynı zamanda baskı bölümünde çalışabilecek personeldir. Benzer olarak 40'ı baskı bölümünde çalışabilecek personel, bunlar arasında 10 kişi aynı zamanda tasarım bölümünde çalışabilecek personel durumundadır. Buna göre, bu fabrikadaki personel sayıları Şekil 4.2'de gibi gösterilebilir. Dikkat edilirse tasarım veya baskı bölümünde çalışabilecek personel sayısı 10+10+30=50'dir. Buna göre 100 – 50=50 kişi tasarım veya baskı bölümünde çalışamayacak personeldir. 50 sayısı şekil 4.2'de $T \cup B$ kümesinin dışında yani tümleyeninde gösterilmektedir.

Şekil 4.2

Personel Dağılımı



İlk olarak aşağıdaki olaylar tanımlansın:

T: Rastgele seçilen personelin tasarım bölümünde çalışabilecek personel olması,
B: Rastgele seçilen personelin baskı bölümünde çalışabilecek personel olması, tanımlanacak olursa;
T∩B: Rastgele seçilen personelin tasarım ve baskı bölümünde çalışabilecek personel olması,
T∪B: Rastgele seçilen personelin tasarım veya baskı bölümünde çalışabilecek personel olması.

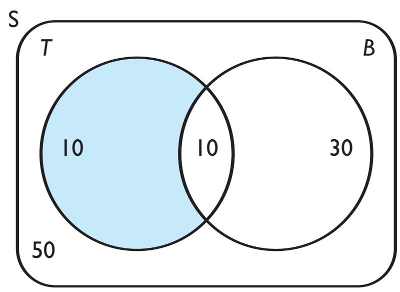
Bu olaylara ilişkin olasılıklar,

$$P(T)=\frac{20}{100}=0.20,\ P(B)=\frac{40}{100}=0.40,\ P(T\cap B)=\frac{10}{100}=0.10$$

şeklindedir.

a) Yalnızca tasarım bölümünde çalışabilecek personel sayısı Şekil 4.3'te taralı alanda, 10 olarak görülmektedir. Dolayısıyla, rastgele seçilen bir personelin yalnız tasarım bölümünde çalışabilecek personel olması olasılığı 10/100=0.10'dur.

Şekil 4.3



Yanlızca tasarım bölümünde çalışabilecek personel sayısı: taralı alan

b) Rastgele seçilen bir kişinin, tasarım veya baskı bölümünde çalışabilecek olması olasılığı, toplama kuralı yardımıyla

$$P(T\cup B)=P(T)+P(B)-P(T\cap B)=0.20+0.40-0.10=0.50$$

şeklinde elde edilir.

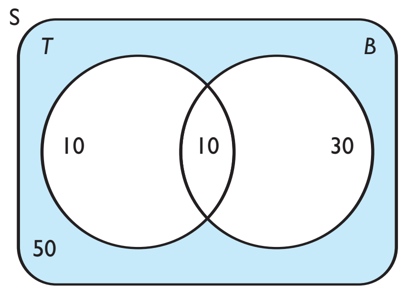
Aynı sonuç, toplama kuralı kullanılmadan sadece venn şemasında görülen bilgiler kullanılarak da bulunabilir. $T\cup B$ olayını içinde 10+10+30=50 sayıda personel vardır. Böylece,

$$P(T\cup B)=\frac{50}{100}=0.50 \text{ dir.}$$

c) Tasarım veya baskı bölümünde çalışamayacak personel sayısı $T\cup B$ olayının tümleyeninde görülmektedir. Şekil 4.4'te istenen olayın sonuç sayısının 50 olduğu görülür. Bu fabrikadan rastgele seçilen bir kişinin tasarım veya baskı bölümünde çalışamayacak personel olması olasılığı 50/100 yani %50'dir.

Şekil 4.4

Tasarım veya baskı bölümünde çalışamayacak personel sayısı: taralı alan



Örnek 24: *Bir firmada çalışan 200 kişiye, firmanın gelecek yıl için aldığı bir kararı onaylayıp onaylamadığı sorulmuş ve aşağıdaki tabloda verilen sonuçlar elde edilmiştir.*

Tablo 4.13 *Çalışanların dağılımı*

	Onaylıyor	Onaylamıyor	Toplam
Erkek	50	20	70
Kadın	60	70	130
Toplam	110	90	200

Çalışanlar arasından rastgele birisi seçildiğinde
a) Kadın ve onaylamıyor olması,
b) Erkek veya onaylıyor olması,
c) Onaylıyor veya onaylamıyor olması olasılıklarını hesaplayınız.
d) Onaylıyor ve onaylamıyor olması olayları bağımsız mıdır?

Çözüm: İstenen olasılıkları hesaplamak için gereken olayları tanımlamak ve bu olaylara ilişkin olasılıkları hesaplamak gerekir. Olaylar

E: Rastgele seçilen kişinin erkek olması,
K: Rastgele seçilen kişinin kadın olması,
O: Rastgele seçilen kişinin onaylıyor olması,
$\bar{O}$: Rastgele seçilen kişinin onaylamıyor olması,

şeklinde tanımlanırsa,

$$P(E) = \frac{70}{200} = 0.35;\ P(K) = \frac{130}{200} = 0.65;\ P(O) = \frac{110}{200} = 0.55;$$

$$P(\bar{O}) = \frac{90}{200} = 0.45$$

olarak yazılır.

a) Tabloda kadın ve onaylamıyor kişi sayısının 70 olduğu görüldüğüne göre

$$P(K \cap \bar{O}) = \frac{70}{200} = 0.35$$

olarak bulunur.

b) Erkek veya onaylıyor olması olasılığı toplama kuralı yardımıyla hesaplanabilir. Toplama kuralı

$$P(E \cup O) = P(E) + P(O) - P(E \cap O)$$

dir. $P(E)$ ve $P(O)$ olasılıkları, yukarıda bulunduğundan $P(E \cap O)$ olasılığının bulunması gerekmektedir. Tabloda verilen değerlere göre, bu olasılık,

$$P(E \cap O) = \frac{50}{200} = 0.25$$

olarak bulunur. Böylece, istenen olasılık

$$P(E \cup O) = P(E) + P(O) - P(E \cap O) = 0.35 + 0.55 - 0.25 = 0.65$$

şeklinde hesaplanır.

c) Rastgele seçilen birinin onaylıyor veya onaylamıyor olması, yine toplama kuralı yardımıyla

$$P(O \cup \bar{O}) = P(O) + P(\bar{O}) - P(O \cap \bar{O}) = 0.55 + 0.45 - 0 = 1$$

olarak bulunur. Burada, dikkat edilmesi gereken $O \cap \bar{O} = \varnothing$ olmasıdır. Bir başka deyişle, O ve $\bar{O}$ olayları ayrık olaylardır. Dolayısıyla,

$P(O \cap \bar{O}) = P(\varnothing) = 0$ dır.

d) O olayı ve $\bar{O}$ olayının bağımsız olup olmadıkları

$$P(O \cap \bar{O}) = P(O)\, P(\bar{O})$$

eşitliğinin sağlanıp sağlanmadığı kontrol edilerek belirlenmelidir.

$$P(O) = \frac{110}{200} = 0.55,\ P(\bar{O}) = \frac{90}{200} = 0.45,\quad P(O \cap \bar{O}) = 0$$

olduğundan bu eşitliğin sağlanmadığı açıktır. Dolayısıyla, O ve $\bar{O}$ olayları bağımlıdır. Buradan önemli bir sonuç ortaya çıkar, olasılıkları 0 olmayan, iki ayrık olay her zaman *bağımlıdır.*

Olasılıkları sıfır olmayan iki ayrık olay her zaman bağımlıdır.

Örnek 25: *Bir öğrencinin Beden Eğitimi (B) ve Kimya (K) derslerinden başarılı olması olasılıkları sırası ile P(B)=0.70 ve P(K)=0.40 dır. Bu öğrencinin beden eğitimi dersinden başarılı olması olayı Kimya dersinden başarılı olması olayından bağımsız ise sözkonusu öğrencinin Beden Eğitimi veya Kimya derslerinden başarılı olması olasılığı nedir?*

Çözüm: Öğrencinin, Beden Eğitimi veya Kimya derslerinden başarılı olması olasılığı toplama kuralı yardımıyla bulunabilir. Toplama kuralı,

$$P(B \cup K) = P(B) + P(K) - P(B \cap K)$$

dir. $P(B)=0.70$ ve $P(K)=0.40$ olarak soruda verildiğinden $P(B \cap K)$'nin bulunması gerekmektedir. $P(B \cap K)$ olasılığı, öğrencinin her iki dersten başarılı olması olasılığıdır. Öğrencinin Beden Eğitimi dersinden başarılı olması olayı, Kimya dersinden başarılı olması olayından bağımsız olduğundan

$$P(B \cap K) = P(B)\, P(K) = (0.70).(0.40) = 0.28$$

olarak hesaplanır. Böylece,

A ve B olayları bağımsız ise $P(A \cap B) = P(A)\, P(B)$ dir.

$$P(B \cup K) = P(B) + P(K) - P(B \cap K) = 0.70 + 0.40 - 0.28 = 0.82$$

Söz konusu öğrencinin Beden Eğitimi veya Kimya dersinden başarılı olması olasılığı 0.82'dir. Diğer bir ifade ile söz konusu öğrencinin Beden Eğitimi veya Kimya dersinden başarılı olması olasılığı %80 dir.

Ayrık Olaylar için Toplama Kuralı

A ve B iki ayrık olaylar olduğunda $A \cap B = \varnothing$ ve $P(\varnothing)=0$ olduğundan dolayı, ayrık olaylar için toplama kuralı

$$P(A \cup B) = P(A) + P(B)$$

şeklindedir.

Örnek 26: *Bir işletmede yaşlara göre yarı zamanlı (part-time) ve tam zamanlı çalışan (full-time) personel sayıları aşağıdaki tabloda verilmiştir.*

Tablo 4.14 *Bir işletmede çalışanların sınıflandırılması*

Yaş Grupları	**Yarı zamanlı**	**Tam zamanlı**	**Toplam**
16-23	10	11	21
24-31	5	35	40
32-39	4	45	49
40-üstü	1	9	10
Toplam	20	100	120

Çalışanlar arasından rastgele birisi seçildiğinde
a) Yarı zamanlı veya tam zamanlı çalışan olması,
b)16-23 veya 32-39 yaş grubunda olan çalışan olması
olasılığı nedir?

Çözüm:
a) Öncelikle

A: Rastgele seçilen kişinin yarı zamanlı çalışan personel olması
B: Rastgele seçilen kişinin tam zamanlı çalışan personel olması

şeklinde tanımlansın. Bu durumda,

$$P(A) = \frac{20}{120} \text{ ve } P(B) = \frac{100}{120}$$

dir. A ve B olaylarının ortak elemanının olmadığı açıktır. Çünkü, bir çalışan hem tam zamanlı hem de yarı zamanlı çalışamaz. Bu durumda, bu olaylar ayrık olaylardır. Rastgele seçilen birinin yarı zamanlı veya tam zamanlı olması olasılığı, yine toplama kuralı yardımıyla

$$P(A \cup B) = P(A) + P(B) = \frac{20}{120} + \frac{100}{120} = \frac{120}{120} = 1$$

olarak elde edilir.
b) Olaylar,

C: Rastgele seçilen kişinin 16-23 yaş grubunda olan personel olması,
D: Rastgele seçilen kişinin 32-39 yaş grubunda olan personel olması

şeklinde tanımlansın. Bu durumda, Tablo 4.14'ten

$$P(C)=\frac{21}{120} \text{ ve } P(D)=\frac{49}{120}$$

dir. C ve D şeklinde tanımlanan olaylar da ayrık olaylardır. Çünkü seçilen kişi ya 16-23 yaş grubunda ya da 32-39 yaş grubunda olacaktır. Dolayısıyla her iki grupta olamaz. Bu durumda, rastgele seçilen kişinin 16-23 veya 32-39 yaş grubunda olması olasılığı toplama kuralı yardımıyla

$$P(C\cup D)=P(C)+P(D)=\frac{21}{120}+\frac{49}{120}=\frac{70}{120}=0.583$$

olarak elde edilir.

Örnek 27: *Hilesiz bir zarın iki kez atılması deneyinde, iki atışın sonunda üste gelen sayıların toplamının 10 veya 12 elde edilmesi olasılığını bulunuz.*

Çözüm: Bir zarın iki kez atılması deneyinin örnek uzayı Tablo 4.15'te verilmiştir. Bir zarın iki kez atılması deneyinde toplam 36 örnek nokta bulunmakta ve dolayısıyla deneyin toplam sonuç sayısı 36 dır. Bu örnek noktalar eşit olasılığa sahip ve her birinin olasılığı 1/36'dır.

		İkinci Atış					
		1	2	3	4	5	6
İlk Atış	1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
	2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,5)
	3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,3)
	4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	**(4,6)**
	5	(5,1)	(5,2)	(5,3)	(5,4)	**(5,5)**	(5,6)
	6	(6,1)	(6,2)	(6,3)	**(6,4)**	(6,5)	**(6,6)**

Tablo 4.15
Bir zarın iki kez atılması deneyi örnek uzayı

A:İki atışın sonunda üste gelen sayıların toplamının 10 olması
B:İki atışın sonunda üste gelen sayıların toplamının 12 olması

olayları

A={(4,6),(6,4),(5,5)}
B={(6,6)}

biçimindedir. A ve B ayrık olaylardır. $A\cap B=\emptyset$ ve $P(\emptyset)=0$'dır.

$P(A)=3/36$ ve $P(B)=1/36$'dır.

A veya B olayının olasılığı

$P(A\cup B)=P(A)+P(B)=3/36+1/36=4/36$

dır.

Eğer soru *iki atışın sonunda üste gelen sayıların toplamının 10 veya 12 veya 9 elde edilmesi olasılığını bulunuz* şeklinde olsa idi. Bu durum üç olay için genişletilecekti.

C:İki atışın sonunda üste gelen sayıların toplamının 9 olması
C={(3,6), (6,3), (4,5),(5,4)} şeklinde ve $P(C)$ = 4/36.

İki atışın sonunda üste gelen sayıların toplamının 10 veya 12 veya 9 elde edilmesi olasılığını, üç ayrık olayın birleşiminin olasılığı

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) = 3/36 + 1/36 + 4/36 = 8/36$$

dır.

SIRA SİZDE 3

1. 1'den 20'ye kadar (20 dahil) olan tam sayılar arasından rastgele seçilen bir sayının 3'e veya 5'e bölünmesi olasılığı nedir?
2. Hilesiz iki farklı renge sahip zarların birlikte atılması deneyinde *A* olayı "*iki zarın toplamının 5 olması*", *B* olayı ise "*iki zarın farkının mutlak değerinin 5 olması*" şeklinde tanımlansın. *A* veya *B* olayının olasılığını bulunuz.
3. Bir konferansta aynı saatte gündüz iki farklı A ve B oturumları yapılmaktadır. Katılımcıların %20'si A oturumuna , %25'i B oturumuna katılmaktadır. Bu konferansta, bir katılımcı aynı saatte iki oturuma katılamamaktadır. A ve B oturumlarına katılanların tamamı akşam olan C oturumuna katılmış ve böylece C'ye %75 katılım olmuştur.
a) Rastgele seçilen bir katılımcının A oturumuna veya B oturumuna katılması olasılığı nedir?
b) Rastgele seçilen bir katılımcının A oturumuna ve C oturumuna katılması olasılığı nedir?
c) Rastgele seçilen bir katılımcının A oturumuna katıldığı bilindiğine göre C oturumuna katılması olasılığı nedir?

Özet

AMAÇ 1

Koşullu olasılıkları hesaplamak.

Bir B olayının gerçekleştiği bilindiğine göre A olayının olasılığı, koşullu olasılık olarak ifade edilir ve $P(A|B)$ şeklinde gösterilir ve bu olasılık

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) \neq 0$$

şeklinde hesaplanır.

AMAÇ 2

Bileşik olasılıkları hesaplamak

A ve B gibi iki olayın arakesitinin (kesişiminin) olasılığı **bileşik olasılık** olarak ifade edilmektedir. Bileşik olasılık $P(A \cap B)$ biçiminde gösterilir ve

$P(A \cap B) = P(B)\,P(A|B)$ ve

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

şeklinde hesaplanır. Buna **çarpma kuralı** adı verilir.

AMAÇ 3

Bağımsız olaylara ilişkin olasılık hesabını uygulamak.

A ve B olayları bağımsız ise $P(A|B) = P(A)$ ve $P(B|A) = P(B)$ dir. Böylece, Çarpma kuralı, A ve B olayları bağımsız ise $P(A \cap B) = P(A)\ P(B)$ şeklindedir.

AMAÇ 4

Bağımsız ve ayrık olaylar arasındaki farkı ayırt etmek.

A ve B olayları ayrık ise $A \cap B = \emptyset$ ve $P(A \cap B) = 0$ dır.

A ve B olayları bağımsız ise $P(A \cap B) = P(A)\ P(B)$ dır.

A ve B olayları ayrık ise ($P(A) \neq 0$ ve $P(B) \neq 0$) her zaman bağımlıdır.

A ve B olayları bağımsız ($P(A) \neq 0$ ve $P(B) \neq 0$) ise hiçbir zaman ayrık değildir.

AMAÇ 5

Olayların birleşiminin olasılığını hesaplaymak.

A veya B olaylarından en az birinin ortaya çıkmasına ilişkin olasılık yani $A \cup B$ olayının olasılığı:

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

A ve B ayrık ise $A \cup B$'nin olasılığı

$P(A \cup B) = P(A) + P(B)$

şeklinde hesaplanır. Buna **toplama kuralı** adı verilir.

Kendimizi Sınayalım

1. 1'den 30'a kadar (30 dahil) olan tam sayılar arasından rastgele seçilen bir sayının 5 ile bölündüğü bilindiğine göre bu sayının 3 ile bölünme olasılığı nedir?

a. 1/30
b. 1/15
c. 1/5
d. 1/3
e. 1

2. $P(B)$= 0.40 ve $P(A \mid B)$= 0.60 değerleri veriliyor ise A ve B olaylarının bileşik olasılığı ($P(A \cap B)$) kaçtır?

a. 0.24
b. 0.40
c. 0.57
d. 0.70
e. 1

3. Bir üniversitede çalışan 1200 kişiye yemekhanede sunulan yemeklerden memnun olup olmadığı sorulmuş ve tablodaki sonuçlar elde edilmiştir.

	Memnun	Memmun değil	Toplam
İdari personel	400	300	700
Akademik personel	350	150	500
Toplam	750	450	1200

Bu üniversitede çalışan bir kişi rassal olarak seçildiğinde, bu kişinin yemekhanede sunulan yemeklerden memnun bir kişi olduğu bilindiğine göre akademik personel olması olasılığı nedir?

a. 0.291
b. 0.467
c. 0.476
d. 0.625
e. 1

4. Bir sepette 3 çürük ve 7 sağlam elma vardır. Bu sepetten seçilen yerine koyulmaksızın art arda iki tane elma rassal olarak seçilmiştir. Seçilen elmalardan birincisinin çürük, ikincisinin sağlam olması olasılığı nedir?

a. 0.047
b. 0.140
c. 0.155
d. 0.210
e. 0.233

5. Belirli bir topluluktaki kan grubu olasılıkları aşağıdaki gibidir:

A	B	AB	0
0.40	0.12	0.06	0.42

Bu topluluktan rassal olarak seçilen iki kişinin kan grupları birbirinden bağımsız ise bu iki kişinin her ikisinin de kan grubunun 0 olması olasılığı nedir?

a. 0.025
b. 0.168
c. 0.176
d. 0.400
e. 0.420

6. Belli bir bölgede halkın %70'i ev sahibi, %30'u kiracıdır. Ev sahibi olanların %60'ı özel sektörde, %40'ı kamuda çalışmakta, kiracıların %70'ı özel sektörde, %30'u kamuda çalışmaktadır. Bu bölgeden rassal olarak bir kişi seçildiğinde kişinin kamuda çalıştığı bilindiğine göre kiracı olması olasılığı nedir?

a. 0.243
b. 0.300
c. 0.332
d. 0.370
e. 0.50

7. Hilesiz bir zarın iki kez atılması deneyinde, iki atışın sonunda zarların toplamının 11 veya 12 olması olasılığını bulunuz.

a. 1/36
b. 2/36
c. 3/36
d. 11/36
e. 12/36

8. Bir şirkette çalışan toplam 70 kişinin eğitim durumuna ilişkin verilerin oluşturduğu tablo aşağıda verilmiştir.

Lise (L)	Üniversite (Ü)	Yüksek lisans (Y)	Toplam
35	28	7	70

Bu şirketten rastgele seçilen bir kişinin üniversite mezunu veya yüksek lisans mezunu olması olasılığı nedir?

a. 7/70
b. 28/70
c. 35/70
d. 42/70
e. 63/70

9. Belirli bir üniversiteden rastgele bir öğrenci seçilsin. A, seçilen öğrencinin A türünde kredi kartına sahip olması olayı ve B, seçilen öğrencinin B türünde kredi kartına sahip olması olayını göstersin. $P(A)= 0.30$, $P(B)= 0.55$ ve $P(A \cap B)= 0.10$ olmak üzere rastgele seçilen bir öğrencinin kartlardan en az birine sahip olması olasılığı nedir? ($A \cup B$ olayının olasılığı nedir?)

a. 0.10
b. 0.25
c. 0.45
d. 0.75
e. 0.85

10. Bir öğrencinin müzik (M) ve fizik (F) derslerinden başarılı olması olasılıkları sırası ile $P(M) = 0.60$ ve $P(F)= 0.40$ dır. Bu öğrencinin müzik dersinden başarılı olması olayı fizik dersinden başarılı olması olayından bağımsız ise müzik veya fizik derslerinden başarılı olması olasılığı nedir?

a. 0.20
b. 0.24
c. 0.40
d. 0.60
e. 0.76

Yaşamın İçinden

“Bir asitli içecek firması, kendisine rakip bir firmanın satışlarını arttırmak için A, B ve C olarak nitelendirilen üç farklı şişe tasarımını piyasaya sürdüğünü öğrenmiştir. Ancak rakip firma, şişe tasarımı yanı sıra üç farklı pazarlama stratejisi de uygulamaya başlamıştır. Pazarlama strateji olarak; reklam, promosyon ve fiyatta indirimi uygulamaya başlamıştır. Firma, rakip firmanın şişe tasarımı ve pazarlama stratejilerine ilişkin çeşitli bilgiler öğrenmek istemektedir. Rakip firmanın, A tasarımını kullanması olasılığı nedir? A tasarımını kullanması ve pazarlama strateji olarak reklam uygulaması olasılığı nedir? B tasarımını kullanması ve pazarlama strateji olarak promosyonu kullanmaları olayları bağımsız mıdır? Yukarıda ifade edilen sorulara, olasılık ve koşullu olasılık ile cevaplar aranır.”

Kendimizi Sınayalım Yanıt Anahtarı

1. d	Yanıtınız yanlış ise “Koşullu Olasılık ve Çarpma Kuralı” konusunu yeniden gözden geçiriniz.
2. a	Yanıtınız yanlış ise “Koşullu Olasılık ve Çarpma Kuralı” konusunu yeniden gözden geçiriniz.
3. b	Yanıtınız yanlış ise “Koşullu Olasılık ve Çarpma Kuralı” konusunu yeniden gözden geçiriniz.
4. e	Yanıtınız yanlış ise “Koşullu Olasılık ve Çarpma Kuralı” konusunu yeniden gözden geçiriniz.
5. c	Yanıtınız yanlış ise “Koşullu Olasılık ve Çarpma Kuralı” konusunu yeniden gözden geçiriniz.
6. a	Yanıtınız yanlış ise “Koşullu Olasılık ve Çarpma Kuralı” konusunu yeniden gözden geçiriniz.
7. c	Yanıtınız yanlış ise “Olayların Birleşiminin Olasılığı” konusunu yeniden gözden geçiriniz.
8. c	Yanıtınız yanlış ise “Olayların Birleşiminin Olasılığı” konusunu yeniden gözden geçiriniz.
9. d	Yanıtınız yanlış ise “Olayların Birleşiminin Olasılığı” konusunu yeniden gözden geçiriniz.
10. e	Yanıtınız yanlış ise “Koşullu Olasılık ve Çarpma Kuralı” ve “Olayların Birleşiminin Olasılığı” konularını yeniden gözden geçiriniz.

Sıra Sizde Yanıt Anahtarı

Sıra Sizde 1

1. Bu deneyin örnek uzayının eleman sayısı, 5 doktor çift yani 10 kişi olduğu için "10" farklı eleman arasından "2" elemanlı grup sayısıdır. Dolayısıyla kombinasyon yardımıyla

$$\binom{10}{2} = \frac{10!}{(10-2)!2!} = \frac{10.9.8!}{8!.2!} = 45$$

olarak bulunur. 45 deneyin toplam sonuç sayısıdır. (n= 45)

45 sayıda 2 elemanlı grup arasında beş evli doktor çift, beş tane sonuç yani beş örnek noktadır. Buna göre *rastgele seçilen iki doktorun evli bir çift olması* olayında beş örnek nokta vardır. (n_A= 5)

Rastgele seçilen iki doktorun evli bir çift olması olasılığı 5/45= 0.11'dir.

2. Bu sınıfta 50-22= 28 erkek öğrenci vardır.

Rassal olarak seçilen bir öğrencinin erkek olması olasılığı 28/50= 0.56'dır.

3. Bir kutudaki bilyelerin %40'ı siyah, %60'ı kırmızı ise rassal olarak çekilen 50 bilyenin %40'ının siyah, %60'ının kırmızı olması beklenir. Buna göre 50'nin %60'ı (50).(0.60) = 30'dur. Böylece 50 bilyeden 30 tanesinin kırmızı bilye olması beklenir.

Sıra Sizde 2

1. İlk olarak aşağıdaki olaylar tanımlansın:

A: *Rastgele seçilen bir müşterinin yüz kremi alması,*

B: *Rastgele seçilen bir müşterinin göz kremi alması,*

ve

$A \cap B$: *Rastgele seçilen bir müşterinin yüz ve göz kremi alması.*

Olaylara ilişkin olasılıklar:

$$P(A) = \frac{40}{100} = 0.40, \ P(B) = \frac{65}{100} = 0.65 \ \text{ve}$$

$$P(A \cap B) = \frac{32}{100} = 0.32 \qquad \text{dir.}$$

a. Rastgele seçilen bir müşterinin göz kremi aldığı bilindiğine göre yüz kremi alması olasılığı

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{0.32}{0.65} = 0.492 \text{ 'dır.}$$

b. Rastgele seçilen bir müşterinin yüz kremi aldığı bilindiğine göre göz kremi alma olasılığı

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{0.32}{0.40} = 0.80 \text{ 'dir.}$$

c. *Yüz kremi alma* ile *göz kremi alma* olayları bağımsız ise $P(A \cap B) = P(A)\, P(B)$ olmalıdır.

$$P(A) = \frac{40}{100} = 0.40, \ P(B) = \frac{65}{100} = 0.65 \ \text{ve}$$

$$P(A \cap B) = \frac{32}{100} = 0.32$$

olduğuna göre 0.32 ≠ (0.40) (0.65)'dir.

Yüz kremi alma olayı ile göz kremi alma olayı bağımsız olaylar değildir. Diğer bir değişle bağımlıdır.

2. Olaylar aşağıdaki gibi tanımlandığında

A: *LCD model televizyon satılması*

B: *Uluslararası garanti satın alınması*

A olayının olasılığı

$P(A)$ = 0.70

şeklinde soruda verilmişti.

%70 LCD model TV'nin %20'si uluslararası garanti de satın aldığına göre $0.70 \times 0.20 = 0.14$ ve satılan TV'lerin %14'ü LCD ve uluslararası garanti satın almıştır.

Böylece $P(A \cap B)$ = 0.14'tür.

%30 Plazmanın %50'si uluslararası garantide satın aldığına göre (0.30) (0.50) = 0.15 ve satılan TV'lerin %15'i Plazma ve uluslararası garanti almıştır.

Bu durumda 0.15 + 0.14 = 0.29'dur. Dolayısıyla toplam müşterilerin %29'u uluslararası garanti satın almıştır. $P(B)$ = 0.29'dur.

Eğer rastgele seçilen bir kişinin uluslararası garanti satın aldığı bilgisine ulaşıldı ise, o kişinin LCD televizyon satın alan bir kişi olması olasılığı

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{0.14}{0.29} = 0.48 \text{ 'dir.}$$

Soruda verilen değerler aşağıdaki tabloya dönüştürülerek de sorunun çözümüne kolayca ulaşılır:

Buna göre:

	LCD	Plazma	Toplam
Uluslararası garanti alan	14	15	29
Uluslararası garanti almayan	56	15	71
Toplam	70	30	100

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{14/100}{29/100} = 0.48$$

şeklinde bulunur.

3. Olaylar ve olaylara ilişkin olasılıklar aşağıdaki gibidir.
A: *Rastgele seçilen bir kişinin sağlık sigortasında 1. planı seçmesi*
B: *Rastgele seçilen bir kişinin hayat sigortasında 2. planı seçmesi*
şeklinde tanımlandığında
A∩B: *Rastgele seçilen bir kişinin sağlık sigortasında 1. planı ve hayat sigortasında 2. planı seçmesi*
Tablo 4.11. de satır ve sütun toplamları eklendiğinde aşağıdaki tablo elde edilir.

	Hayat Sigortası		
Sağlık Sigortası	**1**	**2**	**Toplam**
1	%20	%15	%35
2	%25	%40	%65
Toplam	%45	%55	%100

Dolayısıyla bu şirkettin %35'i sağlık sigortasında 1. planı seçmiş, %55'i hayat sigortasında 2. planı seçmiş, %15'i sağlık sigortasında 1. planı ve hayat sigortasında 2. planı seçmiştir. Buna göre
$P(A) = 0.35$, $P(B) = 0.55$ ve $P(A \cap B) = 0.15$
olduğuna göre, rastgele seçilen bir kişinin hayat sigortasında 2. planı seçtiği bilindiğine göre, sağlık sigortasında 1. planı seçmesi olasılığı

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{0.15}{0.55} = 0.27$$ yani %27'dir.

Sıra Sizde 3

1. Olaylar
A: *Rastgele seçilen bir sayının 3'e bölünmesi*
B: *Ratsgele seçilen bir sayının 5'e bölünmesi*
şeklinde tanımlandığında
A∩B: *Rastgele seçilen bir sayının 3 ve 5'e bölünmesi*
Böylece
$A = \{3, 6, 9, 12, 15, 18\}$ ve $B = \{5, 10, 15, 20\}$ ve $A \cap B = \{15\}$
Olaylara ilişkin olasılıklar:

$$P(A) = \frac{6}{20}, \; P(B) = \frac{4}{20} \text{ ve } P(A \cap B) = \frac{1}{20}$$

Toplama kuralı yardımıyla rastgele seçilen bir sayının 3'e veya 5'e bölünmesi olasılığı

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) =$$

$$\frac{6}{20} + \frac{4}{20} - \frac{1}{20} = \frac{9}{20} = 0.45$$

olarak elde edilir.

2. Bu deneyin örnek uzayı $S = \{(1,1),(1,2),\ldots,(5,6),(6,6)\}$ şeklindedir. Bu örnek uzay Tablo 4.15'te verilmiştir. S'nin 36 örnek noktası vardır.
A olayı iki zarın toplamının 5 olması olayı:
$A = \{(1,4),(2,3),(3,2),(4,1)\}$
B olayı iki zarın farkının mutlak değerinin 5 olması olayı:
$B = \{(1,6),(6,1)\}$
Bu durumda
$A \cap B = \varnothing$'dir.
Olaylara ilişkin olasılıklar

$$P(A) = \frac{4}{36}, \; P(B) = \frac{2}{36} \text{ ve } P(A \cap B) = 0$$

şeklindedir. *A* ve *B* olayları ayrık olaylardır. Ayrık olaylar için toplama kuralı yardımıyla, *A* veya *B* olayının olasılığı

$$P(A \cup B) = P(A) + P(B) = \frac{4}{36} + \frac{2}{36} = \frac{6}{36} = 0.16$$

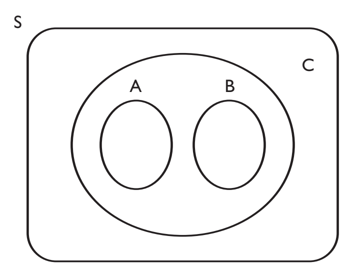
olarak elde edilir.

3. Olaylar
A: *Katılımcının A oturumuna katılması olayı*
B: *Katılımcının B oturumuna katılması olayı*
C: *Katılımcının C oturumuna katılması olayı*
şeklinde ifade edilirse

$P(A) = 0.20$, $P(B) = 0.25$ ve $P(C) = 0.75$'tir.

Soruda dikkat edilecek hususlar:

- A ve B oturumu aynı saatte olduğu için bir katılımcı iki oturuma katılamaz. Bu durumda, A ya katılanlar ve B'ye katılanlar farklı kişilerdir. Dolayısıyla, yukarıda tanımlanan *A* ve *B* olaylarının ortak elemanı (katılımcısı) yoktur ve arakesitini boş kümedir. Yani *A* ve *B* ayrık olaylardır. $A \cap B = \varnothing$
- A ve B oturumlarına katılanların tamamı akşam olan C oturumuna katılmıştır. Yani C oturuma sabahki A ve B oturumlarına katılanların tamamı katılmıştır. Yani *A* ve *B* olaylarının elemanları (katılımcısı) *C* olayının da elemanıdır. Bu durumda *A*, *B* ve *C* olayları aşağıdaki şekildeki gibidir.



a. $P(A \cup B) = P(A) + P(B) = 0.20 + 0.25 = 0.45$
b. $P(A \cap C) = P(A) = 0.20$ (Çünkü $A \cap C = A$ dır.)
c. $$P(C \mid A) = \frac{P(A \cap C)}{P(A)} = \frac{P(A)}{P(A)} = 1$$

Yararlanılan ve Başvurulabilecek Kaynaklar

Akdeniz F. (2007). **Olasılık ve İstatistik,** Adana: Nobel Kitapevi.

Anderson D.R., Sweeney D.J., Williams T.A. (2005). **Statistics for Business and Economics,** China: Thomson-South-Western.

Ben M., Levy H. (1983). **Business Statistics Fundamentals and Applications,** New York, USA: Random House Inc.

McClave J.T., Benson P.G., Sincich T. (2001). **Statistics for Business and Economics,** New Jersey, USA:Prentice-Hall Inc.

Newbold P. (1995). **Statistics for Business and Economics,** New Jersey, USA:Prentice-Hall Inc.

Yüzer A., Ağaoğlu E., Tatlıdil H., Özmen A., Şıklar E. (2006). **İstatistik,** Açıköğretim Fakültesi Yayını.

5

Amaçlarımız

Bu üniteyi tamamladıktan sonra;

- Rassal değişken kavramını tanımlayabilecek,
- Rassal değişkenlerin çeşitlerini (kesikli ve sürekli) ve aralarındaki farkı ayırt edebilecek,
- Olasılık dağılımı kavramını tanımlayabilecek,
- Olasılık ve istatistik teorisinde yaygın olarak kullanılan bazı kesikli dağılımları (Bernoulli, Binom ve Poisson) tanımlayabilecek ve bu dağılımlar yardımıyla olasılık değerlerini hesaplayabilecek,
- Binom dağılımının, Poisson dağılımına yakınsama özelliğini kullanabilecek,
- Kesikli rassal değişkenlerin ortalama, varyans ve standart sapmalarını hesaplayabileceksiniz.

Anahtar Kavramlar

- Rassal Değişken
- Kesikli Olasılık Dağılımı
- Bernoulli, Binom ve Poisson Dağılımları
- Ortalama, Varyans ve Standart Sapma

İçindekiler

İstatistik-I

Kesikli Rassal Değişkenler ve Bazı Kesikli Dağılımlar

Kesikli Rassal Değişkenler ve Bazı Kesikli Dağılımlar

GİRİŞ

İstatistik, genel olarak, rassal bir olayı (ya da deneyi) matematiksel olarak modellemek ve bu model yardımıyla, anakütlenin bilinmeyen karakteristik özellikleri (ortalama, varyans v.b. gibi) hakkında çıkarım yapmak amacıyla kullanılan bir bilim dalı olarak tanımlanır.

Rassal bir olayın modellenmesi, sayısal değerlerle ifade edilen ve rassal değişken olarak adlandırılan değişkenler yardımıyla yapılır. Bu ünitede, rassal değişken kavramı verildikten sonra sürekli ve kesikli olarak adlandırılan rassal değişken çeşitlerinden bahsedilecektir. Gerçek hayat problemlerinde, sıklıkla kullanılan kesikli olasılık dağılımlarının tanımı yapılacak ve özellikleri verilecektir. Kesikli rassal değişkenleri modellemek için kullanılan ve literatürde önemli bir yer tutan *Bernoulli, Binom* ve *Poisson* gibi bazı olasılık dağılımları ayrıntılı olarak incelenecektir. *Binom* dağılımının belli durumlarda *Poisson* dağılımına yakınsamasından bahsedilecektir. Kesikli bir rassal değişken için, *ortalama, varyans ve standart sapma* kavramları tanımlanacaktır. Konu anlatımlarını takiben çeşitli örnekler verilecek, konuların daha iyi anlaşılması sağlanacaktır.

RASSAL DEĞİŞKEN KAVRAMI

Rassal bir olayın (ya da deneyin) sonuçlarını, sayısal değerlerle ifade eden değişkene, **rassal değişken** *(random variable)* denir.

Örneğin, bir para atma deneyi yazı (Y) veya tura (T) ile sonuçlanır. Rassal değişken,

X: T sayısı

olarak tanımlanırsa rassal değişkenin alacağı değerin 0 veya 1 olacağı açıktır. Dolayısıyla, deneyin sayısal olmayan sonuçları (Y ve T), X rassal değişkeni yardımıyla sayısal değerlerle (0 ve 1) ifade edilmiş olur.

Bir zar atma deneyi ise 1, 2, 3, 4, 5 veya 6 ile sonuçlanır. Para atma deneyinden farklı olarak, zar atma deneyinin sonuçları sayısaldır. Bu örnekte, rassal değişken

X: Zarın üzerindeki noktaların sayısı

olarak tanımlanırsa rassal değişkeninin alacağı değerlerin 1, 2, 3, 4, 5 veya 6 olacağı açıktır. Böylelikle, zar atma deneyinde, deneyin sayısal olan sonuçları (1, 2, 3, 4, 5, 6), X rassal değişkeni yardımıyla yine sayısal değerlerle (1, 2, 3, 4, 5, 6) ifade edilmiş olur.

Rassal değişken, tanım kümesi **örnek uzay** (*s*), değer kümesi ise **reel sayılar** (*R*) olan bir fonksiyon olarak tanımlanır.

Şu ana kadar rassal değişken kavramı, sözel olarak ele alınıp tanımlanmıştır. Rassal değişken kavramının matematiksel tanımı ise aşağıda verildiği gibidir.

Rassal değişken, tanım kümesi **örnek uzay** *(sample space-S)*, değer kümesi ise **reel sayılar (R)** olan bir fonksiyon olarak tanımlanır.

DİKKAT

İstatistik terminolojisinde, rassal değişkenler büyük harflerle (*X, Y, Z*,... v.b.), rassal değişkenlerin aldıkları değerler ise küçük harflerle (*x, y, z*,...v.b) gösterilir.

X rassal değişkeni, örnek uzaydaki her bir sonucu sadece ve sadece 1 tane x değerine götürür. Bununla beraber, X rassal değişkeninin, S örnek uzayındaki iki farklı sonucu aynı değere götürdüğü durumlar da olabilir.

Rassal değişken kavramını daha iyi anlayabilmek için aşağıda verilen örneği inceleyelim.

Örnek 1: *Zar atma deneyinde, iki zar aynı anda atılıyor. Bu deneyin olası tüm sonuçlarının bir kümesi olarak ifade edilen örnek uzay (S) aşağıda gösterildiği gibi elde ediliyor:*

$$
\begin{aligned}
S = \{ &(1,1), \quad (1,2), \quad (1,3), \quad (1,4), \quad (1,5), \quad (1,6),\\
&(2,1), \quad (2,2), \quad (2,3), \quad (2,4), \quad (2,5), \quad (2,6),\\
&\qquad\qquad\qquad\qquad \vdots\\
&(6,1), \quad (6,2), \quad (6,3), \quad (6,4), \quad (6,5), \quad (6,6)\}.
\end{aligned}
$$

Bu deneyde rassal değişken,

X: İki zarın üzerindeki noktaların toplamı

olarak tanımlansın. Örneğin, birinci zar 3, ikinci zar 4 gelmişse, bir başka deyişle, deneyin sonucunda s= (3,4) çifti elde edilmişse, X rassal değişkeninin alacağı değer x= 3+4=7 olur. Bu durum,

"X rassal değişkeni, örnek uzaydaki (3,4) elemanını, reel sayıların bir elemanı olan x= 7'ye götürmektedir"

şeklinde ifade edilir. Benzer şekilde, örnek uzayın diğer elemanları için, X rassal değişkeninin aldığı değerler aşağıda gösterilmiştir.

(1,1) $\rightarrow x$= 2
(1,2), (2,1) $\rightarrow x$= 3
(1,3), (3,1), (2,2) $\rightarrow x$= 4
(1,4), (4,1), (2,3), (3,2) $\rightarrow x$= 5
(1,5), (5,1), (2,4), (4,2), (3,3) $\rightarrow x$= 6
(1,6), (6,1), (2,5), (5,2), (3,4), (4,3) $\rightarrow x$= 7
(2,6), (6,2), (3,5), (5,3), (4,4) $\rightarrow x$= 8
(3,6), (6,3), (4,5), (5,4) $\rightarrow x$= 9
(4,6), (6,4), (5,5) $\rightarrow x$= 10
(5,6), (6,5) $\rightarrow x$= 11
(6,6) $\rightarrow x$= 12

X rassal değişkeninin aldığı değerlerin kümesi *"Değer Kümesi"* olarak adlandırılır.

Örneğin, yukarıdaki örnekte, X rassal değişkeninin aldığı değerlerin kümesi,

$$\{2,3,4,5,6,7,8,9,10,11,12\}$$

şeklinde ifade edilir.

RASSAL DEĞİŞKENLERİN ÇEŞİTLERİ

Rassal değişkenler, aldıkları değerlere göre **kesikli** *(discrete)* ya da **sürekli** *(continuous)* olarak adlandırılırlar. *Değer* kümesi **sayılabilir** *(countable)* olan rassal değişkenler kesikli, **sayılamayan** *(uncountable)* olan rassal değişkenler ise sürekli olarak isimlendirilir.

Değer kümesi **sayılabilir** (countable) olan rassal değişkenler **kesikli, sayılamayan** (uncountable) olan rassal değişkenler ise **sürekli** olarak isimlendirilir.

Kesikli rassal değişkenin (X) alabileceği değerlere örnekler:

- $x = 0,1$
- $x = 0,1,2,3,\ldots$
- $x = k, k+1, k+2,\ldots$

Sürekli rassal değişkenin (X) alabileceği değerlere örnekler:

- $0 < x < 1$
- $0 < x < \infty$
- $-\infty < x < \infty$.

Bu örnekler için X rassal değişkeni kesikli ise değer kümesi

$\{0,1\}$, $\{0,1,2,3,\ldots\}$, $\{k, k+1, k+2,\ldots\}$ (X rasal değişkeninin alabileceği değerler sayılabilir)
ve X rassal değişkeni sürekli ise

$(0,1)$, $(0,\infty)$, $(-\infty,\infty)$ (X rasal değişkeninin alabileceği değerler sayılamaz)
şeklinde ifade edilir. Burada, {0,1} elemanları 0 ve 1 olan kümeyi, (0,1) ise 0, 1 aralığını ifade etmektedir.

Örnek 2: *Örnek 1'de verilen zar atma deneyinde, X rassal değişkeninin değer kümesi* {2,3,4,5,6,7,8,9,10,11,12} *olarak bulunmuştur. Değer kümesi sayılabilir olduğundan, X rassal değişkeni kesiklidir.*

Örnek 3: *Kesikli ve sürekli rassal değişkenler arasındaki farkı daha iyi anlayabilmek için gerçek hayattan alınmış bazı örnekler aşağıda verilmiştir.*

Kesikli rassal değişkenler için örnekler:
- Bir kutudaki kusurlu kalem sayısı.
- Atış poligonunda hedefe yapılan isabetli atış sayısı.
- Bir kavşakta meydana gelen trafik kazalarının sayısı.
- Para atma deneyinde tura gelinceye kadar yapılan atış sayısı.
- Bir büfeye gün içinde gelen müşteri sayısı.

Sürekli rassal değişkenler için örnekler:
- Öğrencilerin boy uzunlukları.
- Yarışmacıların 100 metreyi koşma süreleri.
- Domateslerin ağırlığı.
- Rüzgârın hızı.
- Havanın sıcaklığı (diğerlerinden farklı olarak negatif değerler de alır).

Kitabın bu ünitesinde, kesikli rassal değişkenler üzerinde durulacaktır.

SIRA SİZDE 1

1. Aşağıda tanımlanan rassal değişkenlerin hangilerinin kesikli, hangilerinin sürekli olduğunu belirleyiniz.
a) Bir kütüphaneye gelen öğrenci sayısı.
b) Bir markette para ödemek için bekleyen müşterilerin hizmet alma süresi.
c) Bir apartmandaki ailelerin aylık doğal gaz tüketimi.
d) Antalya Gazipaşa'ya havayoluyla gelen yabancı turist sayısı.
2. Negatif değerler alabilen sürekli rassal değişkene örnek veriniz.

OLASILIK DAĞILIMI

Olasılık dağılımı *(probability distribution), P(X= x), X* kesikli rassal değişkeninin aldığı değerler ile bu değerlere karşılık gelen olasılıkları ifade eder.

Örnek 4: *Örnek 1'de verilen zar atma deneyi için olasılık dağılımını elde edelim.* Örneğin,

A olayı: *X rassal değişkeninin 6 değerini alması (X= 6),*

veya bir başka deyişle, iki zarın üzerindeki noktaların toplamının 6 değerini alması olarak tanımlansın. Bu durumda, *A* kümesinin elemanları aşağıda gösterildiği gibidir;

$$A = \{(1,5),(5,1),(2,4),(4,2),(3,3)\} \ ve \ n_A = 5.$$

Bu örnekte, *S* örnek uzayının eleman sayısının 36 olduğu açıktır, bkz. **Örnek 1.** Herhangi bir *A* olayının gerçekleşme olasılığı

$$P(A) = \frac{n_A}{n}$$

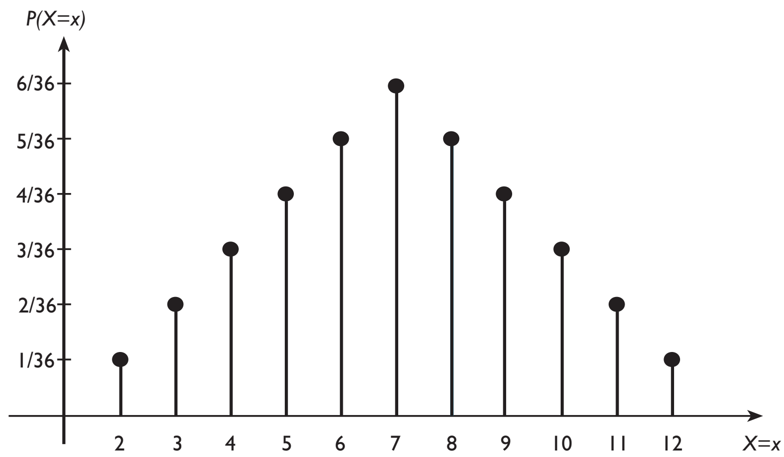
n_A : *A* olayının eleman sayısı
n : *S* örnek uzayının eleman sayısı
olarak tanımlandığından,

$$P(A) = P(X = 6) = \frac{n_A}{n} = \frac{5}{36}$$

olarak bulunur. Benzer şekilde, diğer olasılık değerleri de hesaplanarak aşağıda verilen olasılık dağılımı elde edilir.

X= x	2	3	4	5	6	7	8	9	10	11	12
P(X= x)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

Zar atma deneyine ilişkin olasılık dağılımının grafiği aşağıda gösterildiği gibi elde edilir.



Şekil 5.1

Zar atma deneyine ilişkin olasılık dağılımının grafiği.

Bir fonksiyonun, olasılık dağılımı olarak tanımlanabilmesi için aşağıda verilen özelliklerin sağlanması gerekir:

(i) X kesikli rassal değişkeninin, herhangi bir x'e eşit olma olasılığı, 0 ile 1 arasında değişir. Bir başka deyişle,

$0 \leq P(X = x) \leq 1$, her x değeri için

koşulu sağlanmalıdır.

(ii) X kesikli rassal değişkeninin, x'in tüm olası değerlerine eşit olma olasılıklarının toplamı 1'e eşittir. Bir başka deyişle,

$$\sum_x P\left(X = x\right) = 1$$

koşulu sağlanmalıdır.

Örnek 5: *Örnek 4'te verilen P(X= x) fonksiyonunu ele alalım:*

X= x	2	3	4	5	6	7	8	9	10	11	12
P(X= x)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

P(X= x) fonksiyonunun, yukarıda verilen (i) ve (ii) özelliklerini sağladığını aşağıda gösterelim.

(i) Yukarıdaki tablodan görülebileceği gibi, değer kümesinin {2,3,4,5,6,7,8,9,10,11,12} her bir elemanı için *P(X= x)* olasılığı 0 ile 1 arasında değer almaktadır. Örneğin, *P(X= 2)=1/36*, *P(X= 3)=2/36* ve *P(X= 4)* =3/36'dır. Benzer şekilde diğer olasılıkların da 0 ile 1 arasında olduğu kolayca görülebilir. Bir başka deyişle,

$$0 \leq P(X = x) \leq 1,\ x = 2,3,4,5,6,7,8,9,10,11,12$$

koşulu sağlanır.

(ii) Yukarıdaki tablodan görülebileceği gibi, X rassal değişkeninin alabileceği tüm değerlerin olasılıklarının toplamı 1'e eşittir. Bir başka deyişle,

$$\sum_x P(X=x) = \sum_{x=2}^{12} P(X=x)$$
$$= P(X=2) + P(X=3) + \ldots + P(X=12)$$
$$= \frac{1}{36} + \frac{2}{36} + \ldots + \frac{1}{36} = 1$$

koşulu sağlanır.

Sonuç olarak, *P(X= x)* fonksiyonu, (i) ve (ii) koşullarını sağladığından bir olasılık dağılımıdır.

Kesikli Birikimli Olasılık Dağılımı

Kesikli birikimli olasılık dağılımı *(discrete cumulative probability distribution)*

$$F_X(x) = P\left(X \le x\right) = \sum_{i \le x} P(X=i)$$

olarak tanımlanır. Bir başka deyişle, $F_X(x)$ fonksiyonu, *X* rassal değişkeninin belli bir *x*'ten daha küçük ya da eşit değer alması olasılığını ifade eder. Birikimli olasılık dağılımı, *x*'in bütün değerleri için 0 ile 1 arasında değerler alır. Bir başka deyişle,

$$0 \le F_X\left(x\right) \le 1$$

dir.

Örnek 6: *Bir para atma deneyi yapılıyor. X rassal değişkeni tura (T) sayısı olarak tanımlanıyor. Bu deney için olasılık dağılımı P(X= x)'i ve birikimli olasılık dağılımı* $F_X(x)$*'i bulunuz. Grafiklerini çiziniz. P(X= 1) olasılığını birikimli olasılık dağılımı yardımıyla hesaplayınız.*

Çözüm: Bir para atma deneyinde, örnek uzay

$$S = \left\{Y, T\right\}$$

olur. Deneyin sonucunda, yazı (Y) gelirse *X* rassal değişkeni 0, tura (T) gelirse 1 değerini alır. Bu nedenle, *X* rassal değişkeninin değer kümesi {0,1} olarak ifade edilir.

Bu örnekte ilgilenilen iki tane olay vardır. Birincisi, deneyin sonunda tura gelmemesi (A_1) olayı, ikincisi ise deneyin sonunda tura gelmesi (A_2) olayıdır. Dolayısıyla,

$$A_1 = \left\{Y\right\} \text{ ve } A_2 = \left\{T\right\}$$

olarak tanımlanırsa

$$P(A_1) = P(X=0) = \frac{n_{A_1}}{n} = \frac{1}{2} \text{ ve } P(A_2) = P(X=1) = \frac{n_{A_2}}{n} = \frac{1}{2}$$

bulunur. Bu sonuçlar kullanılarak, olasılık dağılımı

X= x	0	1
P(X= x)	1/2	1/2

şeklinde ifade edilir.

X rassal değişkeni 0 ve 1 değerlerini aldığından, $x<0$ ise

$$F_X\left(x\right) = P\left(X \le x\right) = 0$$

dır. $0 \le x < 1$ ise

$$F_X(x) = P(X \le x) = P(X = 0)$$
$$= \frac{1}{2}$$

x ≥ 1 ise ve $x \ge 1$ ise

$$F_X(x) = P(X \le x)$$
$$= P(X = 0) + P(X = 1)$$
$$= \frac{1}{2} + \frac{1}{2} = 1$$

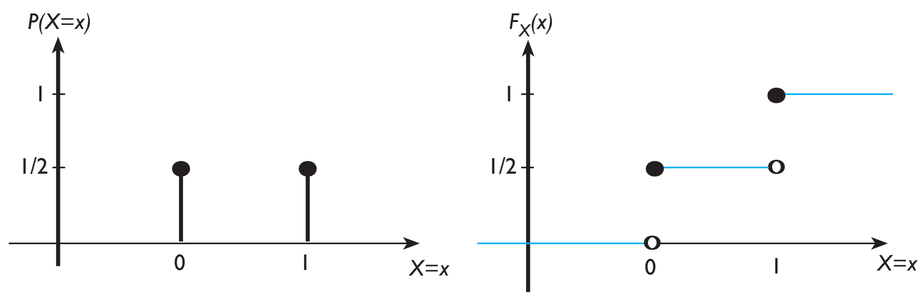
olarak bulunur. Dolayısıyla, birikimli olasılık dağılımı

$$F_X(x) = \begin{cases} 0, & x < 0 \\ 1/2, & 0 \le x < 1 \\ 1, & x \ge 1 \end{cases}$$

şeklinde elde edilir. *P(X= x)* ve $F_X(x)$'in grafikleri Şekil 5.2'de gösterildiği gibidir.

Şekil 5.2

Para atma deneyinde P(X= x) ve $F_X(x)$'in grafikleri



(a) P(X=x)'in grafiği (b) $F_X(x)$'in grafiği

P(X= 1) olasılığı, *X* rassal değişkeninin 1 ve 1'den daha küçük değer alma olasılığı ile *X* rassal değişkeninin 0 ve 0'dan daha küçük değer alma olasılığı arasındaki farka eşittir. Buradan,

$$P(X = 1) = F_X(1) - F_X(0)$$
$$= 1 - \frac{1}{2} = \frac{1}{2}$$

olarak bulunur.

Örnek 7: *Örnek 4'te verilen,*

X= x	2	3	4	5	6	7	8	9	10	11	12
P(X= x)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

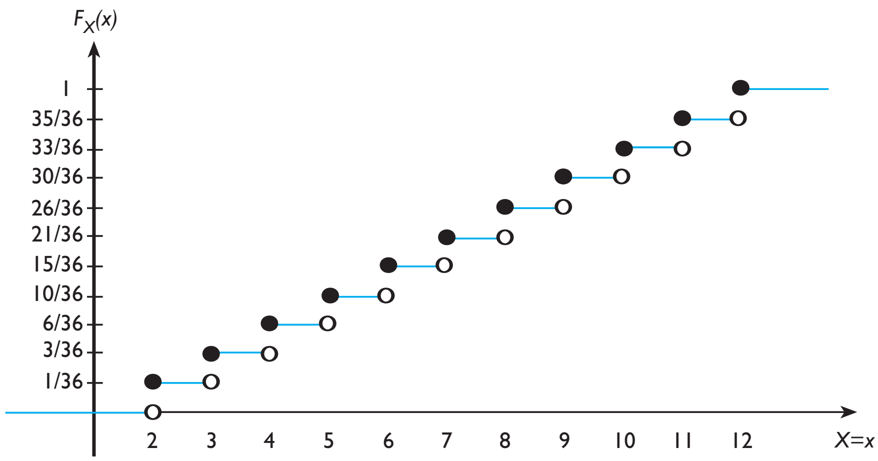
olasılık dağılımı için, X rassal değişkeninin birikimli olasılık dağılımı aşağıda gösterildiği gibi elde edilir.

$$F_X(x) = \begin{cases} 0, & x < 2 \\ 1/36, & 2 \le x < 3 \\ 3/36, & 3 \le x < 4 \\ 6/36, & 4 \le x < 5 \\ 10/36 & 5 \le x < 6 \\ 15/36, & 6 \le x < 7 \\ 21/36, & 7 \le x < 8 \\ 26/36, & 8 \le x < 9 \\ 30/36, & 9 \le x < 10 \\ 33/36, & 10 \le x < 11 \\ 35/36, & 11 \le x < 12 \\ 1, & x \ge 12 \end{cases}$$

$F_X(x)$'in grafiği, Şekil 5.3'te gösterildiği gibidir.

Şekil 5.3

Zar atma deneyinde $F_X(x)$'in grafiği



Bu örnekte, P(X= 6) olasılığını birikimli olasılık dağılımı yardımıyla hesaplayalım. P(X= 6) olasılığı, X rassal değişkeninin 6 ve 6'dan daha küçük değer alma olasılığı ile X rassal değişkeninin 5 ve 5'ten daha küçük değer alma olasılığı arasındaki farka eşittir. Buradan,

$$P(X=6) = F_X(6) - F_X(5)$$
$$= \frac{15}{36} - \frac{10}{36} = \frac{5}{36}$$

olarak bulunur.

Benzer olarak, $P(4 < X \leq 7)$ *olasılığı, birikimli olasılık dağılımı yardımıyla*

$$P(4 < X \leq 7) = F_X(7) - F_X(4) = \frac{21}{36} - \frac{6}{36} = \frac{15}{36}$$

şeklinde hesaplanabilir. Burada, P(4 < X ≤ 7) olasılığı, şekil 5.3'te verilen grafik yardımıyla da hesaplanabilir.

SIRA SİZDE 2

1. *X* rassal değişkeninin aldığı değerler ve bu değerlere karşılık gelen olasılıklar aşağıdaki tabloda verilmiştir.

X= x	*0*	*1*	*2*	*3*	*4*
P(X= x)	*3/20*	*2/10*	*3/10*	*2/10*	*3/20*

a) *P(X= x)* fonksiyonunun olasılık dağılımı olup olmadığını belirleyiniz.
b) *P(X > 0) = 1- P(X=0)* olduğunu tabloda verilen olasılık değerlerini kullanarak gösteriniz.
2. *X* rassal değişkeninin olasılık dağılımı,

X= x	*1*	*2*	*3*	*4*
P(X= x)	*0.12*	*0.16*	*0.54*	*c*

şeklinde veriliyor.
a) *P(X= x)* fonksiyonunun bir olasılık dağılımı olabilmesi için *c* sabiti ne olmalıdır?
b) Birikimli olasılık dağılımı $F_X(x)$'i bulunuz.

BAZI KESİKLİ DAĞILIMLAR

Literatürde rassal değişkenleri modellemek için kullanılan birçok olasılık dağılımı vardır. Olasılık dağılımları, kesikli dağılım olarak da adlandırılır. Bu bölümde, olasılık ve istatistik teorisinde yaygın olarak kullanılan bazı kesikli dağılımlar incelenecektir.

Bernoulli Dağılımı

İki sonucu olan bir deneyi (*Bernoulli denemesi*) modellemek için kullanılan kesikli bir dağılımdır. Genellikle, bu sonuçlar "**başarı** (*success*)" ve "**başarısızlık** (*failure*)" olarak isimlendirilir. *X* rassal değişkeni *"başarı"* durumunda 1, *"başarısızlık"* durumunda ise 0 değerini alır. Bernoulli denemesinin başarı ile sonuçlanma olasılığı "*p*", başarısızlıkla sonuçlanma olasılığı "*1-p*" dir.

X, başarı olasılığı *p=P(X= 1)* olan *Bernoulli* dağılımına sahip rassal bir değişken ise, kısaca *X ~ Bernoulli (p)* olarak ifade edilir.

Bernoulli (p) dağılımına sahip *X* rassal değişkeninin olasılık dağılımı,

$$P(X = x) = \begin{cases} 1 - p, & x = 0 \\ p, & x = 1 \end{cases}$$

veya daha genel bir ifade ile

$$P(X = x) = p^x (1 - p)^{1-x}, \ x = 0,1$$

şeklinde tanımlanır. Burada, yeniden hatırlatmak gerekirse

p: Başarı olasılığı

x: Başarı sayısı

dır.

Örnek 8: *Hilesiz (ağırlık merkezi ile oynanmamış) bir para atılıyor. Bu deneyin yazı (Y) ve tura (T) olmak üzere iki sonucu vardır. T gelmesi "başarı" olarak kabul ediliyor ve*

X: T sayısı

olarak tanımlanıyor. Başarı olasılığını, bir başka deyişle p=P(X=1)'i bulmak için A olayının ve S örnek uzayının belirlenmesi gerekir. Bu deneyde, örnek uzay,

S= {T, Y}

ve A olayı örnek uzayda başarı olarak tanımlanan elemanların bir kümesi olduğundan

A= {T}

dir. Burada,

$n_A = 1$ ve $n = 2$

dir. Dolayısıyla,

$$p = P(X=1) = P(A) = \frac{n_A}{n} = \frac{1}{2} \text{ ve } 1-p = P(X=0) = 1 - P(X=1) = \frac{1}{2}$$

olur.

Sonuç olarak, X ~ Bernoulli (1/2) dir ve X rassal değişkeninin olasılık dağılımı,

$$P(X=x) = \begin{cases} 1/2, & x=0 \\ 1/2, & x=1 \end{cases}$$

veya daha genel bir ifade ile

$$P(X=x) = \left(\frac{1}{2}\right)^{x} \left(\frac{1}{2}\right)^{1-x}, \; x = 0,1$$

olarak elde edilir.

Örnek 9: *Hilesiz (ağırlık merkezi ile oynanmamış) iki zar atılıyor. Bu deneyde oyuncunun oyunu kazanabilmesi için (6,6) atması gerekiyor. Oyuncu için başarı (6,6) atmak (X= 1), başarısızlık ise (6,6) atamamak (X= 0) olarak tanımlanmaktadır.*

Bu deneyde,

A= {(6, 6)} ve $n_A = 1$

dir. Örnek uzay S ve S'nin eleman sayısı için bkz. ***Örnek 1.*** *Dolayısıyla, X= 1 olasılığı*

$$p = P(X=1) = P(A) = \frac{n_A}{n} = \frac{1}{36}$$

ve X=0 olasılığı da

$$1-p = P(X=0) = 1 - P(X=1) = \frac{35}{36}$$

olarak bulunur.

Bu örnekte, $X \sim Bernoulli\left(\frac{1}{36}\right)$ *ve X rassal değişkeninin olasılık dağılımı,*

$$P(X=x)=\left(\frac{1}{36}\right)^{x}\left(\frac{35}{36}\right)^{1-x},\ x=0,1$$

şeklinde ifade edilir.

Örnek 10: *Kalem üreticisi bir firmanın ürettiği 100 kalemden 97 tanesinin kusursuz olduğu biliniyor. Üretim anında rastgele seçilen bir kalemin kusursuz olması başarı (X= 1), kusurlu olması ise başarısızlık (X= 0) olarak tanımlanmaktadır.*

Bu deneyde, K kusursuz kalemleri, KL de kusurlu kalemleri göstermek üzere

$$A=\{K_1,K_2,...,K_{97}\}\ ve\ S=\{K_1,K_2,...,K_{97},KL_1,KL_2,KL_3\}$$

dir. Buradan,

n_A *= 97 ve n= 100*

bulunur. Dolayısıyla,

$$p=P(X=1)=P(A)=\frac{n_A}{n}=\frac{97}{100}=0.97$$

ve

$$1-p=P(X=0)=1-P(X=1)=0.03$$

tür.

Bu örnekte, X ~ Bernoulli (0.97) ve X rassal değişkeninin olasılık dağılımı,

$$P(X=x)=(0.97)^{x}(0.03)^{1-x},\ x=0,1$$

şeklinde ifade edilir.

Binom Dağılımı

Binom dağılımı, *n* tane **bağımsız ve aynı dağılımlı** *(independently and identically distributed) Bernoulli* rassal değişkeninden elde edilen başarı sayısını modellemek için kullanılan kesikli bir dağılımdır. Burada, aynı dağılımlı kelimesi, her bir Bernoulli denemesi için başarı (ya da başarısızlık) olasılığının aynı kaldığı anlamındadır.

Binom dağılımı, uygulama problemlerinde oldukça sık karşılaşılan bir dağılım olduğundan, kesikli dağılımlar içinde önemli bir yer tutar.

Matematiksel olarak ifade etmek gerekirse $X_i \sim Bernoulli\ (p)\ (i=1,2,...,n)$ dağılımına sahip bağımsız rassal değişkenler olmak üzere

$Y = X_1+X_2+...+X_n$

olsun. Bu durumda, *Y* rassal değişkeninin dağılımı *Binom(n,p)*'dir ve kısaca, $Y \sim Binom(n,p)$ olarak gösterilir.

Binom(n,p) dağılımına sahip *Y* rassal değişkeninin olasılık dağılımı,

$$P(Y=y)=\binom{n}{y}p^{y}(1-p)^{n-y},\ y=0,1,\ldots,n$$

olarak tanımlanır. Burada, yeniden hatırlatmak gerekirse

n: Bernoulli denemelerinin sayısı

y: n Bernoulli denemesinden elde edilen başarı sayısı

p: Her bir Bernoulli denemesindeki başarı olasılığı

dır.

Binom(n,p) dağılımının elde edilişini daha iyi anlayabilmek için aşağıda verilen örneği inceleyelim.

Örnek 11: *3 tane bağımsız Bernoulli denemesi yapılıyor. Başarı olasılığının p olması halinde, 2 denemenin başarı ile sonuçlanma olasılığının*

$$P(Y=2)=\binom{3}{2}p^2(1-p)^1$$

olduğunu gösteriniz.

Çözüm: $X_i \sim$ *Bernoulli (p) (i= 1,2,3)* olmak üzere, $Y=\sum_{i=1}^{3} X_i$ olarak tanımlayalım. 2 tane *Bernoulli* denemesinin başarı ile sonuçlanması, bir başka deyişle $Y=2$ olabilmesi için olası tüm durumlar aşağıdaki tabloda gösterilmiştir.

Durum	x_1	x_2	x_3	Y
1	1	1	0	2
2	1	0	1	2
3	0	1	1	2

Örneğin, Durum 1'de, $x_1=1$, $x_2=1$ ve $x_3=0$ olarak ifade edilen Bernoulli denemelerinin sonuçları, 1. ve 2. denemelerde başarı, 3. denemede ise başarısızlık elde edildiği anlamında kullanılmıştır. Diğer durumlar da benzer şekilde yorumlanmalıdır.

Aşağıda, her bir durum için, $P(Y=2)$ olasılığı hesaplanmıştır. Bütün durumlarda, X_1, X_2 ve X_3 rassal değişkenlerinin bağımsız olmalarından dolayı olasılığın çarpım kuralından yararlanılmıştır.

Durum 1.
$$\begin{aligned} P(Y=2) &= P(X_1=1, X_2=1, X_3=0) \\ &= P(X_1=1)P(X_2=1)P(X_3=0) \\ &= p.p.(1-p) \\ &= p^2.(1-p) \end{aligned}$$

Durum 2.
$$\begin{aligned} P(Y=2) &= P(X_1=1, X_2=0, X_3=1) \\ &= P(X_1=1)P(X_2=0)P(X_3=1) \\ &= p.(1-p).p \\ &= p^2.(1-p) \end{aligned}$$

Durum 3.
$$\begin{aligned} P(Y=2) &= P(X_1=0, X_2=1, X_3=1) \\ &= P(X_1=0)P(X_2=1)P(X_3=1) \\ &= (1-p).p.p \\ &= p^2.(1-p) \end{aligned}$$

Görüldüğü ve beklenildiği gibi, bütün durumlarda 2 tane başarı ve 1 tane başarısızlık elde etme olasılığı aynıdır ve $p^2(1\text{-}p)$ ye eşittir. Bernoulli denemelerinden elde edilen başarı ve başarısızlıkların sırası değil, sayısı önemli olduğundan $P(Y= 2)$

olasılığı, 3 farklı durum için elde edilen *$p^2(1-p)$* olasılıklarının toplamına eşittir. Buradan,

$$P(Y=2)=3.p^2.(1-p)$$

dir. Buradaki 3 çarpanı, 3 *Bernoulli* denemesinden 2 tanesinin başarı ile sonuçlanmasının, bir başka deyişle, 3'ün 2'li kombinasyonu kadar farklı şekilde gerçekleşebileceğini gösteren sabittir. Dolayısıyla,

$$P(Y=2)=\binom{3}{2}p^2(1-p)$$

olarak ifade edilir.

Örnek 12: *Aynı anda hilesiz (ağırlık merkezi ile oynanmamış) 3 para atılıyor. Y rassal değişkeni tura (T) sayısı olarak tanımlanırsa,* $Y \sim Binom\left(n=3, p=\frac{1}{2}\right)$ *olur. Bu durumda, Y rassal değişkeninin olasılık dağılımı,*

$$P(Y=y)=\binom{3}{y}\left(\frac{1}{2}\right)^y\left(\frac{1}{2}\right)^{3-y}, \; y=0,1,2,3$$

olarak ifade edilir. Bu örnek için,

a) Hiç T gelmemesi olasılığını bulunuz.

b) En az 1 tane T gelmesi olasılığını bulunuz.

Çözüm:

a) Para atma deneyinde hiç T gelmemesi olasılığı, *Y* rassal değişkeninin 0 değerini alması ile eşdeğerdir. Dolayısıyla,

P (3 paranın 3'ünün de T gelmemesi)= *P(Y= 0)*

$$=\binom{3}{0}\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^{3-0}=\frac{1}{8}$$

olur.

Yorum: Aynı anda atılan hilesiz 3 paranın 3'ünün de T gelmemesi olasılığı 1/8'dir.

b) Para atma deneyinde, en az 1 tane T gelmesi olasılığı iki farklı yaklaşımla hesaplanabilir.

i) Birinci yaklaşım:

P (3 paradan en az 1 inin T gelmesi) = *$P(Y\geq 1)$*

$= P(Y=1)+P(Y=2)+P(Y=3)$

$$=\frac{3}{8}+\frac{3}{8}+\frac{1}{8}=\frac{7}{8}$$

dir.

ii) İkinci yaklaşım:

P (3 paradan en az 1 inin T gelmesi)= *P*($Y \geq 1$)

P (3 paradan en az 1 inin T gelmesi)= *P(Y=1)+P(Y=2)+P(Y=3)*

dir. *P(Y= y)* olasılık dağılımı olduğundan

P(Y=0)+P(Y= 1)+P(Y=2)+P(Y=3)=1

eşitliği sağlanmaktadır. Buradan,

P(Y= 1)+P(Y=2)+P(Y=3)=1-P(Y=0)

olduğu görülür. Dolayısıyla,

$$1 - P(Y = 0) = 1 - \frac{1}{8} = \frac{7}{8}$$

olarak bulunur.

Her iki yaklaşımla da aynı sonuca ulaşıldığından, hesaplama kolaylığı bakımından kolay olan tercih edilir. (i) ve (ii)'de kullanılan olasılık değerleri, aşağıda verilen olasılıklar yardımıyla hesaplanmıştır.

$$P(Y = 0) = \binom{3}{0}\left(\frac{1}{2}\right)^{0}\left(\frac{1}{2}\right)^{3-0} = \frac{1}{8}$$

$$P(Y = 1) = \binom{3}{1}\left(\frac{1}{2}\right)^{1}\left(\frac{1}{2}\right)^{3-1} = \frac{3}{8}$$

$$P(Y = 2) = \binom{3}{2}\left(\frac{1}{2}\right)^{2}\left(\frac{1}{2}\right)^{3-2} = \frac{3}{8}$$

ve

$$P(Y = 3) = \binom{3}{3}\left(\frac{1}{2}\right)^{3}\left(\frac{1}{2}\right)^{3-3} = \frac{1}{8}$$

Yorum: Aynı anda atılan hilesiz 3 paradan en az 1 tanesinin T gelmesi olasılığı 7/8'dir.

Örnek 13: *Yeni geliştirilen bir füze, hedefin 50 m yakınına düştüğünde hedefi imha etmektedir. Füzenin hedefi imha etme olasılığı 0.40'tır. Prototip olarak üretilen 5 tane füze yapay bir hedefe atılıyor. Buna göre, hedefe atılan*

a) *5 füzeden 1 tanesinin hedefi imha etme olasılığını bulunuz.*

b) *5 füzeden en az 4 tanesinin hedefi imha etme olasılığını bulunuz.*

Çözüm: Bu örnekte,

Y: Hedefin imha edilme sayısı

olarak tanımlanırsa, *Y ~ Binom(n=5, p=0.40)* olduğu görülür.

a) Füze atma deneyinde,

$$P(\text{hedefin 1 kez imha edilmesi}) = P(Y=1)$$
$$= \binom{5}{1}(0.40)^1(0.60)^4$$
$$= \frac{5!}{1!4!}\cdot(0.40)(0.1296)$$
$$= 0.2592$$

olarak bulunur.

Yorum: 5 füzeden 1 tanesinin hedefi imha etme olasılığı 0.2592 ya da denk olarak %25.92'dir.

b) Füze atma deneyinde,

$$P(\text{hedefin en az 4 kez imha edilmesi}) = P(Y \geq 4)$$
$$= P(Y=4) + P(Y=5)$$
$$= \binom{5}{4}(0.40)^4(0.60)^1 + \binom{5}{5}(0.40)^5(0.60)^0$$
$$= \frac{5!}{4!1!}\cdot(0.0256).(0.60) + \frac{5!}{5!0!}\cdot(0.0102).(1.0)$$
$$= 0.0768 + 0.0102$$
$$= 0.0870$$

olarak bulunur.

Yorum: 5 füzeden en az 4 tanesinin hedefi imha etme olasılığı 0.0870 ya da denk olarak %8.70'tir.

Örnek 14: *Bir yumurta firmasının satışa sunduğu 30'luk kolilerdeki yumurtaların %95'i kırılmadan tüketiciye ulaşıyor. Bir müşteri, bu firmaya ait 4 koli yumurta satın alıyor ve eve gittiğinde her bir koliden rassal olarak 1 tane yumurta seçiyor. Bu yumurtalardan en fazla 3 tanesinin sağlam olma olasılığını bulunuz.*

Çözüm: Bu örnekte,
Seçilen
Y: Sağlam yumurta sayısı

olarak tanımlanırsa, *Y ~ Binom (n=4, p=0.95)* olduğu görülür. Rassal olarak seçilen 4 yumurtadan en fazla 3 tanesinin sağlam olma olasılığı *P(Y ≤ 3)* olarak ifade edilir. Dolayısıyla,

$$P(Y \leq 3) = P(Y=0) + P(Y=1) + P(Y=2) + P(Y=3)$$
$$= 1 - P(Y > 3)$$
$$= 1 - P(Y=4)$$
$$= 1 - \binom{4}{4}(0.95)^4(0.05)^0$$
$$= 1 - \frac{4!}{4!0!}\cdot(0.8145).(1.0)$$
$$= 1 - 0.8145$$
$$= 0.1855$$

olarak bulunur.

Yorum: 4 yumurtadan en fazla 3 tanesinin sağlam olma olasılığı 0.1855 ya da denk olarak %18.55'tir.

SIRA SİZDE 3

1. Belli bir hastalık türü, 0-1 yaş arası bebeklerin %10'unda kalıcı hasar bırakıyor. Bu hastalığın,
a) 4 bebekten, 2 tanesinde kalıcı hasar bırakması olasılığını bulunuz.
b) 4 bebekten, en az 1 en fazla 3 tanesinde kalıcı hasar bırakması olasılığını bulunuz.
2. Bir bilgi yarışmasında, 5 tane sorunun 3 tanesini yapan finale kalıyor. Sorular a, b, c ve d olmak üzere 4 seçenekten oluşuyor. Yarışmacıların her bir soruya doğru cevap verme olasılığı 0.25'tir. Bu yarışmaya katılan bir yarışmacının finale kalma olasılığını bulunuz.

Poisson Dağılımı

Poisson dağılımı, bir olayın, belirlenen bir zaman ya da uzay (uzunluk, alan, hacim gibi) aralığında gerçekleşme sayısını modellemek için kullanılan kesikli bir dağılımdır.

Poisson dağılımı, olasılık ve istatistik teorisinde yaygın olarak kullanılan kesikli bir dağılımdır. Bir olayın, belirlenen bir zaman ya da uzay (uzunluk, alan, hacim gibi) aralığında gerçekleşme sayısını modellemek için kullanılır, bkz. Simeon Denis Poisson (1837). İlgilenilen aralık uzunluğu, bir "birim" olarak ifade edilirse zamanla ilgili aralıklar "birim zaman", uzayla ilgili aralıklar ise "birim uzay" olarak ifade edilir.

- *Birim zamana örnek olarak;*
  Bir hafta, altı ay, bir yıl
- *Birim uzaya örnek olarak ise;*
  Bir metre (uzunluk), bir dönüm (alan), 1/2 metre küp (hacim) v.b.

verilebilir.

Aşağıda, Poisson dağılımı kullanılarak modelleme yapılabilecek bazı olaylara örnekler verilmiştir.

- Dünyaya, *bir haftada* (birim zaman) düşen göktaşı sayısı.
- Bir kavşakta, *altı ayda* (birim zaman) meydana gelen trafik kazası sayısı.
- Bir maden ocağında, *bir yılda* (birim zaman) meydana gelen ve yaralanmayla sonuçlanan kaza sayısı.
- *Bir metre* (birim uzunluk) uzunluğunda, bir çelik halattaki üretimden kaynaklanan hata sayısı.
- *İki dönüm* (birim alan) büyüklüğünde bir domates serasındaki hastalıklı fide sayısı.
- 1/2 *metreküp* (birim hacim) büyüklüğünde bir akvaryumdaki hasta Japon balığı sayısı.

Örneklerden de anlaşılabileceği üzere, Poisson dağılımı nadir (seyrek) gerçekleşen olayların modellenmesinde kullanılan bir dağılımdır.

Poisson Dağılımının Kullanımına İlişkin Bazı Varsayımlar

- Birim zaman ya da birim uzayda gerçekleşen olayların sayısı **ayrık** *(disjoint)* olmak kaydıyla diğer bir birim zaman ya da uzayda gerçekleşen olayların sayısından bağımsızdır.
- Birim zaman ya da birim uzayda gerçekleşen olayların ortalama sayısı **düzgün** *(uniform)* dür. Örneğin, birim zamanda ortalama x kadar olay gerçekleşiyorsa birim zamanın yarısında ortalama $x/2$, birim zamanın iki katında ise ortalama $2x$ kadar olay gerçekleşir.
- Kısa bir zaman aralığında bir olayın gerçekleşme olasılığı, birim zamanın toplam uzunluğu ile orantılıdır. Bir başka ifadeyle, birim zaman çok kısa ise, birim zamanda birden fazla olayın gerçekleşme olasılığı sıfıra yaklaşır.

Bkz., Bradley (2007).

X, Poisson dağılımına sahip bir rassal değişken ise, kısaca $X \sim$ *Poisson* (λ) olarak gösterilir. *Poisson* (λ) dağılımına sahip *X* rassal değişkeninin olasılık dağılımı,

$$P(X = x) = \frac{e^{-\lambda}\lambda^{x}}{x!},\ x = 0,1,2,...;\ \lambda > 0$$

olarak tanımlanır. Burada,

λ: Birim zaman ya da birim uzayda gerçekleşen ortalama olay sayısı

x: Birim zaman ya da birim uzayda gerçekleşen olay sayısı

e: Euler sayısı (yaklaşık değeri ≅ 2.718...)

dir. (λ Yunan alfabesinde bir harftir ve *"lamda"* şeklinde okunur.)

Örnek 15: *Ankara Tandoğan meydanında altı ayda ortalama 5 kaza olduğu biliniyorsa, önümüzdeki altı ayda en az 2 kaza olma olasılığını bulunuz.*

Çözüm: Bu örnekte,

X: Tandoğan Meydanı'nda altı ayda meydana gelen kaza sayısı

olarak tanımlanırsa, $X \sim$ *Poisson* (λ= 5) olduğu görülür. Önümüzdeki altı ayda Tandoğan Meydanı'nda en az 2 kaza olma olasılığı

$$\begin{aligned} P(\text{en az 2 kaza}) &= P(X \geq 2) \\ &= P(X=2) + P(X=3) + P(X=4) + ... \end{aligned}$$

dır.

P(X= x) olasılık dağılımı olduğundan

$P(X=0)+P(X= 1)+P(X=2)+P(X=3)+...=1$

dir. Buradan, $P(X=2)+P(X=3)+...=1-\{P(X= 0)+P(X= 1)\}$ olduğu görülür. Dolayısıyla,

$$\begin{aligned} 1-\{P(X=0)+P(X=1)\} &= 1-\left\{\frac{e^{-5}5^{0}}{0!}+\frac{e^{-5}5^{1}}{1!}\right\} \\ &= 1-6e^{-5} \\ &= 0.9596 \end{aligned}$$

olarak hesaplanır.

Yorum: Ankara Tandoğan Meydanı'nda önümüzdeki 6 ayda en az 2 kaza olma olasılığı 0.9596 ya da denk olarak %95.96'dır.

Örnek 16: *Domates seralarına ekilen fidelerin dönüm başına ortalama 3 tanesinin fire verdiği (kuruduğu) bilinmektedir. Buna göre, Antalya'nın Gazipaşa ilçesinde seracılık yapan bir çiftçinin 2 dönüm büyüklüğündeki serasında*

a) *5 tane fidenin kuruma olasılığını bulunuz.*

b) *En fazla 2 tane fidenin kuruma olasılığını bulunuz.*

Çözüm: Bu örnekte,

X: Hastalıklı fide sayısı

olarak tanımlanırsa, $X \sim$ *Poisson* (λ= 3) olduğu görülür. "1 dönüm" büyüklüğünde bir serada ortalama 3 tane domates fidesi fire veriyorsa "2 dönüm" büyüklüğünde bir serada 2x3=6 tane domates fidesi fire verir. Dolayısıyla, "birim alan" olarak “2 dönüm” büyüklüğünde bir sera düşünülürse *X* rassal değişkeninin dağılımının *X* ~ *Poisson* (λ= 6) olduğu görülür.

a) *X* ~ *Poisson* (λ= 6) dağılımının olasılık dağılımı kullanılarak,

$$\begin{aligned} P(5 \text{ tane hastalıklı fide}) &= P(X=5) \\ &= \frac{e^{-6}6^5}{5!} \\ &= 0.1606 \end{aligned}$$

olarak bulunur.

Yorum: Çiftçinin 2 dönüm büyüklüğündeki serasında 5 tane domates fidesinin kuruma olasılığı 0.1606'ya da denk olarak %16.06'dır.

b) a seçeneğine benzer şekilde

$$\begin{aligned} P(\text{en fazla 2 tane hastalıklı fide}) &= P(X \le 2) \\ &= P(X=0)+P(X=1)+P(X=2) \\ &= \frac{e^{-6}6^0}{0!}+\frac{e^{-6}6^1}{1!}+\frac{e^{-6}6^2}{2!} \\ &= 0.0620 \end{aligned}$$

olarak bulunur.

Yorum: Çiftçinin 2 dönüm büyüklüğündeki serasında en fazla 2 tane domates fidesinin kuruması olasılığı 0.0620'ya da denk olarak %6.20'dir.

SIRA SİZDE 4

1. Araba lastiği üreten bir fabrikada haftada beş gün üretim yapılmaktadır. Haftalık bazda, üretilen lastiklerin ortalama 4 tanesinin kusurlu olduğu bilinmektedir. Bu fabrikada, haftanın belli bir gününde üretilen lastiklerin en az 3 tanesinin kusurlu olma olasılığını bulunuz.

2. Bir oyuncak üreticisinin, bir günde ürettiği oyuncakların, ortalama 5 tanesi üretim hatası nedeniyle iade ediliyor. Belli bir günde, bu oyuncaklardan, en az 2 en fazla 4 tanesinin iade edilme olasılığını bulunuz.

Binom Dağılımının Poisson Dağılımına Yakınsaması

Binom (n,p) dağılımında, bağımsız *Bernoulli* denemelerinin sayısı olarak tanımlanan n değeri çok büyük, buna karşılık başarı olasılığı olarak tanımlanan p değeri çok küçük olabilir. Bu gibi durumlarda, hesaplama zorlukları kaçınılmaz olduğundan, Binom dağılımından elde edilen olasılıklar yerine daha basit ve daha kolayca hesaplanabilen Poisson dağılımından elde edilen olasılıklar kullanılabilir.

Binom (n,p) dağılımında, ortalama (beklenen) başarı sayısı $\lambda=np$ olarak ifade edilir ve bu eşitlik kullanılarak p yerine $\frac{\lambda}{n}$ yazılırsa *Binom (n,p)* dağılımı

$$P(Y=y)=\binom{n}{y}\left(\frac{\lambda}{n}\right)^y\left(1-\frac{\lambda}{n}\right)^{n-y}, \quad y=0,1,2,\ldots,n$$

şeklinde ifade edilir. Buradan, *Y* rassal değişkeninin olasılık dağılımının n değeri sonsuza yaklaşırken, *Poisson* (λ) dağılımına yakınsadığı, bir başka deyişle,

$$P(Y=y)=\frac{e^{-\lambda}\lambda^y}{y!} \qquad y=0,1,2,\ldots$$

olduğu görülebilir.

DİKKAT

Büyük *n* ve küçük *p* değerleri için *Binom (n,p)* dağılımı, *Poisson* (λ) dağılımına yakınsar.

DİKKAT

- **n'in yeterince büyük olduğu durumlarda, Poisson dağılımından elde edilen olasılıklar, Binom dağılımından elde edilen olasılıklara yaklaşık olarak eşit, n değeri sonsuza yaklaşırken ise tam olarak eşittir.**
- **Uygulama problemlerinde, Binom dağılımının, Poisson dağılımına yakınsama özelliğini kullanabilmek için $\lambda = np \leq 7$ koşulunun sağlanması gerektiği genel kabul görmüştür, bkz. Newbold ve ark. (2010).**

Örnek 17: *Y rassal değişkeninin, n=1000 ve p=0.002 olan Binom dağılımına sahip olduğu bilinmektedir. Buna göre, P(Y=4) olasılığını bulunuz.*

Çözüm: Bu örnek iki farklı yaklaşımla çözülebilir. Birinci yaklaşımda *Binom* dağılımı, ikinci yaklaşımda ise *Binom* dağılımının *Poisson* dağılımına yakınsaması özelliği kullanılmıştır.

i) *Birinci yaklaşım:*

Binom dağılımından yararlanarak,

$$P(Y=4) = \binom{1000}{4}(0.002)^4(0.998)^{996}$$
$$= 0.0903$$

bulunur.

ii) *İkinci yaklaşım:*

n=1000 yeterince büyük ve *p*=0.002 başarı olasılığı yeterince küçük olduğundan, *Binom* dağılımının, *Poisson* dağılımına yakınsaması özelliğinden yararlanarak *P(Y=4)* olasılığı hesaplanabilir. Binom dağılımının Poisson dağılımına yakınsama özelliği kullanılarak, $\lambda = np = 1000.(0.002) = 2$ bulunur. Poisson dağılımında, λ yerine 2 yazılırsa

$$P(Y=4) = \frac{e^{-2}2^4}{4!}$$
$$= 0.0902$$

elde edilir.

Görüldüğü gibi, her iki yaklaşım da benzer sonuçlar vermiştir. Bununla beraber, Binom dağılımı kullanılarak *P(Y=4)* olasılığını hesaplamak, $\binom{1000}{4}$, $(0.002)^4$ ve $(0.998)^{996}$ ifadelerinden dolayı oldukça zaman alıcıdır. (ii) de verilen yaklaşım ise hesaplama kolaylığı bakımından oldukça avantajlıdır.

Örnek 18: *Bir firma seramik saksılar üretiyor. Üretilen saksıların kusurlu olma olasılığı 0.004 ise üretilen 100 tane saksıdan en az 3 tanesinin kusurlu olma olasılığını, Binom dağılımının Poisson dağılımına yakınsama özelliğinden yararlanarak bulunuz.*

Çözüm: Bu örnekte,

Y: Kusurlu saksı sayısı

olarak tanımlanırsa *Y ~ Binom (n=100, p=0.004)* olduğu görülür.

$$P(Y \geq 3) = P(Y = 3) + P(Y = 4) + \ldots$$
$$= 1 - \{P(Y = 0) + P(Y = 1) + P(Y = 2)\}$$

olasılığının hesaplanabilmesi için

$$P(Y = 0) = \binom{100}{0}(0.004)^0(0.996)^{100}$$

$$P(Y = 1) = \binom{100}{1}(0.004)^1(0.996)^{99}$$

ve

$$P(Y = 2) = \binom{100}{2}(0.004)^2(0.996)^{98}$$

olasılıklarının hesaplanması gereklidir. Ancak, n yeterince büyük ve p olasılığı da yeterince küçük olduğundan, istenen olasılık değerini, Binom dağılımının Poisson dağılımına yakınsaması özelliğinden yararlanarak çözmek hesap kolaylığı bakımından daha avantajlıdır. Bu durumda,

$\lambda = np = 100.(0.004) = 0.4$

olur. Buradan,

$$P(Y \geq 3) = 1 - \{P(Y = 0) + P(Y = 1) + P(Y = 2)\}$$
$$= 1 - \left\{\frac{e^{-0.4}(0.4)^0}{0!} + \frac{e^{-0.4}(0.4)^1}{1!} + \frac{e^{-0.4}(0.4)^2}{2!}\right\}$$
$$= 0.0079$$

olarak bulunur.

Yorum: Firma tarafından üretilen 100 saksıdan en az 3 tanesinin kusurlu olma olasılığı 0.0079 ya da denk olarak %0.79'dur.

SIRA SİZDE 5

1. Bir firma, hırsızlara karşı alarm cihazı üretmektedir. Üretilen cihazın kusurlu olma olasılığı 0.01 ise üretilen 400 tane alarm cihazından en fazla 2 tanesinin kusurlu olma olasılığını, *Binom* dağılımının *Poisson* dağılımına yakınsama özelliğinden yararlanarak bulunuz.
2. Yılın herhangi bir gününde elektrik kesintisi yaşanması olasılığı 0.005 ise 1 yılda en az 1, en fazla 3 gün elektrik kesintisi yaşanma olasılığını, *Binom* dağılımının *Poisson* dağılımına yakınsama özelliğinden yararlanarak bulunuz. (Not: Bir yıl 365 gün olarak alınacaktır.)

KESİKLİ RASSAL DEĞİŞKENLERİN ORTALAMA, VARYANS VE STANDART SAPMASI

Dağılımı karakterize etmeleri bakımından, **ortalama** *(mean)* ve **varyans** *(variance)* kavramları son derece önemlidir. Ortalama ve varyans, sırasıyla verinin *konumunun (location)* ve *değişkenliğinin (variation)* birer ölçüsü olarak tanımlanırlar.

Ortalama yerine beklenen değer (expectation) ifadesi de kullanılır ve $E(X)$ sembolü ile gösterilir. Verideki homojenliğin (ya da heterojenliğin) bir ifadesi olan **varyans** ise $V(X)$ sembolü ile gösterilir. Varyans değeri büyük ise, verideki değerlerin (veya eşdeğer olarak rassal değişkenin aldığı değerlerin) birbirinden oldukça farklı (heterojen) olduğu, varyans küçük ise verideki değerlerin birbirine benzer (homojen) olduğu söylenir.

Ortalama ve varyans kavramları matematiksel olarak

i) $E(X) = \sum_x xP(X = x)$

ii) $V(X) = E(X^2) - [E(X)]^2$

şeklinde tanımlanır. (ii) eşitliğindeki *$E(X^2)$*,

$$E(X^2) = \sum_x x^2 P(X = x)$$

formülü yardımıyla hesaplanır.

Varyansın karekökü, *standart sapma (standard deviation)* olarak tanımlanır ve *sd(X)* sembolü ile gösterilir. Bir başka deyişle,

$$sd(X) = \sqrt{V(X)}$$

dir.

Bu gösterimlere alternatif olarak ortalama, varyans ve standart sapma, sırasıyla

μ, σ^2 ve σ

sembolleri ile de gösterilir. Literatürde, her iki gösterim de yaygın olarak kullanılır.

Varyans, yukarıda verilen formüle alternatif olarak

$$\sigma^2 = V\left(X\right) = E\left[(X - \mu)^2\right] = \sum_x (x - \mu)^2 P(X = x)$$

formülü yardımıyla da hesaplanabilir. Matematiksel işlemler yapıldığında iki formülün denk olduğu görülebilir.

Şimdi, standart sapma ve varyans kavramlarını daha iyi anlayabilmek için aşağıda verilen örneği inceleyelim.

Örnek 19: *Anadolu Üniversitesi, İstatistik ve Matematik Bölümlerinde okuyan 1. sınıf öğrencilerine güz döneminin sonunda İngilizce sınavı yapılıyor. İstatistik Bölümü öğrencilerinin İngilizce sınav sonuçları 100 üzerinden 70 ile 80 aralığında yoğunlaşıyor. 60'ın altında ve 90'ın üzerinde not alan kimse çıkmıyor. Matematik Bölümü öğrencilerinin İngilizce sınav sonuçları ise herhangi bir aralıkta belirgin bir yoğunlaşma olmadan, 0 ile 100 arasında rassal olarak değişiyor.*

Bu örnekte, İstatistik Bölümü öğrencilerinin İngilizce sınav sonuçlarındaki değişkenliğin (varyansın ya da standart sapmanın) Matematik Bölümü öğrencilerinin İngilizce sınav sonuçlarındaki değişkenlikten daha az olduğu söylenir. Bir başka deyişle, İstatistik Bölümü öğrencileri, İngilizce bilgi düzeyleri bakımından Matematik Bölümü öğrencilerine göre daha benzerdir (homogeneous) denir.

Şimdi, herhangi bir kesikli *X* rassal değişkeninin ortalama, varyans ve standart sapmasının nasıl hesaplandığını görelim.

Örnek 20: *X kesikli rassal değişkeninin olasılık dağılımı aşağıda verilmiştir. Buna göre, X rassal değişkeninin*

a) *Beklenen değerini (ortalamasını) bulunuz.*

b) *Varyansını ve standart sapmasını bulunuz.*

$X = x$	1	2	3	4	5
$P(X = x)$	2/14	5/14	4/14	2/14	1/14

Çözüm:

a) *X* kesikli rassal değişkeninin beklenen değeri (ortalaması),

$$\begin{aligned}\mu = E(X) &= \sum_x xP(X=x),\ x=1,2,3,4,5\\ &= \sum_{x=1}^{5} xP(X=x)\\ &= 1\cdot P(X=1)+2\cdot P(X=2)+3\cdot P(X=3)+4\cdot P(X=4)+5\cdot P(X=5)\\ &= 1\cdot\frac{2}{14}+2\cdot\frac{5}{14}+3\cdot\frac{4}{14}+4\cdot\frac{2}{14}+5\cdot\frac{1}{14}\\ &= \frac{37}{14}=2.6429\end{aligned}$$

dur.

b) *X* kesikli rassal değişkeninin varyansını hesaplamak için öncelikle $E(X^2)$'nin hesaplanması gereklidir.

$$\begin{aligned}E(X^2) &= \sum_{x=1}^{5} x^2P(X=x)\\ &= 1^2\cdot P(X=1)+2^2\cdot P(X=2)+3^2\cdot P(X=3)+4^2\cdot P(X=4)+5^2\cdot P(X=5)\\ &= 1\cdot\frac{2}{14}+4\cdot\frac{5}{14}+9\cdot\frac{4}{14}+16\cdot\frac{2}{14}+25\cdot\frac{1}{14}\\ &= \frac{115}{14}=8.2143\end{aligned}$$

bulunur. $E(X^2)$ ve $E(X)$ ifadelerinin eşitleri varyans formülünde yerlerine yazılırsa

$$\sigma^2 = V(X) = E(X^2)-\left[E(X)\right]^2 = \frac{115}{14}-\left(\frac{37}{14}\right)^2 = \frac{241}{196}=1.2296$$

olur. Buradan, *X* kesikli rassal değişkeninin standart sapmasının

$$\sigma = sd(X) = \sqrt{V(X)} = 1.2296$$

olduğu görülür.

Örnek 21: *X ~ Bernoulli(p= 0.5)* ise *X* rassal değişkeninin beklenen değerini (ortalamasını), varyansını ve standart sapmasını bulunuz.

Çözüm: *X ~ Bernoulli(p= 0.5)* ise *X* rassal değişkeninin *olasılık dağılımı*

$$\begin{aligned}P(X=x) &= p^x(1-p)^{1-x},\ x=0,1\\ &= (0.5)^x(0.5)^{1-x},\ x=0,1\end{aligned}$$

olarak ifade edilir. Burada,

$$\begin{aligned}P(X=0) &= (0.5)^0(0.5)^1\\ &= 0.5\end{aligned}$$

ve

$$\begin{aligned}P(X=1) &= (0.5)^1(0.5)^0\\ &= 0.5\end{aligned}$$

olarak bulunur. *X* rassal değişkeninin beklenen değeri (ortalaması)

$$\begin{aligned}\mu = E(X) &= \sum_{x=0}^{1} xP(X=x)\\ &= 0.P(X=0)+1.P(X=1)\\ &= 0.(0.5)+1.(0.5)\\ &= 0.5\end{aligned}$$

tir. X rassal değişkeninin varyansını hesaplamak için öncelikle $E(X^2)$ nin hesaplanması gerekir.

$$\begin{aligned}E(X^2) &= \sum_{x=0}^{1} x^2 P(X=x)\\ &= 0^2.P(X=0)+1^2.P(X=1)\\ &= 0.(0.5)+1.(0.5)\\ &= 0.5\end{aligned}$$

olarak hesaplanır. $E(X^2)$ ve $E(X)$ ifadelerinin eşitleri varyans formülünde yerlerine yazılırsa

$$\sigma^2 = V(X) = E(X^2) - \left[E(X)\right]^2 = 0.5 - (0.5)^2 = 0.25$$

olur. Buradan, X kesikli rassal değişkeninin standart sapması

$$\sigma = sd(X) = \sqrt{0.25} = 0.5$$

olarak bulunur.

DİKKAT

Bernoulli (p) dağılımının ortalama ve varyansı, aşağıdaki eşitlikler yardımıyla kısa yoldan hesaplanabilir.
Ortalama: $\mu = E(X) = p$
Varyans: $\sigma^2 = V(X) = p(1-p)$

Örnek 22: *$Y \sim$ Binom ($n=3$, $p=0.5$) ise Y rassal değişkeninin beklenen değerini (ortalamasını), varyansını ve standart sapmasını bulunuz.*

Çözüm: *$Y \sim$ Binom ($n=3$, $p=0.5$)* ise Y rassal değişkeninin *olasılık dağılımı*

$$P(Y=y) = \binom{3}{y}(0.5)^y \left(1-0.5\right)^{3-y}, \; y = 0,1,2,3$$

şeklinde ifade edilir. Burada,

$$P(Y=0)=\frac{1}{8},\; P(Y=1)=\frac{3}{8},\; P(Y=2)=\frac{3}{8} \; ve \; P(Y=3)=\frac{1}{8}$$

olarak bulunur, bkz. **Örnek 12.** Y rassal değişkeninin beklenen değeri (ortalaması)

$$\begin{aligned}\mu = E(Y) &= \sum_{y=0}^{3} yP(Y=y)\\ &= 0.P(Y=0)+1.P(Y=1)+2.P(Y=2)+3.P(Y=3)\\ &= 0.\frac{1}{8}+1.\frac{3}{8}+2.\frac{3}{8}+3.\frac{1}{8} = 1.5\end{aligned}$$

bulunur. Y rassal değişkeninin varyansını hesaplamak için öncelikle $E(Y^2)$'nin hesaplanması gerekir.

$$E(Y^2) = \sum_{y=0}^{3} y^2 P(Y = y)$$
$$= 0^2.P(Y=0) + 1^2.P(Y=1) + 2^2.P(Y=2) + 3^2.P(Y=3)$$
$$= 0.\frac{1}{8} + 1.\frac{3}{8} + 4.\frac{3}{8} + 9.\frac{1}{8}$$
$$= 3$$

olarak hesaplanır. $E(Y)$ ve $E(Y^2)$ ifadelerinin eşitleri varyans formülünde yerlerine yazılırsa

$$\sigma^2 = V(Y) = E(Y^2) - \left[E(Y)\right]^2 = 3 - (1.5)^2 = 0.75$$

olduğu görülür. Buradan, Y kesikli rassal değişkeninin standart sapması

$$\sigma = sd(Y) = \sqrt{V(Y)} = \sqrt{0.75} = 0.866$$

bulunur.

DİKKAT

***Binom(n,p)* dağılımının ortalama ve varyansı, aşağıdaki eşitlikler yardımıyla kısa yoldan hesaplanabilir.**
Ortalama: $\mu = E(X) = np$
Varyans: $\sigma^2 = V(X) = np(1-p)$.

Örnek 23: *X ~ Poisson (λ= 5) ise X rassal değişkeninin beklenen değerini (ortalamasını), varyansını ve standart sapmasını bulunuz.*

Çözüm: X ~ *Poisson* (λ= 5) ise X rassal değişkeninin *olasılık dağılımı*

$$P(X = x) = \frac{e^{-5}5^x}{x!}, \; x = 0,1,2,...$$

olarak ifade edildiğinden,

$$P(X=0) = \frac{e^{-5}5^0}{0!}, \; P(X=1) = \frac{e^{-5}5^1}{1!}, \; P(X=2) = \frac{e^{-5}5^2}{2!}, ...$$

dir. X rassal değişkeninin beklenen değeri (ortalaması) ve varyansını bulmak için

$$E(X) = \sum_{x=0}^{\infty} xP(X = x)$$
$$= 0.P(X=0) + 1.P(X=1) + 2.P(X=2) + 3.P(X=3) + ...$$
$$= 0.\frac{e^{-5}5^0}{0!} + 1.\frac{e^{-5}5^1}{1!} + 2.\frac{e^{-5}5^2}{2!} + 3.\frac{e^{-5}5^3}{3!} + ...$$

ve

$$E(X(X-1)) = \sum_{x=0}^{\infty} x(x-1)P(X = x)$$
$$= 0.(0-1).P(X=0) + 1.(1-1).P(X=1) + 2.(2-1).P(X=2) + ...$$
$$= \frac{e^{-5}5^2}{(2-2)!} + \frac{e^{-5}5^3}{(3-2)!} + \frac{e^{-5}5^4}{(4-2)!} + ...$$

eşitliklerinin bulunması gereklidir. Bu eşitliklerin bulunması detaylı matematiksel işlemler gerektirdiğinden, *Poisson* (λ) dağılımının ortalama ve varyansını, aşağıdaki eşitlikler yardımıyla kısa yoldan hesaplamak daha uygun olacaktır:

Ortalama: $\mu = E(X) = \lambda$

Varyans: $\sigma^2 = V(X) = \lambda$.

Bu eşitlikler kullanılarak, λ parametresi 5 olan Poisson dağılımının beklenen değeri (ortalaması) ve varyansı sırasıyla,

$$\mu = E(X) = 5$$
$$\sigma^2 = V(X) = 5$$

şeklinde elde edilir. Buradan, X rassal değişkeninin standart sapması

$$\sigma = sd(X) = \sqrt{V(X)} = \sqrt{5} = 2.2361$$

bulunur.

SIRA SİZDE 6

1. *X* rassal değişkeninin olasılık dağılımı,

$X = x$	0	1	2	3
$P(X = x)$	0.22	c	0.29	0.06

olarak veriliyor. *X* rassal değişkeninin ortalamasını, varyansını ve standart sapmasını bulunuz.

2. Aşağıda verilen *X* rassal değişkenlerinin ortalamasını, varyansını ve standart sapmasını bulunuz.

a) $X \sim Binom\,(n=50,\ p=0.4)$

b) $X \sim Poisson\,(\lambda = 4)$

Özet

AMAÇ 1

Rassal değişken kavramını tanımlamak.
Rassal bir olayın (ya da deneyin) sonuçlarını, sayısal değerlerle ifade eden değişkene, **rassal değişken** (*random variable*) denir.
Rassal değişken, tanım kümesi **örnek uzay** (*sample space-S*), değer kümesi ise **reel sayılar** (***R***) olan bir fonksiyon olarak tanımlanır.

AMAÇ 2

Rassal değişkenlerin çeşitlerini (kesikli ve sürekli) ve aralarındaki farkı ayırt etmek.
Rassal değişkenler, aldıkları değerlere göre **kesikli** (discrete) ya da sürekli (continuous) olarak adlandırılırlar. Değer kümesi **sayılabilir** (countable) olan rassal değişkenler kesikli, **sayılamayan** (uncountable) olan rassal değişkenler ise sürekli olarak isimlendirilir.

AMAÇ 3

Olasılık dağılımı kavramını tanımlamak.
Olasılık dağılımı (*probability distribution*), X kesikli rassal değişkeninin aldığı değerler ile bu değerlere karşılık gelen olasılıkları ifade eder.
Bir fonksiyonun, olasılık dağılımı olarak tanımlanabilmesi için aşağıda verilen özelliklerin sağlanması gerekir:
(i) X kesikli rassal değişkeninin, herhangi bir x'e eşit olma olasılığı, 0 ile 1 arasında değişir. Bir başka deyişle,

$0 \leq P(X=x) \leq 1$, her x değeri için

koşulu sağlanmalıdır.
(ii) X kesikli rassal değişkeninin, x'in tüm olası değerlerine eşit olma olasılıklarının toplamı 1'e eşittir. Bir başka deyişle,

$$\sum_x P(X=x)=1$$

koşulu sağlanmalıdır.
Kesikli birikimli olasılık dağılımı (*discrete cumulative probability distribution*)

$$F_X(x)=P(X \leq x)=\sum_{i \leq x} P(X=i)$$

olarak tanımlanır. Bir başka deyişle, $F_X(x)$ fonksiyonu, X rassal değişkeninin belli bir x'ten daha küçük ya da eşit değer alması olasılığını ifade eder. Birikimli olasılık dağılımı, x'in bütün değerleri için 0 ile 1 arasında değer alır. Bir başka deyişle, $0 \leq F_X(x) \leq 1$'dir.

AMAÇ 4

Olasılık ve istatistik teorisinde yaygın olarak kullanılan bazı kesikli dağılımları (Bernoulli, Binom ve Poisson) tanımlamak ve bu dağılımlar yardımıyla olasılık değerlerini hesaplamak.
Bernoulli dağılımı, iki sonucu olan bir deneyi (*Bernoulli denemesi*) modellemek için kullanılan kesikli bir dağılımdır. Genellikle, bu sonuçlar "**başarı** (*success*)" ve "**başarısızlık** (*failure*)" olarak isimlendirilir. X rassal değişkeni "*başarı*" durumunda 1, "*başarısızlık*" durumunda ise 0 değerini alır. Bernoulli denemesinin başarı ile sonuçlanma olasılığı "p", başarısızlıkla sonuçlanma olasılığı "1-p" dir. Bernoulli (p) dağılımına sahip X rassal değişkeninin olasılık dağılımı,

$$P(X=x)=p^x(1-p)^{1-x},\ x=0,1$$

şeklinde tanımlanır. Burada, yeniden hatırlatmak gerekirse

p: Başarı olasılığı
x: Başarı sayısı

dır.
Binom dağılımı, n tane **bağımsız ve aynı dağılımlı** (*independently and identically distributed*) *Bernoulli* rassal değişkeninden elde edilen başarı sayısını modellemek için kullanılan kesikli bir dağılımdır. Burada, aynı dağılımlı kelimesi, her bir *Bernoulli* denemesi için başarı (ya da başarısızlık) olasılığının aynı kaldığı anlamındadır. *Binom* dağılımı, uygulama problemlerinde oldukça sık karşılaşılan bir dağılım olduğundan dolayı, kesikli dağılımlar içinde önemli bir yer tutar. *Binom* (n,p) dağılımına sahip Y rassal değişkeninin olasılık dağılımı,

$$P(Y=y)=\binom{n}{y}p^y(1-p)^{n-y},\ y=0,1,\ldots,n$$

olarak tanımlanır. Burada, yeniden hatırlatmak gerekirse

n: Bernoulli denemelerinin sayısı
y: n Bernoulli denemesinden elde edilen başarı sayısı
p: Her bir Bernoulli denemesindeki başarı olasılığıdır.
Poisson dağılımı, bir olayın, belirlenen bir zaman ya da uzay (uzunluk, alan, hacim gibi) aralığında ger-

çekleşme sayısını modellemek için kullanılan kesikli bir dağılımdır. *Poisson* dağılımı nadir (seyrek) gerçekleşen olayların modellenmesinde kullanılan bir dağılımdır. *X*, Poisson dağılımına sahip rassal bir değişken ise kısaca X~ *Poisson* (λ) olarak gösterilir. *Poisson* (λ) dağılımına sahip *X* rassal değişkeninin olasılık dağılımı,

$$P(X=x)=\frac{e^{-\lambda}\lambda^{x}}{x!},\quad x=0,1,2,...;\ \lambda>0$$

olarak tanımlanır. Burada,

λ: Birim zaman ya da birim uzayda gerçekleşen ortalama olay sayısı

x: Birim zaman ya da birim uzayda gerçekleşen olay sayısı

e: Euler sayısı (yaklaşık değeri ≅ 2.718...)

dir.

AMAÇ 5

Binom dağılımının, Poisson dağılımına yakınsama özelliğini kullanmak.

Binom (*n,p*) dağılımında, bağımsız *Bernoulli* denemelerinin sayısı olarak tanımlanan *n* değeri çok büyük, buna karşılık başarı olasılığı olarak tanımlanan *p* değeri çok küçük olabilir. Bu gibi durumlarda, hesaplama zorlukları kaçınılmaz olduğundan, Binom dağılımından elde edilen olasılıklar yerine daha basit ve daha kolayca hesaplanabilen Poisson dağılımından elde edilen olasılıklar kullanılabilir.

Binom (*n,p*) dağılımında, ortalama (beklenen) başarı sayısı λ=*np* olarak ifade edilir ve bu eşitlik kullanılarak *p* yerine $\frac{\lambda}{n}$ yazılırsa, *Binom* (*n,p*) dağılımı

$$P(Y=y)=\binom{n}{y}\left(\frac{\lambda}{n}\right)^{y}\left(1-\frac{\lambda}{n}\right)^{n-y}\quad ,y=0,1,2,...,n$$

şeklinde ifade edilir. Buradan, *Y* rassal değişkeninin olasılık dağılımının *n* değeri sonsuza yaklaşırken, *Poisson* (λ) dağılımına yakınsadığı, bir başka deyişle,

$$P(Y=y)=\frac{e^{-\lambda}\lambda^{y}}{y!},\ y=0,1,2,...$$

olduğu görülebilir.

AMAÇ 6

Kesikli rassal değişkenlerin ortalama, varyans ve standart sapmalarını hesaplamak.

Ortalama ve varyans, sırasıyla verinin **konumunun** (*location*) ve **değişkenliğinin** (*variation*) birer ölçüsü olarak tanımlanırlar.

Ortalama yerine **beklenen değer** (*expectation*) ifadesi de kullanılır ve *E*(*X*) sembolü ile gösterilir. Verideki homojenliğin (ya da heterojenliğin) bir ifadesi olan varyans ise *V*(*X*) sembolü ile gösterilir. Varyans değeri büyük ise, verideki değerlerin (veya eşdeğer olarak rassal değişkenin aldığı değerlerin) birbirinden oldukça farklı (heterojen) olduğu, varyans küçük ise verideki değerlerin birbirine benzer (homojen) olduğu söylenir.

Ortalama ve varyans kavramları matematiksel olarak

i) $E(X)=\sum_{x} xP(X=x)$

ii) $V(X)=E(X^2)-[E(X)]^2$

şeklinde tanımlanır. (ii) eşitliğindeki $E(X^2)$,

$$E(X^2)=\sum_{x} x^2P(X=x)$$

formülü yardımıyla hesaplanır.

Varyansın karekökü, **standart sapma** (*standard deviation*) olarak tanımlanır ve *sd*(*X*) sembolü ile gösterilir. Bir başka deyişle,

$$sd(X)=\sqrt{V(X)}$$

dir. Bu gösterimlere alternatif olarak ortalama, varyans ve standart sapma, sırasıyla

μ, σ^2 ve σ

sembolleri ile de gösterilir.

Bernoulli (*p*) dağılımının ortalama ve varyansı, aşağıdaki eşitlikler yardımıyla kısa yoldan hesaplanabilir.

Ortalama: $\mu=E(X)=p$

Varyans: $\sigma^2=V(X)=p(1-p)$.

Binom (*n,p*) dağılımının ortalama ve varyansı, aşağıdaki eşitlikler yardımıyla kısa yoldan hesaplanabilir.

Ortalama: $\mu=E(X)=np$

Varyans: $\sigma^2=V(X)=np(1-p)$.

Poisson (λ) dağılımının ortalama ve varyansı, aşağıdaki eşitlikler yardımıyla kısa yoldan hesaplanabilir.

Ortalama: $\mu=E(X)=\lambda$

Varyans: $\sigma^2=V(X)=\lambda$.

Kendimizi Sınayalım

1. Aşağıdakilerden hangisi kesikli bir rassal değişken kullanılarak **tanımlanamaz?**

a. Bir hastanedeki hasta sayısı
b. Bir maden ocağındaki kazaların sayısı
c. Bir şehirdeki hava sıcaklığı değerleri
d. Bir kasadaki çürük domateslerin sayısı
e. Yeni basılan bir kitabın her bir sayfasındaki yazım hatalarının sayısı

2. X rassal değişkeninin olasılık dağılımı aşağıda verildiği gibidir:

X	1	2	3	4	5
$P(X=x)$	0.08	0.3	0.16	0.06	c

Buna göre, c sabitinin değeri aşağıdakilerden hangisidir?

a. 0.3
b. 0.4
c. 0.5
d. 0.6
e. 0.7

3. X rassal değişkeninin olasılık dağılımı aşağıda verildiği gibidir:

X	0	1	2	3	4
$P(X=x)$	0.44	0.25	0.11	0.18	0.02

Buna göre, $P(1<X\leq3)$ olasılığı aşağıdakilerden hangisidir?

a. 0.29
b. 0.54
c. 0.69
d. 0.44
e. 0.66

4. Piyasa araştırması yapan bir bilgisayar firması, tüketicilerin %40'ının kendi ürettikleri bilgisayar markasını kullandıklarını tespit etmiştir. Tüketiciler arasından rassal olarak seçilen 5 kişiden 2 tanesinin bu firmanın ürettiği bilgisayar markasını kullanma olasılığı aşağıdakilerden hangisidir?

a. 0.0712
b. 0.3296
c. 0.2304
d. 0.1607
e. 0.3456

5. Bir otobüs firması, Ankara-İzmir arasında düzenlediği seferlerin %60'ında tam doluluk oranını yakalamaktadır. Ankara-İzmir hattında çalışan otobüslerden haftanın her günü rassal olarak 1 tanesi seçiliyor. Buna göre, seçilen 7 otobüsten, 4 tanesinin tam dolu olma olasılığı aşağıdakilerden hangisidir?

a. 0.1935
b. 0.4502
c. 0.2215
d. 0.2903
e. 0.4188

6. Bir firmanın, bir günde ürettiği bardakların ortalama 2 tanesi kusurludur. Bu firmanın, haftanın belli 2 günü (örneğin Pazartesi ve Çarşamba) ürettiği bardaklardan 3 tanesinin kusurlu olma olasılığı aşağıdakilerden hangisidir?

a. 0.3286
b. 0.1954
c. 0.8046
d. 0.2945
e. 0.1011

7. Bir hastanenin acil servisine, bir haftada trafik kazası vakası gelme olasılığı 0.007'dir. Acil servise bir haftada gelen 1000 vakadan en az 3 tanesinin trafik kazası vakası olma olasılığı aşağıdakilerden hangisidir?

a. 0.8162
b. 0.0296
c. 0.9704
d. 0.7667
e. 0.5549

8. X rassal değişkeninin olasılık dağılımı aşağıda verildiği gibidir:

X	1	2	3	4
$P(X=x)$	0.1	0.2	0.3	0.4

Buna göre, X rassal değişkeninin beklenen değeri (μ) ve varyansı (σ^2) aşağıdakilerden hangisidir?

a. $\mu= 1$, $\sigma^2=3$
b. $\mu=3$, $\sigma^2=10$
c. $\mu=3$, $\sigma^2=1$
d. $\mu=3$, $\sigma^2=9$
e. $\mu=1$, $\sigma^2=10$

9. Temmuz ayında herhangi bir günde Bolu'ya yağmur yağma olasılığı 0.10'dur. Temmuz ayı içerisinde rassal olarak seçilen 10 günden 3 tanesinde Bolu'ya yağmur yağma olasılığı aşağıdakilerden hangisidir?

a. 0.8993
b. 0.0574
c. 0.1642
d. 0.0315
e. 0.9426

10. *X* kesikli rassal değişkeni *Bernoulli(p=0.3)* dağılımına sahip ise ortalama μ ve varyansı σ^2 aşağıdakilerden hangisidir?

a. $\mu=0.3$, $\sigma^2=0.2100$
b. $\mu=0.7$, $\sigma^2=0.2100$
c. $\mu=0.3$, $\sigma^2=0.4583$
d. $\mu=0.21$, $\sigma^2=0.2100$
e. $\mu=0.7$, $\sigma^2=0.4583$

Kendimizi Sınayalım Yanıt Anahtarı

1. c	Yanıtınız yanlış ise "Rassal Değişkenlerin Çeşitleri" konusunu yeniden gözden geçiriniz.
2. b	Yanıtınız yanlış ise "Olasılık Dağılımı" konusunu yeniden gözden geçiriniz.
3. a	Yanıtınız yanlış ise "Olasılık Dağılımı" konusunu yeniden gözden geçiriniz.
4. e	Yanıtınız yanlış ise "Binom Dağılımı" konusunu yeniden gözden geçiriniz.
5. d	Yanıtınız yanlış ise "Binom Dağılımı" konusunu yeniden gözden geçiriniz.
6. b	Yanıtınız yanlış ise "Poisson Dağılımı" konusunu yeniden gözden geçiriniz.
7. c	Yanıtınız yanlış ise "Binom Dağılımının Poisson Dağılımına Yakınsaması" konusunu yeniden gözden geçiriniz.
8. c	Yanıtınız yanlış ise "Kesikli Rassal Değişkenlerin Ortalama, Varyans ve Standart Sapması" konusunu yeniden gözden geçiriniz.
9. b	Yanıtınız yanlış ise "Binom Dağılımı" konusunu yeniden gözden geçiriniz.
10. a	Yanıtınız yanlış ise "Kesikli Rassal Değişkenlerin Ortalama, Varyans ve Standart Sapması" konusunu yeniden gözden geçiriniz.

Sıra Sizde Yanıt Anahtarı

Sıra Sizde 1

1. (a) ve (d) kesikli, (b) ve (c) süreklidir.

2. Ek hesabı olan bir banka müşterisinin hesabındaki para miktarı hem pozitif hem de negatif değerler alabilen bir rassal değişkendir.

Sıra Sizde 2

1. a)

$$0 \le P\left(X=x\right) \le 1, \; her \;\; x \;\; için \;\; ve$$

$$\sum_{x=0}^{4} P(X=x) = P(X=0)+P(X=1)+\ldots+P(X=4)=1$$

olduğundan, *P(X=x)* olasılık dağılımıdır.

b)

$$\begin{aligned} P(X>0) &= P(X=1)+P(X=2)+P(X=3)+P(X=4) \\ &= 2/10+3/10+2/10+3/20 = 17/20 \;\; ve \end{aligned}$$

$$1-P(X=0) = 1-3/20 = 17/20$$

olduğundan *P(X>0)=1-P(X=0)* 'dir.

2. a) *P(X=x)* in olasılık dağılımı olabilmesi için

$$\begin{aligned} \sum_{x=1}^{4} P(X=x) &= P(X=1)+P(X=2)+P(X=3)+P(X=4)=1 \\ &= 0.12+0.16+0.54+c=1 \end{aligned}$$

olmalıdır. Buradan, *c*= 0.18 olarak bulunur.

b)

$$F_X(x) = \begin{cases} 0, & x<1 \\ 0.12, & 1 \le x < 2 \\ 0.28, & 2 \le x < 3 \\ 0.82, & 3 \le x < 4 \\ 1, & x \ge 4 \end{cases}$$

Sıra Sizde 3

1. "Y: Hastalığın kalıcı hasar bıraktığı bebek sayısı" olarak tanımlanırsa, *Y ~ Binom(4,0.10)* olduğu görülür. Bu durumda,

a) *P(*4 bebekten 2 tanesinde kalıcı hasar bırakması*)* *=P(X=2)*

$$= \begin{pmatrix} 4 \\ 2 \end{pmatrix} (0.1)^2 (0.9)^2 = 0.0486$$

bulunur.

b) *P(*4 bebekten en az 1 en fazla 3'ünde kalıcı hasar bırakması*)*= *P(1≤X≤3)*

$$= P\left(X=1\right)+P\left(X=2\right)+P\left(X=3\right)$$

$$= \begin{pmatrix} 4 \\ 1 \end{pmatrix} (0.1)^1 (0.9)^3 + \begin{pmatrix} 4 \\ 2 \end{pmatrix} (0.1)^2 (0.9)^2 + \begin{pmatrix} 4 \\ 3 \end{pmatrix} (0.1)^3 (0.9)^1 = 0.3438$$

bulunur.

2. "*Y:* Doğru cevaplanan soru sayısı" olarak tanımlanırsa *Y ~ Binom (5,0.25)* olduğu görülür. Bu durumda, *P(*5 sorudan 3'ünü doğru yapma)= P(X=3)=

$$\begin{pmatrix} 5 \\ 3 \end{pmatrix} (0.25)^3 (0.75)^2 = 0.0879$$

bulunur.

Sıra Sizde 4

1. 5 günde ortalama 4 tane kusurlu lastik üretiliyorsa bir günde ortalama $\lambda = \frac{4}{5} = 0.8$ tane kusurlu lastik üretilir. "*X:* Belli bir günde üretilen ortalama kusurlu lastik sayısı" olarak tanımlanırsa, *X~ Poisson (λ*= 0.8) olduğu görülür. Buradan,

*P(*belli bir günde üretilen lastiklerin en az 3 tanesinin kusurlu olması)= *P (X≥3)*

$$\begin{aligned} &= 1-\left\{P(X=0)+P(X=1)+P(X=2)\right\} \\ &= 1-\left\{\frac{e^{-0.8}(0.8)^0}{0!}+\frac{e^{-0.8}(0.8)^1}{1!}+\frac{e^{-0.8}(0.8)^2}{2!}\right\} \\ &= 0.0474 \end{aligned}$$

bulunur.

2. "*X:* Belli bir günde iade edilen ortalama oyuncak sayısı" olarak tanımlanırsa, *X ~ Poisson (λ=5)* olduğu görülür. Bu durumda,

*P(*belli bir günde en az 2 en fazla 4 oyuncağın iade edilmesi)= *P(2≤X≤4)*

$$\begin{aligned} &= P(X=2)+P(X=3)+P(X=4) \\ &= \frac{e^{-5}5^2}{2!}+\frac{e^{-5}5^3}{3!}+\frac{e^{-5}5^4}{4!} \\ &= 0.4001 \end{aligned}$$

bulunur.

Sıra Sizde 5

1. "*Y*: Kusurlu cihaz sayısı" olarak tanımlanırsa *Y ~ Binom (400,0.01)* olduğu görülür. $\lambda=np=400(0.01)=4$ olmak üzere, aranan olasılık değeri,

P(400 tane alarm cihazından en fazla 2 tanesinin kusurlu olması)= $P(Y\leq 2)$

$$= P(Y=0)+P(Y=1)+P(Y=2)$$
$$= \frac{e^{-4}4^0}{0!}+\frac{e^{-4}4^1}{1!}+\frac{e^{-4}4^2}{2!}$$
$$= 0.2381$$

şeklinde bulunur.

2. "*Y*: Bir yılda elektrik kesintisi yaşanan gün sayısı" olarak tanımlanırsa, *Y ~ Binom (365,0.005)* olduğu görülür. $\lambda=np= 365.(0.005)=1.825$ olmak üzere, aranan olasılık değeri,

P(1 yılda en az 1 en fazla 3 gün elektrik kesintisi yaşanması)= $P(1\leq X\leq 3)$

$$= P(X=1)+P(X=2)+P(X=3)$$
$$= \frac{e^{-1.825}(1.825)^1}{1!}+\frac{e^{-1.825}(1.825)^2}{2!}+\frac{e^{-1.825}(1.825)^3}{3!}$$
$$= 0.7260$$

şeklinde bulunur.

Sıra Sizde 6

1. $P(X=x)$ olasılık dağılımı olduğundan,

$$\sum_{x=0}^{3} P(X=x) = P(X=0)+P(X=1)+P(X=2)+$$
$$P(X=3)=1'dir.\ \text{Buradan, } c=0.43 \text{ bulunur.}$$

$$\mu = E(X) = \sum_{x=0}^{3} x.P(X=x) = 0.(0.22)+1.(0.43)+$$
$$2.(0.29)+3.(0.06)=1.19$$

$$E(X^2) = \sum_{x=0}^{3} x^2.P(X=x) = 0^2.(0.22)+$$
$$1^2.(0.43)+2^2.(0.29)+3^2.(0.06)=2.13$$

$$\sigma^2 = V(X) = E(X^2)-\left(E(X)\right)^2 = 2.13-(1.19)^2 = 0.71 \ ve$$
$$\sigma = \sqrt{V(X)} = \sqrt{0.71} = 0.84 \text{ bulunur.}$$

2.

a) $\mu = E(X) = np = 50.(0.40) = 20;$

$$\sigma^2 = V(X) = np(1-p) = 50.(0.4).(0.6) = 12 \ ve$$
$$\sigma = \sqrt{V(X)} = \sqrt{np(1-p)} = \sqrt{12} = 3.46$$

b) $\mu = E(X) = \lambda = 4;\ \sigma^2 = V(X) = \lambda = 4 \ ve$

$$\sigma = \sqrt{V(X)} = \sqrt{\lambda} = \sqrt{4} = 2$$

Yararlanılan ve Başvurulabilecek Kaynaklar

Bradley T. (2007). **Essential Statistics for Economics, Business and Management.** *John Wiley & Sons, Ltd., England.*

Newbold P, Carlson W. L. ve Thorne B. (2010). **Statistics for Business and Economics.** *Pearson: Boston.*

Poisson S.D. (1837). **Probabilite des jugements en matiere criminelle et en matiere civile, precedees des regles generales du calcul des probabilities** *(Paris, France: Bachelier), page 206.*

6

Amaçlarımız

Bu üniteyi tamamladıktan sonra;

- Sürekli rassal değişken kavramını açıklayabileceksiniz,
- Düzgün dağılımı genel hatlarıyla inceleyecek, günlük yaşamda düzgün dağılımın yerini belirleyebilecek ve düzgün dağılıma ilişkin olasılıkları hesaplayabileceksiniz,
- Normal dağılımı ana hatlarıyla öğrenecek, normal dağılımın gündelik hayattaki yerini belirleyebilecek ve normal dağılımla ilgili olasılıkları hesaplayabileceksiniz,
- Binom dağılımıyla normal dağılım arasındaki bağlantıyı ilişkilendirebileceksiniz.

Anahtar Kavramlar

- Sürekli Rassal Değişken
- Düzgün Dağılım
- Normal Dağılım
- Standart Normal Dağılım
- Binom Dağılımına Normal Dağılım Yaklaşımı

İçindekiler

İstatistik-I

Sürekli Rassal Değişkenler ve Olasılık Dağılımları

- GİRİŞ
- SÜREKLİ RASSAL DEĞİŞKENLER
- DÜZGÜN (UNİFORM) DAĞILIM
- NORMAL DAĞILIM
- BİNOM DAĞILIMINA NORMAL DAĞILIM YAKLAŞIMI

Sürekli Rassal Değişkenler ve Olasılık Dağılımları

GİRİŞ

Bundan önceki bölümde kesikli rassal değişkenler ile binom ve Poisson dağılımları anlatılmıştı. Hatırlanacağı gibi, kesikli rassal değişkenler sonlu (örneğin; 0, 1, 2,...,10) veya sayılabilir sonsuzlukta (örneğin; 0, 1, 2,...) değerler alabilmektedir. Fakat birçok uygulamada, rassal değişkenin aldığı değerler sayılamayacak sonsuzlukta olabilmektedir.

Örneğin, bir işletmede otomatik makine yardımıyla 1.5 metre uzunluğunda bir ürün üretiliyor olsun. Çeşitli nedenlerden dolayı üretilen ürünlerin uzunluklarının birbirinden farklı olduğu bilinsin. Bu durumda, ürünlerin uzunluğu bir rassal değişkendir ve bu rassal değişkenin alabileceği değerler ise 1.40, 1.41, 1.43, ... , 1.57, 1.58 metre şeklinde sayılamayacak (sonsuz) kadar çoktur. Sonuç olarak, bu örnekteki X rassal değişkeni, 1.40 ile 1.58 arasındaki her bir değeri alabileceği için sürekli bir rassal değişkendir. Bu rassal değişkeninin alabileceği değerlerinin bulunduğu 1.40-1.58 aralığı ise tanım aralığı olarak ifade edilir.

Bu bölümde ilk olarak günlük yaşamda geniş bir uygulama alanı olan sürekli rassal değişkenler ele alınacaktır. Daha sonra, sürekli rassal değişkenlerin uygulandığı pek çok olasılık dağılımı arasında en basiti olan düzgün (uniform) dağılım üzerinde durulacaktır. Ardından, gerek yaygın kullanımı gerekse istatistiksel çıkarsamalarda temel dağılım olan normal dağılım incelenecektir. Son olarak, binom dağılımında büyük n değerleri için olasılık hesaplamalarında bize kolaylıklar sağlayan binom dağılımına normal dağılım yaklaşımı ele alınacaktır.

SÜREKLİ RASSAL DEĞİŞKENLER

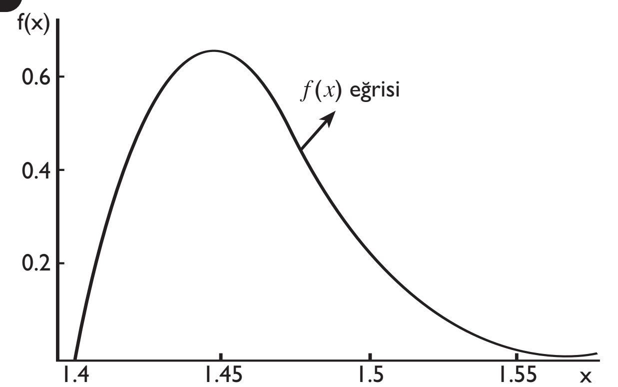
Sürekli rassal değişkenin tanımı üzerinde tekrar durulacak olursa belli bir aralıkta veya aralıklarda her değeri alabilen rassal değişkene, **sürekli rassal değişken** denir. Bir başka ifadeyle, sürekli rassal değişken, alabileceği değerleri sayılamayacak (sonsuz) kadar çok olan rassal değişkendir. Ayrıca, sürekli rassal değişkenin değerleri genellikle, sayım yoluyla elde edilen kesikli rassal değişkenlerin aksine ölçüm yoluyla elde edilmektedir.

Sürekli rassal değişkenlerle ilgili verilebilecek örneklerden bazıları şu şekildedir; bir hisse senedinin fiyatı, okul giderleri için yapılan harcama, faiz oranı, bir arabanın yakıt tüketimi, rüzgâr hızı, bir elektronik eşyanın dayanma süresi, yağış miktarı, bir ürünün tamamlanma süresi, ana okulundaki çocukların ağırlığı ve üretilen demir çubuklarının uzunluğu gibi değişkenler sürekli rassal değişkenlerdir.

Sürekli rassal değişkenlerle ilgili olasılık hesabı, kesikli rassal değişkenlerden farklılık göstermektedir. Bu durum, giriş bölümünde verilen ürünlerin uzunluklarıyla ilgili örnek üzerinde anlatılsın. Buna göre işletmede çalışan yöneticiler, elde edilen bu ürünler arasından rassal olarak seçilen bir ürünün uzunluğunun 1.45 ile 1.50 metre arasında olması olasılığını bulmak istesinler. Bu durumda, X rassal değişkeninin 1.45 ile 1.50 metre arasında alabileceği değerler de sayılamayacak çokluktadır. Bundan dolayı, sürekli rassal değişkenlerle ilgili olasılıklar, kesikli rassal değişkenlerde olduğu gibi tek tek hesaplanamaz. Bu durumda sürekli rassal değişkenlerle ilgili olasılıkları belirlemek için **alan** kavramı kullanılır.

Şekil 6.1

Ürünlerin uzunluklarına uygun $f(x)$ eğrisi.



Sürekli rassal değişkenlerle ilgili olasılıkların (alanların) hesaplanabilmesi ve çeşitli yorumların yapılabilmesi için **olasılık yoğunluk fonksiyonu** veya **olasılık dağılım eğrisi** kullanılır. Kullanılan bu olasılık yoğunluk fonksiyonu $f(x)$ şeklinde gösterilir. Ele alınan her bir sürekli rassal değişken için, söz konusu olasılık dağılım eğrisinin şekli değişmektedir. Örneğin, bir hisse senedinin fiyatı için başka, bir ürünün tamamlanma süresi için daha başka şekilde dağılım eğrileri elde edilmektedir. Ürünlerin uzunlukları için elde edilen $f(x)$ olasılık yoğunluk fonksiyonunun grafiği veya $f(x)$ eğrisi Şekil 6.1'de verilmiştir.

$f(x)$ olasılık yoğunluk fonksiyonu olmak üzere, her x için $f(x) \geq 0$ dır. $f(x)$ eğrisi altında kalan ve x-ekseniyle sınırlandırılmış alan (olasılık) 1'e eşittir.

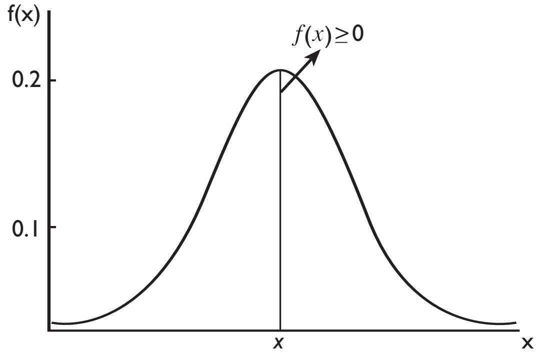
Sürekli bir rassal değişkenin $f(x)$ olasılık yoğunluk fonksiyonu aşağıdaki özellikleri sağlaması gerekir.

i. Her x için $f(x) \geq 0$'dır.

ii. $f(x)$ eğrisi altında kalan ve x-ekseniyle sınırlandırılmış alan veya olasılık 1'e eşittir.

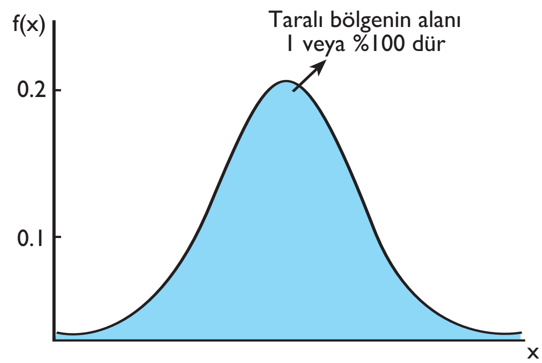
Birinci özellik, $-\infty < x < \infty$ aralığındaki her x değeri için, $f(x)$ olasılık yoğunluk fonksiyonunun aldığı değerlerin pozitif veya sıfıra eşit olmasıdır. Bir başka deyişle, $f(x)$ fonksiyonun grafiği daima x ekseninin üst kısmındadır (Şekil 6.2). İkinci özellikse, $f(x)$ olasılık yoğunluk fonksiyonunun grafiği altında kalan ve x-ekseniyle sınırlandırılmış alanın veya olasılığın her zaman 1 veya %100 olmasıdır (Şekil 6.3).

Şekil 6.2



Bir olasılık yoğunluk fonksiyonunun pozitif veya sıfır olması.

Şekil 6.3



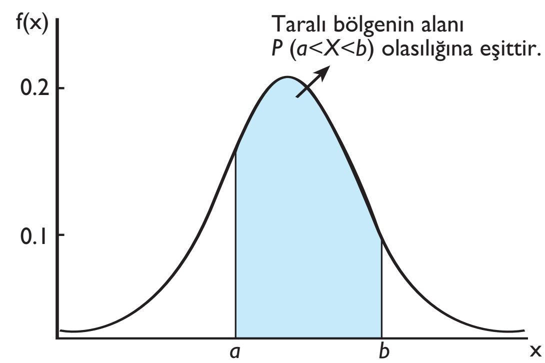
Bir olasılık yoğunluk fonksiyonunun grafiği altında kalan toplam alanın 1 olması.

Sürekli X rassal değişkeninin herhangi a ve b $(a \leq b)$ değerleri arasında bulunması **olasılığı,** $P(a < X < b)$ biçiminde gösterilir. Bu olasılık değeri ise $f(x)$ eğrisi altında a ve b değerleri arasında kalan bölgenin alanına eşittir. Bir başka ifadeyle, $P(a < X < b)$ olasılığı, Şekil 6.4'te gösterilen taralı bölgenin alanına eşittir. $f(x)$ eğrisi altındaki söz konusu taralı bölgenin alanını bulmak için, belirli integral $\int_a^b f(x)dx$ işleminden yararlanılması gerekir. Buna göre, X sürekli rassal değişkeninin a ve b $(a \leq b)$ değerleri arasında olması olasılığı,

$$\text{Taralı Bölgenin Alanı} = P(a < X < b) = \int_a^b f(x)dx$$

şeklinde bulunur. Ancak bu kitapta ele alınacak olan sürekli rassal değişkenlerle ilgili olasılık hesaplamaları, integral işlemleri yerine, basit alan hesaplamaları veya hazır tablo yardımıyla yapılacaktır.

Şekil 6.4



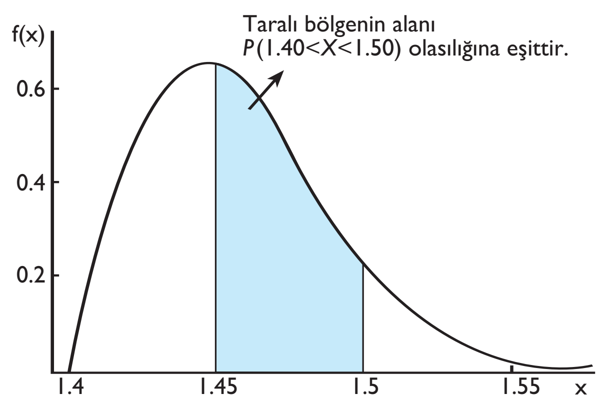
Olasılığın, $f(x)$ eğrisi altındaki alan olarak ifadesi.

DİKKAT

Sürekli X rassal değişkeninin herhangi a ve b $(a \leq b)$ değerleri arasında olması olasılığı, $P(a < X < b)$, her zaman 0 ile 1 arasındadır. Bir başka ifade ile $0 \leq P(a < X < b) \leq 1$'dir.

Bu bilgiler ışığında ürünlerin uzunluklarıyla ilgili X rassal değişkeninin 1.45 ve 1.50 değerleri arasında olması olasılığı, uzunlukların $f(x)$ eğrisi altında 1.45 ve 1.50 değerleri arasında kalan bölgenin alanına eşittir. Bu durum Şekil 6.5'te görselleştirilmiştir.

Şekil 6.5



X rassal değişkeninin 1.40 ve 1.50 değerleri arasında olması olasılığı.

Sürekli X rassal değişkeninin tek bir x değerini ($x = a$ gibi) alması olasılığı her zaman sıfırdır (Şekil 6.6). Bunun nedeni, verilen bir noktanın alanının sıfır olmasıdır. Buna göre, sürekli X rassal değişkenin sadece a veya sadece b değerini alması olasılığı,

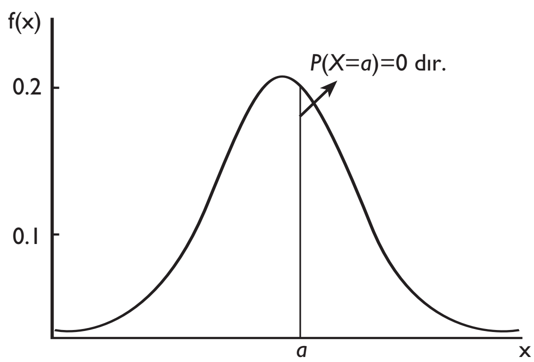
$$P(X = a) = 0 \text{ veya } P(X = b) = 0$$

olduğundan, X rassal değişkeni için

$$P(a < X < b) = P(a \le X \le b) = P(a < X \le b) = P(a \le X < b)$$

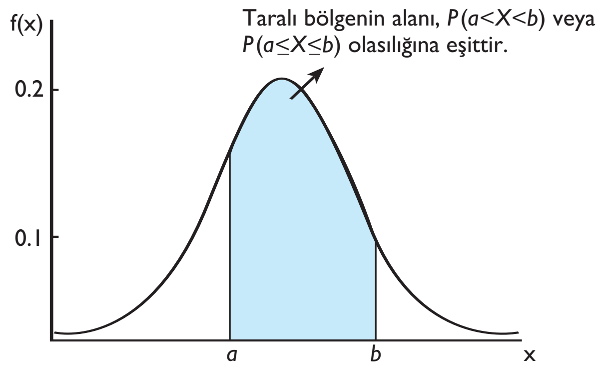
eşitliği yazılabilir (Şekil 6.7). Bu nedenle, sürekli X rassal değişkeninin a ve b arasında değer alması olasılığı hesaplanırken, a ve b'nin bu aralığa dahil edilmesi hesaplama açısından önemli değildir.

Şekil 6.6



Sürekli rassal değişkenin tek bir değer alması olasılığı.

Şekil 6.7



Sürekli bir rassal değişken için $P(a < X < b)$ veya $P(a \le X \le b)$ olasılığı.

SIRA SİZDE 1

1. Sürekli rassal değişkeni tanımlayınız ve aşağıda verilen rassal değişkenlerden hangilerinin sürekli rassal değişken olduğunu belirleyiniz.

a. Bir firmanın herhangi bir ayda sattığı cep telefonu sayısı

b. Bir ailenin geliri

c. Bir günde üretilen kusurlu ürün sayısı

d. Yeni doğan bebeklerin boy uzunluğu

2. Sürekli bir rassal değişkenin $f(x)$ olasılık yoğunluk fonksiyonunun sağlaması gereken iki özelliğini kısaca yazınız.

3. Aşağıda $a < b$ olmak üzere, sürekli X rassal değişkeniyle ilgili verilenlerden hangilerinin doğru olduğunu belirleyiniz.

a) $P(a < X < b) = P(a \le X \le b)$

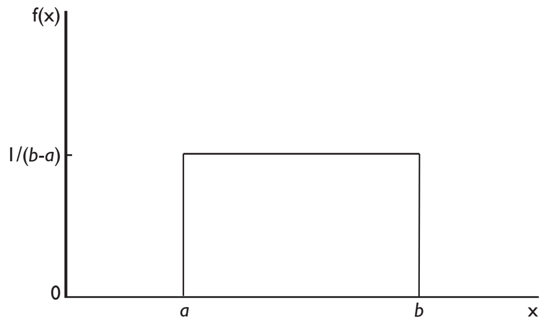
b) $P(X = a) = 0.5$

c) $P(a \le X \le b) = 1.05$

DÜZGÜN (UNIFORM) DAĞILIM

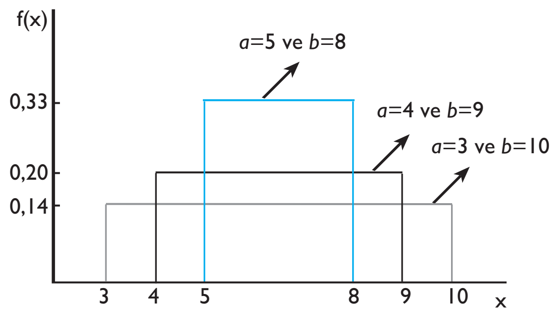
Sürekli rassal değişkenler için en basit dağılımlardan biri, düzgün (uniform) dağılımdır. Düzgün dağılım, sürekli bir rassal değişkenin tanımlı olduğu aralıkta belirlenen eşit uzunluktaki aralıkların olasılıklarının eşit olduğu bir dağılımdır. Örneğin; bir uçağın bir yerden başka bir yere olan uçuş süresi, belli uzunluktaki bir borunun arızalandığı noktadaki mesafesi gibi rassal değişkenler yaklaşık olarak düzgün dağılıma sahiptir.

Şekil 6.8



Düzgün dağılımın olasılık yoğunluk fonksiyonunun şekli.

Şekil 6.9



Düzgün dağılımda farklı a ve b değerleri için f (x) eğrileri.

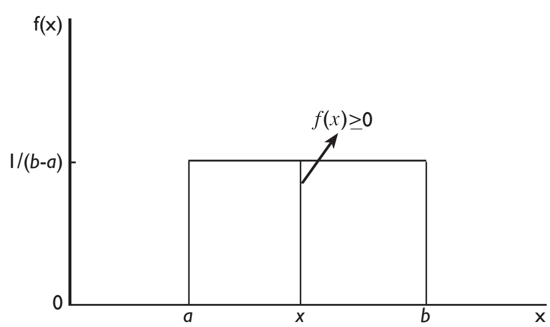
Sürekli X rassal değişkeninin olasılık yoğunluk fonksiyonu,

$$f(x) = \frac{1}{b-a}, \quad a \leq x \leq b$$

şeklinde ise X rassal değişkenine **düzgün dağılıma** sahiptir denir ve kısaca $X \sim U(a, b)$ biçiminde gösterilir. Burada "~" sembolü dağılmıştır olarak okunur. Şekil 6.8'de görüldüğü gibi, $f(x)$ olasılık yoğunluk fonksiyonunun grafiği veya düzgün dağılım eğrisi belli bir $a \leq x \leq b$ aralığında düz bir doğru biçimindedir. Bu $a \leq x \leq b$ aralığının dışındaki x değerleri için $f(x)$ fonksiyonu sıfır değerini alır. Burada a değeri, rassal değişkenin alabileceği değerlerin en küçüğünü (minimumunu), b değeri ise en büyüğünü (maksimumunu) göstermektedir. Her farklı a ve b değeri için, farklı $f(x)$ olasılık yoğunluk fonksiyonu elde edilmektedir. Üç farklı a ve b değeri için elde edilen $f(x)$ fonksiyonlarının grafikleri Şekil 6.9'da verilmiştir.

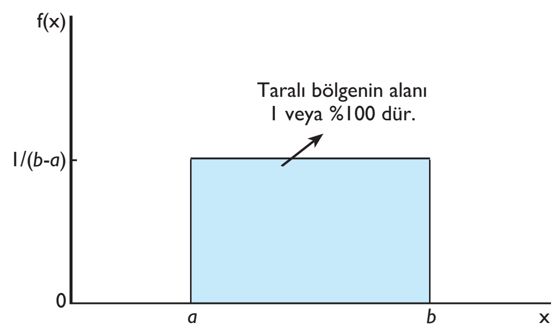
Düzgün dağılımın olasılık yoğunluk fonksiyonunun, bir önceki bölümde belirttiğimiz iki özelliği de sağladığı Şekil 6.10 ve Şekil 6.11'de gösterilmiştir.

Şekil 6.10



Düzgün dağılımın olasılık yoğunluk fonksiyonunun pozitif olması.

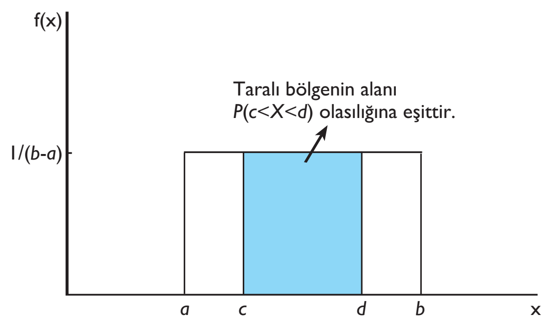
Şekil 6.11



Düzgün dağılım eğrisi altında kalan toplam alanın 1 olması.

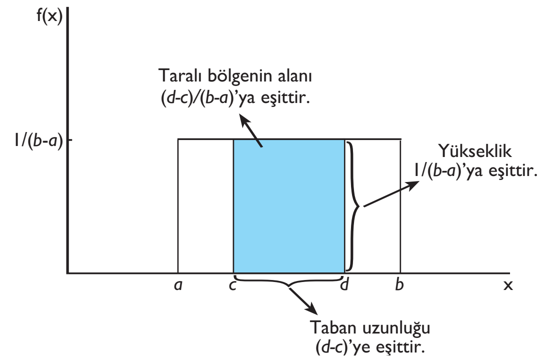
Sürekli X rassal değişkeni $a \leq x \leq b$ aralığında düzgün dağılıma sahip olduğunda, bu rassal değişkenin c ve d ($a \leq c < d \leq b$) değerleri arasında olması olasılığı $P(c < X < d)$ biçiminde gösterilir. Bu olasılık değeri ise düzgün dağılım eğrisi altında c ve d değerleri arasında kalan bölgenin alanına eşittir (Şekil 6.12).

Şekil 6.12



Düzgün dağılım eğrisi altında x = c ile x = d noktaları arasındaki alan.

Şekil 6.13



Düzgün dağılım eğrisi altında kalan alanın değeri.

Şekil 6.12'de belirtilen taralı bölgenin alanı veya $P(c < X < d)$ olasılığı,

Dikdörtgen alanı= (Yükseklik) (Taban uzunluğu)

formülü ile bulunur. Diğer bir ifadeyle, Şekil 6.13'ten görüldüğü gibi, taralı bölgenin (dikdörtgenin) yüksekliği $1 / (b - a)$'ya, taban uzunluğu $(d - c)$'ye eşittir. Bu bilgilere göre, taralı bölgenin alanı veya $P(c < X < d)$ olasılığı,

Taralı Bölgenin Alanı = $P(c < X < d) = (d - c) / (b - a)$

formülüyle hesaplanır. Düzgün dağılıma sahip X rassal değişkeniyle ilgili olasılıkların bulunması aşağıdaki örneklerle verilmiştir.

DİKKAT

$a \leq x \leq b$ aralığında düzgün dağılıma sahip sürekli X rassal değişkenin, $a \leq c < d \leq b$ olmak üzere c ve d değerleri arasında olması olasılığı, $P(c < X < d) = (d - c)/(b - a)$ formülüyle hesaplanır.

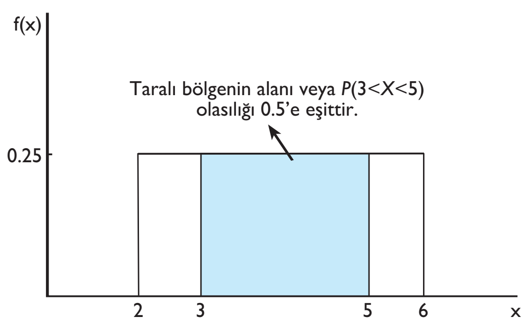
Örnek 1: *Sürekli X rassal değişkeni, 2 ile 6 değerleri arasında düzgün dağılıma sahip olsun. X rassal değişkeninin 3 ile 5 arasında değer alması olasılığını bulunuz.*

Çözüm: Verilen bilgilere göre, düzgün dağılıma sahip sürekli rassal değişkenin olasılık yoğunluk fonksiyonu:

$$f(x) = \frac{1}{b-a} = \frac{1}{6-2} = \frac{1}{4}, \; 2 \leq x \leq 6$$

Şekil 6.14

Düzgün dağılım eğrisi altında; x = 3 ve x = 5 arasında kalan alan.



olarak elde edilir. Şekil 6.14'ten görüleceği gibi, X rassal değişkeninin 3 ve 5 arasında değer alması olasılığı, taralı bölgenin alanına eşittir ve

$P(c < X < d) = (d - c) / (b - a)$

formülüyle bulunur.

Buna göre, aranan olasılık değeri,

$$P\,(3 < X < 5) = \frac{d-c}{b-a} = \frac{5-3}{6-2} = \frac{2}{4} = 0.5$$

elde edilir. Ayrıca, bir önceki bölümde ifade edildiği gibi, sürekli bir rassal değişkenin tek bir değer alması olasılığı sıfırdır. Bu nedenle söz konusu örnekte,

$$P\,(3 < X < 5) = P\,(3 \leq X \leq 5) = P\,(3 \leq X < 5) = P\,(3 < X \leq 5) = 0.5$$

eşitliği de geçerlidir.

Örnek 2: *Bir firma 5.5 metre uzunluğunda elektrik kablosu üretmektedir. Bu üretilen kabloların 0 ile 5.5 metre arasındaki herhangi bir yerinde arıza çıkmasının düzgün dağılıma sahip olduğu bilinmektedir. Bu üretilen kablolardan rassal olarak seçilen bir kabloda;*

a. 3.25 ile 4.05 metre arasında arıza çıkması,

b. 2.15 metreden önce arıza çıkması,

olasılıklarını bulunuz.

Çözüm: Bu örnekte sürekli X rassal değişkeni kablonun arızalandığı yerin uzunluğunu göstermekte ve 0 ile 5.5 değerleri arasında düzgün dağılmaktadır. Kısaca $X \sim U(0,5.5)$ olarak ifade edilir. X rassal değişkenin olasılık yoğunluk fonksiyonu da,

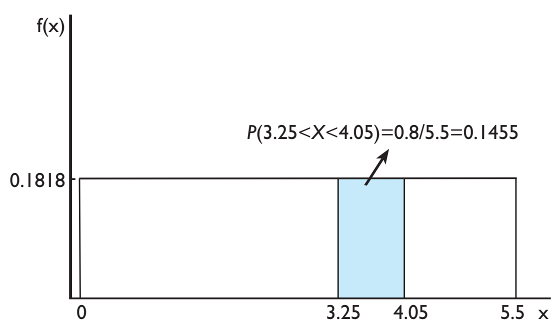
$$f(x) = \frac{1}{b-a} = \frac{1}{5.5-0} = \frac{1}{5.5},$$

$$0 \leq x \leq 5.5$$

eşitliğiyle ifade edilir.

a. Şekil 6.15'ten görüleceği gibi, X rassal değişkeninin $x = 3.25$ ve $x = 4.05$ değerleri arasında olması olasılığı, düzgün dağılım eğrisi altında verilen sınır değerleri arasındaki alana eşittir. Bu durumda, rassal olarak seçilen bir kabloda 3.25 ile 4.05 metre arasında arıza çıkması olasılığı,

Şekil 6.15



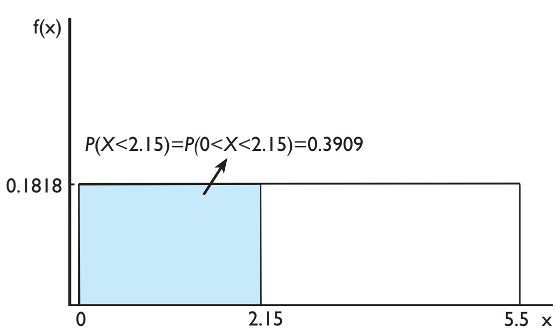
f (x) eğrisi altında x = 3.25 ve x = 4.05 değerleri arasında kalan alan.

$$P\,(3.25 < X < 4.05) = \frac{d-c}{b-a} = \frac{4.05-3.25}{5.5-0} = \frac{0.8}{5.5} = 0.1455$$

olarak elde edilir.

b. Burada X rassal değişkeninin $x = 2.15$'ten önce (solunda) değerler alması olasılığı, bir başka deyişle $P\,(X < 2.15)$ olasılığının değeri sorulmaktadır (Şekil 6.16). Daha önce belirtildiği gibi, X rassal değişkenin aldığı değerler $0 \leq x \leq 5.5$ arasındadır. Bu nedenle aranan $P\,(X < 2.15)$ olasılığı, $P\,(0 < X < 2.15)$ olasılığına eşit olur. Bu olasılık,

Şekil 6.16



f (x) eğrisi altında x = 2.15 değerinin solunda kalan alan.

$$P(0 < X < 2.15) = \frac{2.15 - 0}{5.5 - 0} = \frac{2.15}{5.5} = 0.3909$$

olarak bulunur. Sonuç olarak, X rassal değişkeninin $x = 2.15$'in solundaki değerleri alması olasılığı,

$$P(X < 2.15) = P(0 < X < 2.15) = 0.3909$$

biçiminde de gösterilir.

DİKKAT

$a \leq x \leq b$ aralığında düzgün dağılıma sahip sürekli X rassal değişkeni için $a \leq c \leq b$ olmak üzere $P(X < c)$ olasılığı, $P(X < c) = P(a < X < c) = (c - a) / (b - a)$ şeklinde bulunur.

DİKKAT

$a \leq x \leq b$ aralığında düzgün dağılıma sahip sürekli X rassal değişkeni için $a \leq d \leq b$ olmak üzere $P(X > d)$ olasılığı, $P(X > d) = P(d < X < b) = (b - d) / (b - a)$ şeklinde bulunur.

Düzgün Dağılımın Ortalaması ve Standart Sapması

Burada $a \leq x \leq b$ aralığında düzgün dağılıma sahip sürekli X rassal değişkeninin ortalamasının (μ) ve standart sapmasının (σ) bulunabilmesi için kullanılan formüller aşağıda verilmiştir.

$$\mu = \frac{a+b}{2} \quad \text{ve} \quad \sigma = \frac{b-a}{\sqrt{12}}$$

Formülerde, a değeri, rassal değişkenin alabileceği değerlerin en küçüğünü (minimumunu), b değeri ise en büyüğünü (maksimumunu) göstermektedir.

Örnek 3: *Sürekli X rassal değişkeni, $1 \leq x \leq 7$ aralığında düzgün dağılıma sahip olsun. X rassal değişkeninin ortalama ve standart sapma değerlerini bulunuz.*

Çözüm: Verilen bilgilere göre, X rassal değişkeninin tanım aralığı $1 \leq x \leq 7$ biçimindedir. Bu nedenle, X'in alabileceği minimum değer $a = 1$ ve maksimum değer ise $b = 7$'dir. Bu değerler yukarıda belirtilen formüllerde yerine konulduğunda, X rassal değişkeninin ortalaması ve standart sapması sırasıyla

$$\mu = \frac{a+b}{2} = \frac{1+7}{2} = 4\ , \quad \sigma = \frac{b-a}{\sqrt{12}} = \frac{7-1}{\sqrt{12}} = 1.7321$$

olarak elde edilir.

SIRA SİZDE 2

1. Sürekli X rassal değişkeni, $4 \leq x \leq 9$ aralığında düzgün dağılmaktadır. Bu bilgilere göre,
a. $P(4 \leq X \leq 6)$
b. $P(6 \leq X \leq 9)$
olasılıklarını bulunuz.
2. Bir üretim bandında belli bir parçanın montaj süresi 30 ile 37 saniye arasında düzgün dağılmaktadır. Üretim bandından rassal olarak seçilen bir parçanın montaj süresinin,
a. 32 -36 saniye arasında olması,
b. 34 saniyeden fazla olması,
olasılıklarını bulunuz.
3. Düzgün dağılıma sahip X rassal değişkeninin ortalamasının 42 ve tanım aralığının $28 \leq x \leq b$ olduğu bilindiğine göre,
a. b değerini bulunuz.
b. X rassal değişkeninin standart sapmasını hesaplayınız.

NORMAL DAĞILIM

Sürekli rassal değişkenler için en önemli dağılımlardan biri, normal dağılımdır. Bunun nedeni, gerek günlük yaşamımızda gözlenen sürekli rassal değişkenlerin büyük çoğunluğunun (yaklaşık olarak) normal dağılıma uyması, gerekse istatistiksel çıkarsamalarda temel dağılım olarak normal dağılımın kullanılmasıdır. Örneğin, bir yatırım aracının aylık getirileri, bir şirketin haftalık satışı, üretilen ürünlerin ağırlıkları, bir deneyde yapılan rassal ölçüm hataları, zeka testi sonuçları, yeni doğan bebeklerin ağırlıkları ve boy uzunlukları gibi rassal değişkenler yaklaşık olarak normal dağılırlar.

Sürekli *X* rassal değişkeninin olasılık yoğunluk fonksiyonu,

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}, \quad -\infty < x < \infty$$

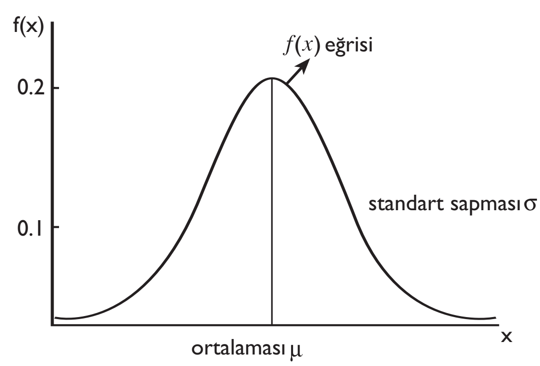
şeklinde ise *X* rassal değişkenine **normal dağılıma** sahiptir denir ve kısaca $X \sim N(\mu,\sigma^2)$biçiminde gösterilir. Burada μ ve σ^2 parametreleri, sırasıyla normal dağılıma sahip rassal değişkenin ortalamasını ve varyansını göstermektedir. Dolayısıyla, σ normal dağılıma sahip rassal değişkenin standart sapmasıdır. e; Euler sayısını (e ≈ 2.71) ve π; pi sayısını ($\pi \approx 3.14$) ifade eder. Normal dağılımın $f(x)$ olasılık yoğunluk fonksiyonun grafiği veya normal dağılım eğrisi ($f(x)$ eğrisi), Şekil 6.17'de gösterildiği gibi çan eğrisi şeklindedir.

DİKKAT

Normal dağılımın μ ortalaması sıfır dahil olmak üzere pozitif ve negatif değerler alabilirken ($-\infty < \mu < \infty$); σ standart sapması ise sadece pozitif değerler almaktadır ($\sigma > 0$).

Şekil 6.17'den görüldüğü gibi, μ parametresi yatay eksen üzerindeki normal dağılım eğrisinin merkezini (konumunu), σ parametresi ise dağılımın yayılımını (değişkenliğini) ifade etmektedir. Söz konusu μ ve σ parametrelerinin değerlerine bağlı olarak farklı normal dağılım eğrileri elde edilebilir. Şekil 6.18'de ortalama ve standart sapma değerleri farklı olan normal dağılım eğrileri verilmiştir.

Şekil 6.17

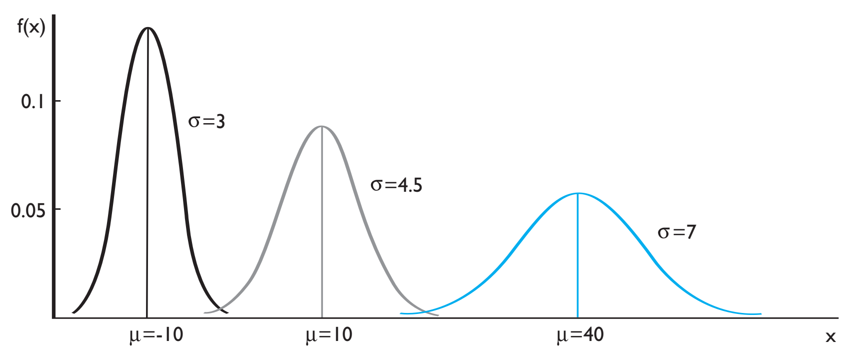


μ ortalama ve σ standart sapmaya sahip normal dağılım eğrisi.

DİKKAT

Normal dağılımın olasılık yoğunluk fonksiyonun grafiği veya normal dağılım eğrisi, çan eğrisi şeklindedir.

Şekil 6.18



Ortalama ve standart sapma değerleri farklı olan normal dağılım eğrileri.

Normal dağılımın $f(x)$ olasılık yoğunluk fonksiyonu aşağıdaki özellikleri sağlar.

i. Her x için $f(x) \geq 0$.

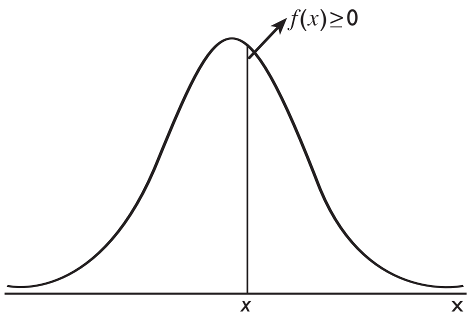
ii. $f(x)$ eğrisi altında kalan ve x-ekseniyle sınırlandırılmış alan 1'e eşittir.

iii. $f(x)$ eğrisi $x = \mu$'ye göre simetriktir.

iv. $f(x)$ eğrisinin iki ucu (kuyruğu) sonsuza gitmektedir.

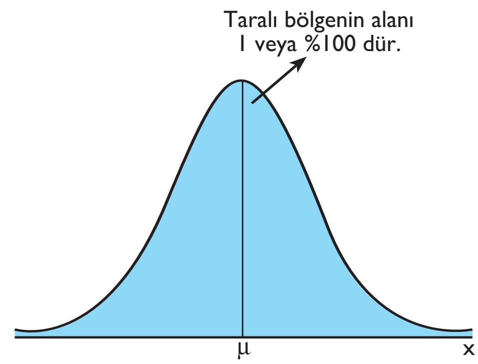
Birinci özellik, normal dağılımın tanımlı olduğu $-\infty < x < \infty$ aralığında, $f(x)$ olasılık yoğunluk fonksiyonunun aldığı değerlerin pozitif veya sıfıra eşit olmasıdır (Şekil 6.19). İkinci özellik, normal dağılımın $f(x)$ eğrisi altında kalan ve x-ekseniyle sınırlandırılmış toplam alanın veya olasılığın 1 olmasıdır (Şekil 6.20).

Şekil 6.19



Normal dağılımın olasılık yoğunluk fonksiyonun pozitif olması.

Şekil 6.20



Normal dağılım eğrisi altındaki toplam alan.

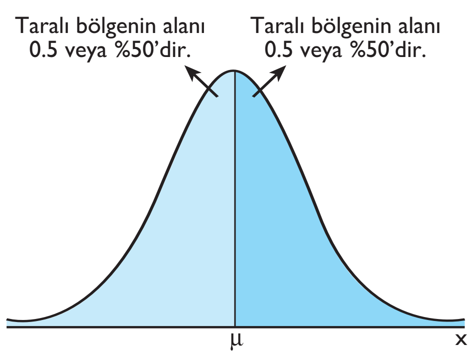
Normal dağılım eğrisi ortalama μ'ye göre simetriktir. Ayrıca, μ'nün sağındaki alan 0.5'e, solundaki alanda 0.5'e eşittir ve eğri altındaki toplam alan 1'dir.

Üçüncü özellik, **normal dağılım eğrisi**nin şeklinin $x = \mu$'nün solunda ve sağında aynı olmasıdır. Ayrıca, ortalama (μ), eğri altındaki toplam alanı iki eşit parçaya ayırır (Şekil 6.21). Dördüncü özellik ise normal dağılımın tanım aralığı $-\infty < x < \infty$ olduğu için normal dağılım eğrisinin iki kuyruğu (ucu) x eksenine dokunmayacak ve kesmeyecek bir biçimde sonsuza kadar uzamasıdır (Şekil 6.22). Fakat uygulamalarda, hemen hemen tüm x değerlerinin $\mu - 3\sigma < x < \mu + 3\sigma$ aralığında değer aldığı görülür ve bu aralığın dışında kalan bölgenin eğri altındaki alanın çok küçük olmasından dolayı sıfır olduğu kabul edilmektedir.

DİKKAT

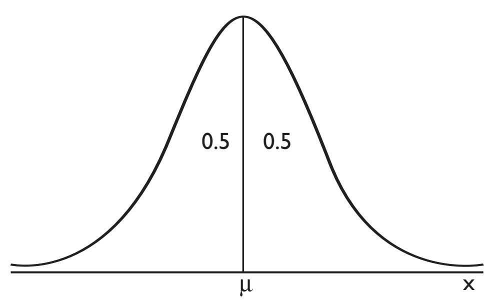
Normal dağılım eğrisinin iki kuyruğu (ucu) x eksenine dokunmayacak ve kesmeyecek bir biçimde sonsuza kadar uzamaktadır.

Şekil 6.21



Normal dağılım eğrisinin μ'ye göre simetrik olması.

Şekil 6.22



Normal dağılım eğrisinin iki kuyruğu.

Daha önce de söz edildiği gibi, sürekli rassal değişkenlerle ilgili olasılık hesabı, $f(x)$ eğrisi altında kalan alanın bulunmasına dayanmaktadır. Normal dağılım için bu alanın bulunması, düzgün dağılımda olduğu kadar kolay değildir. Bu nedenle, olasılık (alan) değerlerini hesaplamak için daha önceden hazırlanmış olan Ek 1'deki tablodan yararlanılır. Söz konusu tablo, $\mu = 0$ ve $\sigma = 1$ olan **normal dağılım (standart normal dağılım)** eğrisi altında ve belli değerler arasında kalan alanları vermektedir.

Bu nedenle, ortalaması μ ve standart sapması σ olan normal dağılıma sahip X rassal değişkeniyle ilgili olasılık hesaplamalarına geçmeden önce standart normal dağılım ve onunla ilgili örnekler üzerinde durulacaktır.

Normal dağılıma sahip rassal bir değişkenle ilgili olasılık hesaplamaları için, daha önceden hazırlanmış olan Ek 1'deki standart normal dağılım tablosundan yararlanılır.

DİKKAT

Standart Normal Dağılım

Standart normal dağılım, normal dağılımın özel bir halidir. Bir başka ifadeyle, ortalaması $\mu = 0$ ve standart sapması $\sigma = 1$ olan normal dağılıma, **standart normal dağılım** denir. Standart normal dağılıma sahip sürekli bir **rassal değişken Z** ile gösterilmektedir. Kısaca $Z \sim N(0,1)$ olarak ifade edilir. Yukarıda verilen normal dağılımın $f(x)$ olasılık yoğunluk fonksiyonunda $\mu = 0$ ve $\sigma = 1$ yazıldığında, standart normal dağılıma sahip Z rassal değişkenin olasılık yoğunluk fonksiyonu:

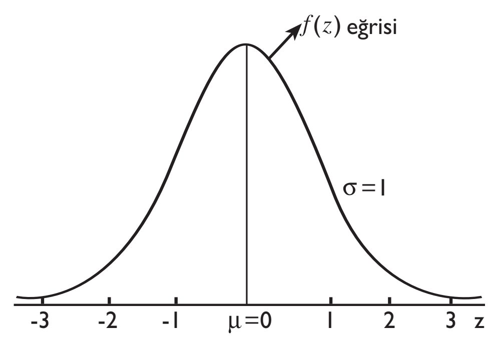
Ortalaması $\mu = 0$ ve standart sapması $\sigma = 1$ olan normal dağılıma, **standart normal dağılım** denir.

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \quad -\infty < z < \infty$$

şeklinde elde edilir. $f(z)$ olasılık yoğunluk fonksiyonunun grafiği veya standart normal dağılım eğrisi Şekil 6.23'te verilmiştir.

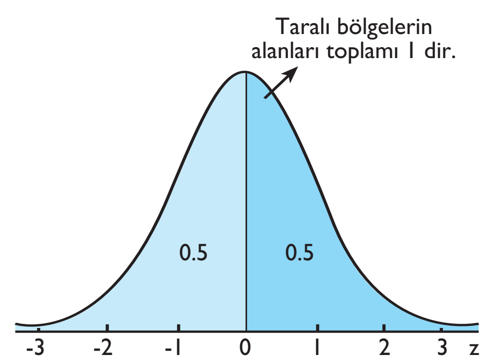
Şekil 6.23'te görüldüğü gibi, standart normal dağılım eğrisinin yatay ekseni (z ekseni) üzerinde işaretlenmiş değerlere, z değerleri veya z skorları denir. Ayrıca, ortalamanın sağında kalan z değerleri pozitif, solunda kalanlar ise negatiftir.

Standart normal dağılım eğrisi altındaki toplam alanın 1 olması ve simetriklik özelliğinden dolayı ortalamanın her iki tarafındaki alanın da 0.5 olması, Şekil 6.24'te özetlenmiştir.



Şekil 6.23

Standart normal dağılım eğrisi.



Şekil 6.24

Standart normal dağılım eğrisi altında kalan toplam alan.

Alan kavramı nedeniyle, $f(z)$ eğrisi altında $z = 0$ ile negatif z değerleri arasında kalan alan pozitiftir.

DİKKAT

Daha önce de belirtildiği gibi, standart normal dağılım eğrisi altında belli iki nokta arasında kalan alanın değerini bulmak için, Ek 1'de verilmiş olan standart normal dağılım tablosundan yararlanılır. Bu tabloda, standart normal dağılım eğrisi altında $z = 0$ ile z'nin 0.00'dan 3.09'a kadar olan değerleri arasındaki alanlar yer almaktadır. Bir başka ifadeyle, $f(z)$ eğrisi altında kalan ve standart normal dağılımın ortalaması olan $z = 0$ değeri ile pozitif z değerleri arasındaki alanlar bu tablodan bulunmaktadır. Standart normal dağılım eğrisi ortalamaya göre simetrik olduğu için, $z = 0$ ile pozitif z değeri ($a > 0$ için $z = a$ gibi) arasındaki alan, negatif z değeri ile $z = 0$ arasındaki alana eşittir. Dolayısıyla, $z = 0$ ile pozitif z değerleri arasındaki alanları veren Ek 1'deki tablo değerleri, negatif z değerleri ile $z = 0$ arasındaki alanlar için de kullanılır.

Ek 1'de verilen tablodan yararlanarak, standart normal dağılım eğrisi altında kalan alanların bulunmasıyla ilgili aşağıda örnekler verilmiştir.

DİKKAT

Ek 1'deki tabloda, standart normal dağılım eğrisi altında $z = 0$ ile z'nin 0.00'dan 3.09'a kadar olan değerleri arasındaki alan veya olasılık değerleri yer almaktadır.

Örnek 4: *Sürekli Z rassal değişkeni standart normal dağılıma sahip olduğu bilindiğine göre, standart normal dağılım eğrisi altında,*

a. $z = 0$ ile $z = 1.68$ arasındaki

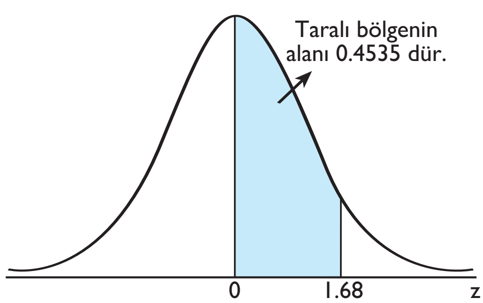
b. $z = -0.73$ ile $z = 0$ arasındaki

alanları bulunuz.

Çözüm: a. Şekil 6.25'ten de görüleceği gibi, burada istenen alan, standart normal dağılım eğrisi altında $z = 0$ ile $z = 1.68$ değerleri arasında kalan taralı bölgenin alanıdır.

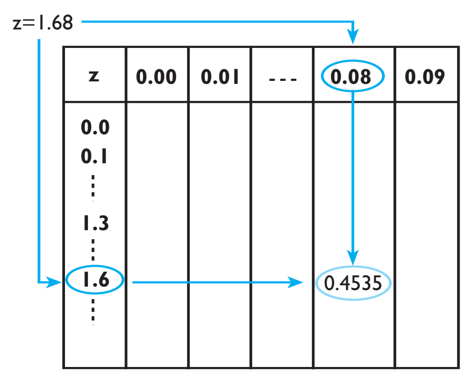
Ek 1'deki standart normal dağılım tablosundan yararlanarak söz konusu alanın sayısal sonucu bulunur. Buna göre, $z = 1.68$ değerine karşılık gelen alanın tabloda bulunması Şekil 6.26'da görselleştirilmiştir.

Şekil 6.25



Standart normal eğri altında $z = 0$ ile $z = 1.68$ arasında kalan alan.

Şekil 6.26



Standart normal dağılım tablosundan $z = 0$ ile $z = 1.68$ arasındaki alanın bulunması.

Şekil 6.26'dan görüldüğü gibi $z = 1.68$ değerinin, ondalık noktasının solundaki ve sağındaki ilk hane (1.6), tablonun ilk sütunu olarak verilen değerlerden seçilir. Ondalık noktasının sağındaki ikinci hane ise (0.08) tablonun ilk satırından seçilir. Daha sonra seçilmiş olan satır ve sütunun kesişim noktasındaki değer, $z = 0$ ile $z = 1.68$ arasında kalan alanının değeridir. Bu değer 0.4535 olarak bulunur.

Ayrıca bu sonuç, standart normal dağılıma sahip olan Z rassal değişkeninin $z = 0$ ile $z = 1.68$ arasında değer alması olasılığıdır ve aşağıdaki gibi de ifade edilir.

$P(0 < Z < 1.68) = 0.4535.$

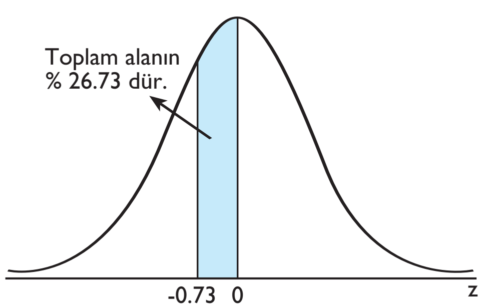
b. Burada aranan alan, standart normal dağılım eğrisi altında $z = -0.73$ ve $z = 0$ noktaları arasında kalan alandır (Şekil 6.27). Daha önce de belirtildiği gibi, standart normal dağılım ortalamaya göre simetrik olduğu için, $z = 0$ ile $z = 0.73$ arasındaki alan, $z = -0.73$ ile $z = 0$ arasındaki alana eşittir. Bundan dolayı sadece $z = 0$ ile z'nin pozitif değerleri arasındaki alanları veren Ek 1'deki tablo, söz konusu alanın bulunması için de kullanılabilir.

Bu durumda, Ek 1'de verilmiş olan tabloda $z = 0.73$ değerinin; 0.7 kısmı ilk sütundan, 0.03 kısmı ise ilk satırdan bulur (Şekil 6.28). Bu seçilmiş olan satır ve sütunun kesişim noktasındaki 0.2673 değeri, $z = -0.73$ ile $z = 0$ veya $z = 0$ ile $z = 0.73$ arasında kalan alanın değeridir. Ayrıca bu değer,

$P(-0.73 < Z < 0) = 0.2673$

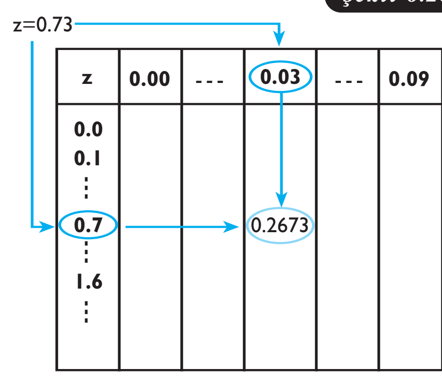
olarak da gösterilebilir.

Şekil 6.27



Standart normal eğri altında $z = -0.73$ ile $z = 0$ arasında kalan alan.

Şekil 6.28



Standart normal dağılım tablosundan $z = -0.73$ ile $z = 0$ arasında kalan alanın bulunması.

Örnek 5: *Standart normal dağılıma sahip Z rassal değişkeni için aşağıda verilen olasılıkları bulunuz.*

a. $P(Z \geq 1.23)$

b. $P(Z \leq -2.70)$

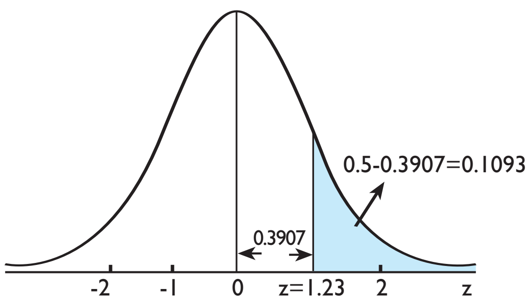
Çözüm: a. Aranan $P(Z \geq 1.23)$ olasılık değeri, standart normal dağılım eğri altında ve $z = 1.23$'ün sağında kalan alandır (Şekil 6.29). Burada dikkat edilmesi gereken husus, standart normal dağılım tablosunda yer alan değerler, ortalama ($z = 0$) ile verilen z değeri arasındaki alanlardır. Ancak soruda, $z = 1.23$ değerinin sağındaki alan sorulmaktadır.

Bu nedenle, $z = 0$ ile $z = 1.23$ arasındaki alan değeri tablodan bulunur ve ortalamanın ($z = 0$'ın) sağındaki toplam alan değeri olan 0.5'ten çıkarılırsa, $z = 1.23$'ün sağındaki alan değeri elde edilir. Buna göre, $z = 0$ ile $z = 1.23$ arasındaki alan 0.3907'dir ve Şekil 6.29' da gösterilen taralı bölgenin alanı veya $P(Z \geq 1.23)$ olasılık değeri,

Taralı Bölgenin Alanı = $P(Z \geq 1.23) = 0.5 - 0.3907 = 0.1093$

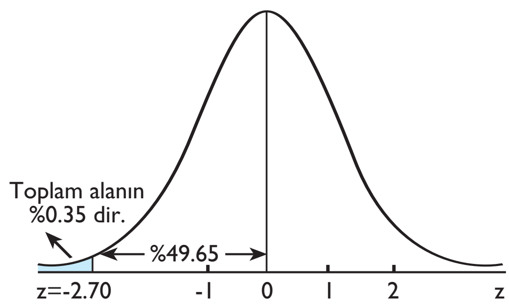
olarak elde edilir.

Şekil 6.29



P (Z > 1.23) olasılık değeri.

Şekil 6.30



P (–2.70 ≤ Z) olasılık değeri.

b. Şekil 6.30'dan görüleceği gibi, $P(Z \leq -2.70)$ olasılık değeri $z = -2.70$'in solundaki taralı bölgenin alanına eşittir. Buna göre ilk olarak, Ek 1'de verilen tablodan $z = -2.70$ ile $z = 0$ arasındaki alan değeri bulunur. Sonra bu değer ortalamanın solundaki toplam alan değeri olan 0.5'ten çıkarılırsa, aranan $P(Z \leq -2.70)$ olasılık değeri elde edilir.

$P(Z \leq -2.70) = 0.5 - 0.4965 = 0.0035.$

Örnek 6: *Standart normal dağılıma sahip Z rassal değişkeni için aşağıda verilen olasılıkların değerlerini bulunuz.*

a. $P(-2.58 < Z < -0.46)$

b. $P(-1.85 \leq Z \leq 1.69)$

Çözüm: a. Burada istenen olasılık değeri, standart normal eğri altında ve $z = -2.58$'den $z = -0.46$'ya kadar olan alana eşittir. Şekil 6.31'den de görüleceği gibi, verilen iki nokta arasındaki alan, ortalamanın sol tarafındadır.

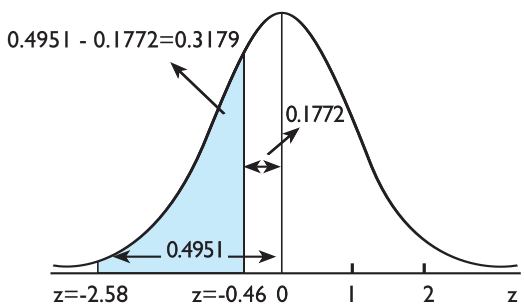
Bu durumda, istenen olasılık değerinin bulunabilmesi için önce $z = -2.58$ ile $z = 0$ arasındaki ve $z = -0.46$ ile $z = 0$ arasındaki alanlar bulunur. Sonra büyük alanın değerinden küçük alanın değeri çıkartılarak, istenen olasılık değeri elde edilir. Sonuç olarak,

$z = -2.58$ ile $z = 0$ arasındaki alan: $P(-2.58 < Z < 0) = 0.4951$,
$z = -0.46$ ile $z = 0$ arasındaki alan: $P(-0.46 < Z < 0) = 0.1772$,
$z = -2.58$ ile $z = -0.46$ arasındaki alan: $P(-2.58 < Z < -0.46) = 0.4951 - 0.1772 = 0.3179$
değerine ulaşılır.

DİKKAT

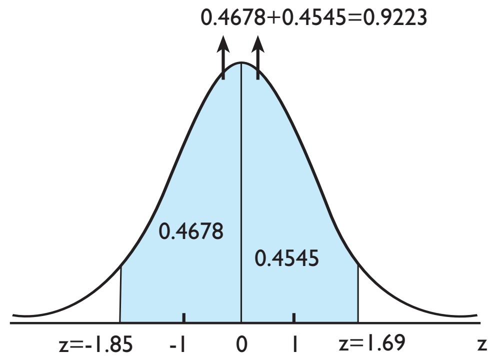
$Z \sim N(0,1)$ için $P(a < Z < b)$ olasılığı sorulduğunda, eğer a ve b noktalarının her ikisi ortalamanın solunda veya sağında ise önce ortalama ile verilen bu noktalar arasındaki alan değerleri bulunur. Sonra büyük alan değerinden küçük alan değeri çıkarılır.

Şekil 6.31



P (–2.58 < Z < –0.46) olasılık değeri.

Şekil 6.32



P (–1.85 ≤ Z ≤ 1.69) olasılık değeri.

b. Şekil 6.32'den görüldüğü gibi, $P(-1.85 \leq Z \leq 1.69)$ olasılık değeri $z = -1.85$ ile $z = 1.69$ değerleri arasındaki alana eşittir. Burada dikkat edilmesi gereken durum, verilen iki noktanın, ortalamanın iki farklı tarafında (solunda ve sağında) olmasıdır.

Bu nedenle, aranan olasılık değerinin bulunabilmesi için önce $z = -1.85$ ile $z = 0$ ve $z = 0$ ile $z = 1.69$ arasındaki alanlar bulunur:

$P(-1.85 < Z < 0) = 0.4678$ ve $P(0 < Z < 1.69) = 0.4545$.

Sonra bu iki alan değeri toplanarak aranan olasılık değeri elde edilir.

$P(-1.85 \leq Z \leq 1.69) = 0.4678 + 0.4545 = 0.9223$.

DİKKAT

$Z \sim N(0,1)$ için $P(a < Z < b)$ olasılığı sorulduğunda, eğer a ve b noktalarının biri ortalamanın solunda ve diğeri sağında ise önce ortalama ile verilen bu noktalar arasındaki alan değerleri bulunur. Sonra bulunan iki alan değeri toplanır.

Örnek 7: *Z rassal değişkeni standart normal dağılım göstermektedir. Bu durumda aşağıdaki olasılıkları bulunuz.*

a. $P(Z < 1.52)$

b. $P(-4.87 < Z < 3.93)$

Standart normal dağılıma sahip **Z rassal değişkeni** için $P(Z < 0) = 0.5$ ve $P(Z > 0) = 0.5$ dir.

Çözüm: a. Burada $P(Z < 1.52)$ olasılık değeri, $z = 1.52$'nin solundaki alana eşittir (Şekil 6.33). Bu durumda istenen alan, $z = 0$ ile $z = 1.52$ arasındaki alan ve $z = 0$'ın solundaki tüm alan olacak şekilde, iki bölümde düşünülebilir. İlk bölümün alanı,

$P(0 < Z < 1.52) = 0.4357$

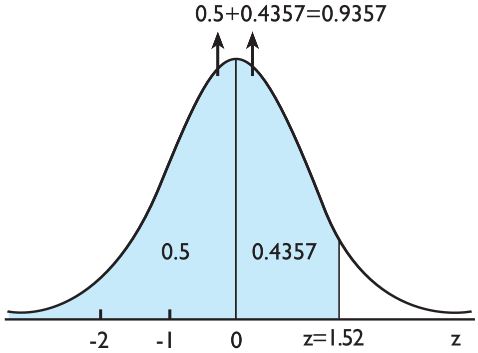
dır ve z = 0'ın veya ortalamanın solundaki alan da

$P(Z < 0) = 0.5$

değerine eşittir. Bu iki alan değeri toplanarak, $P(Z < 1.52)$ olasılık değeri elde edilir.

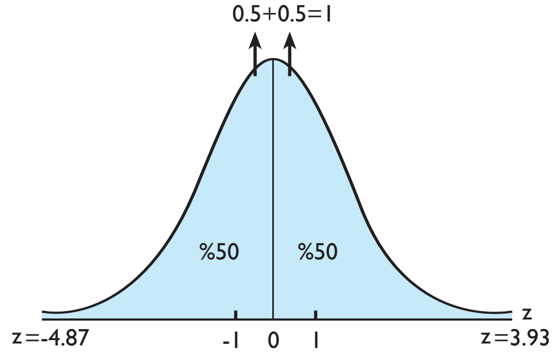
$P(Z < 1.52) = 0.5 + 0.4357 = 0.9357$.

Şekil 6.33



P (Z < 1.52) olasılık değeri.

Şekil 6.34



P (–4.87 < Z < 3.93) olasılık değeri.

b. P (-4.87 < Z < 3.93) değeri, standart normal dağılım eğrisi altında z = -4.87 ile z = 3.93 noktaları arasındaki alana eşittir (Şekil 6.34). Bu durumda, verilen iki nokta, ortalamanın iki farklı tarafında olduğu için, z = -4.87 ile z = 0 arasındaki ve z = 0 ile z = 3.93 arasındaki alanların bulunması gerekir.

Gerekli alanların bulunması için Ek 1'de verilen standart normal dağılım tablosundan yararlanıldığında bir sorunla karşılaşılır. Bu sorun, tabloda en son z = 3.09 değerine kadar olan olasılık (alan) değerinin bulunabilecek olmasıdır. Eğer verilen nokta, z = 3.09'dan büyük (veya z = -3.09'dan küçük) olduğu durumda, z = 0 ile bu nokta arasındaki olasılık değeri 0.5 veya %50 olarak kabul edilmektedir. Buna göre,

z = - 4.87 ile z = 0 arasındaki alan: $P(-4.87 < Z < 0) = 0.5$,

z = 0 ile z = 3.93 arasındaki alan: $P(0 < Z < 3.93) = 0.5$,

dır ve aranan olasılık,

$$P(-4.87 < Z < 3.93) = 0.5 + 0.5 = 1$$

olarak bulunur.

DİKKAT

Eğer verilen nokta z = 3.09'dan büyük (veya z = 3.09'dan küçük) ise ortalamayla bu nokta arasındaki olasılık veya alan değeri 0.5 (%50) olarak kabul edilir.

Normal Dağılım Uygulamaları

Bundan önceki kısımda, Ek 1'de verilen standart normal dağılım tablosundan yararlanılarak, standart normal dağılıma sahip Z rassal değişkeni ile ilgili çeşitli olasılıkların (alanların) bulunması anlatıldı. Fakat gerçek yaşam uygulamalarında ilgilenilen rassal değişkenler, ortalaması $\mu \neq 0$ ve standart sapması $\sigma \neq 1$ olacak şekilde normal dağılım gösterebilmektedir. Bu gibi durumlarda, bu rassal değişkenle ilgili olasılıkların bulunabilmesi için μ ve σ parametrelerine bağlı olarak birçok normal dağılım tablosunun hazırlanması gerekmektedir. Bu zorluğu aşmak için standart normal dağılım tablosunun kullanılabilmesini sağlayan bir dönüşüm yapılarak, olasılık hesaplamaları yapılmaktadır. Normal dağılıma sahip X rassal değişkeninin, standart normal dağılıma sahip Z rassal değişkenine dönüştürülmesine **standartlaştırma** adı verilmektedir.

Ortalaması μ ve standart sapması σ olacak şekilde normal dağılıma sahip X rassal değişkeninin herhangi bir x değeri,

$$z = \frac{x - \mu}{\sigma}$$

formülü yardımıyla ***z* değerine** dönüştürülür. Bir başka ifadeyle, normal dağılım gösteren X rassal değişkeninden, μ ortalaması çıkarılıp σ standart sapmasına bölünerek, standart normal dağılım gösteren $Z = (X - \mu)/\sigma$ rassal değişkeni elde edilir. Bu dönüşüm sonucunda, X rassal değişkeniyle ilgili olasılık hesabı Ek 1'de verilen standart normal dağılım tablosundan yararlanarak yapılır.

μ ortalama ve σ standart sapmayla normal dağılıma sahip X rassal değişkeniyle ilgili olasılıkların bulunması aşağıdaki örneklerle verilmiştir.

Örnek 8: *X rassal değişkeni, μ =100 ve σ = 20 ile normal dağılım göstermektedir. Bu durumda aşağıda verilen olasılıkları bulunuz.*

a. $P(60 < X < 100)$

b. $P(125 < X < 150)$

Çözüm: a. Burada X rassal değişkenin dağılımı $X \sim N(100,20^2)$ olarak da gösterilebilir. Aranan olasılık değeri ise normal eğri altında x = 60 ile x = 100 değerleri arasındaki alanın değerine eşittir (Şekil 6.35). Bu alanın veya olasılığın değeri, yukarıda belirtilen z dönüşümü uygulanarak bulunabilir. Buna göre ilk olarak, verilmiş olan x = 60 ve x = 100 değerlerini standart normal dağılımın z değerlerine dönüştürülmesi gerekir.

x = 60 için z değeri: $z = \frac{x - \mu}{\sigma} = \frac{60 - 100}{20} = -2$

x = 100 için z değeri: $z = \frac{x - \mu}{\sigma} = \frac{100 - 100}{20} = 0.$

Bu dönüşüm yardımıyla aranan olasılık, standart normal dağılım eğrisi altında z = -2 ile z = 0 değerleri arasındaki alan olur:

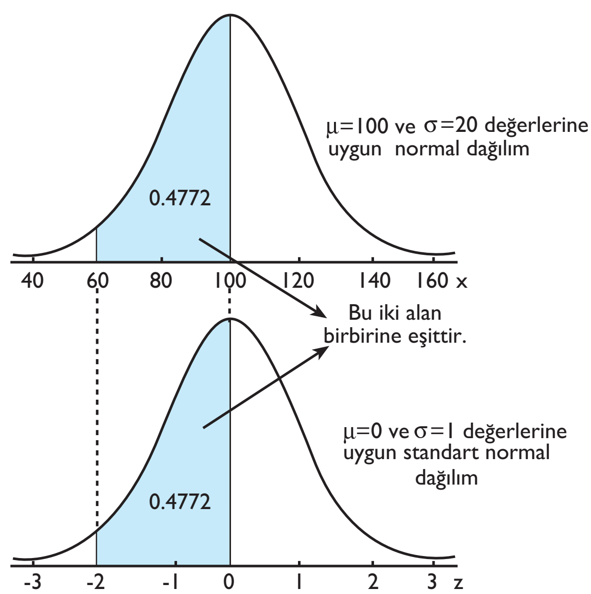
$P(60 < X < 100) = P(-2 < Z < 0)$.

Bu alan değerinin bulunması için, Ek 1'de verilen standart normal dağılım tablosundan yararlanıldığında,

$P(60 < X < 100) = P(-2 < Z < 0) = 0.4772$

sonucu elde edilir.

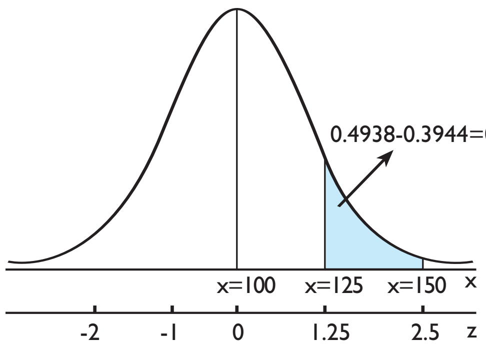
Şekil 6.35



Normal eğri altındaki alanın standart normal eğri yardımıyla gösterimi.

b. Benzer şekilde $P(125 < X < 150)$'in değerini bulmak için ilk olarak, verilmiş olan $x = 125$ ve $x = 150$ değerleri,

Şekil 6.36



P (125 < X < 150) olasılık değeri.

$x = 125$ için z: $z = \dfrac{125-100}{20} = 1.25$

$x = 150$ için z: z: $z = \dfrac{150-100}{20} = 2.5$

biçiminde standart normal z değerlerine dönüştürülür. Sonra $z = 1.25$ ile $z = 2.5$ arasındaki alan, Ek 1'deki tablo yardımıyla bulunur.

Şekil 6.36'dan da görüleceği gibi, burada verilen değerlerin her ikisi ortalamanın sağında olduğu için, $z = 0$ ile $z = 2.5$ arasındaki alandan $z = 0$ ile $z = 1.25$ arasındaki alan çıkartılarak, istenen olasılık değeri bulunur.

$P(100 < X < 125) = P(0 < Z < 1.25) = 0.3944$
$P(100 < X < 150) = P(0 < Z < 2.50) = 0.4938$
$P(125 < X < 150) = P(1.25 < Z < 2.50) = 0.4938 - 0.3944 = 0.0994.$

Örnek 9: *X rassal değişkeni, 11.22 ortalama ve 2.68 standart sapmayla normal dağılıma sahiptir. X rassal değişkeninin,*

a. $x = 9.36$'dan küçük değer alması,
b. $x = 10.58$'den büyük değer alması,

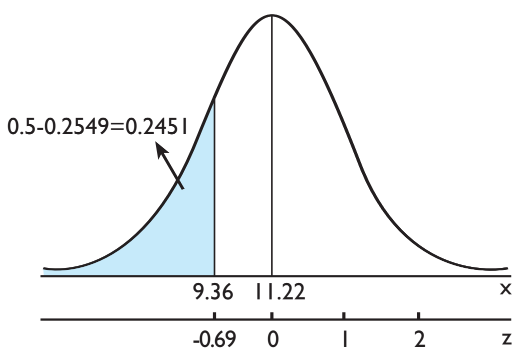
olasılıklarını bulunuz.

Çözüm: Burada normal dağılımının ortalaması $\mu = 11.22$ ve standart sapması $\sigma = 2.68$ olarak verilmiştir. İstenen olasılıklar aşağıdaki gibi bulunur.

a. Burada X rassal değişkeninin 9.36'dan küçük değer alması olasılığı ile $P(X < 9.36)$ olasılık değeri sorulmaktadır. Bu aranan olasılık değeri de, normal dağılım eğrisi altında $x = 9.36$'nın solunda kalan alana eşittir. Buna göre Ek 1'de verilen tablodan yaralanabilmek için, $x = 9.36$ değerinin z değerine dönüştürülmesi gerekir.

Şekil 6.37

Normal eğri altında x = 9.36 noktasının solunda kalan alan.



$x = 9.36$ için z değeri:

$$z = \frac{9.36 - 11.22}{2.68} = -0.69$$

Bu durumda aranan alan, $z = -0.69$ değerinin sol tarafındaki alan olur (Şekil 6.37). Bu nedenle, $z = 0$'ın solundaki alan değeri olan 0.5'ten $z = -0.69$ ile $z = 0$ arasındaki alanın değeri çıkartıldığında, aranan olasılık değerine ulaşılır.

$P(X < 9.36) = P(Z < -0.69) = 0.5 - 0.2549 = 0.2451.$

b. Benzer biçimde, X'in 10.58'den büyük değerler alması olasılığını ($P(X > 10.58)$) bulmak için, z standart değeri,

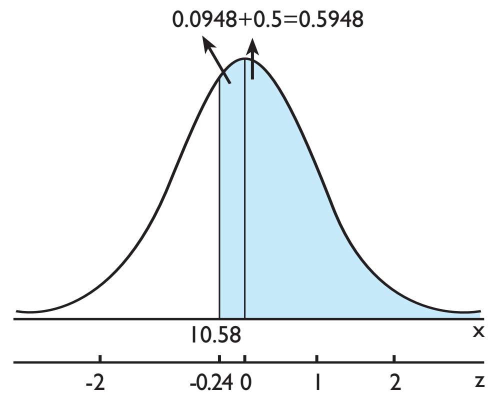
x = 10.58 için z değeri:

$$z = \frac{10.58 - 11.22}{2.68} = -0.24$$

elde edilir. Şekil 6.38'den de görülebileceği gibi, z = -0.24'ün sağında kalan alan iki parçadan oluşmaktadır. Bu durumda, ortalamanın sağındaki tüm alan ile z = -0.24 ve ortalama arasındaki alan toplanarak, istenen olasılık elde edilir.

$$P(X > 10.58) = P(Z > -0.24) = 0.5 + 0.0948 = 0.5948.$$

Şekil 6.38



Normal eğri altında x =10.58 noktasının sağında kalan alan.

Örnek 10: *Ortalaması 20 ve standart sapması 1.25 olacak şekilde normal dağılmakta olan X rassal değişkeni için aşağıdaki olasılıkları bulunuz.*
a. $P(X \geq 23.625)$
b. $P(X \geq 25.125)$

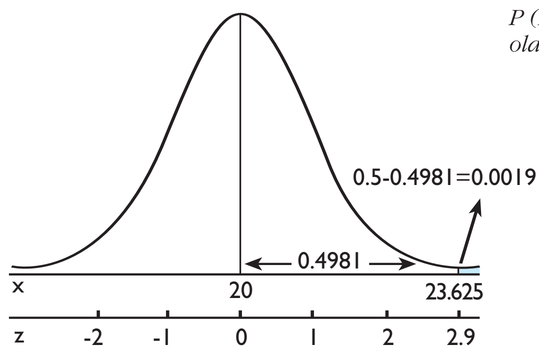
Çözüm: Burada normal dağılım için μ = 20 ve σ = 1.25 olarak verilmiştir ve ayrıca $X \sim N(20,1.25^2)$ olarak da gösterilebilir.

a. $P(X \geq 23.625)$ değerini bulmak için, x = 23.625 değerinin z değerine dönüştürülmesinden,

$$z = \frac{23.625 - 20}{1.25} = 2.9$$

elde edilir. Bu durumda aranan olasılık değeri, eğrinin sağ ucundaki taralı bölgenin alanıdır (Şekil 6.39). Bu bölgenin alanı, ortalamanın sağında kalan alanın değeri olan 0.5'ten z = 0 ile z = 2.9 arasındaki alan değeri çıkartılarak elde edilir. Sonuç olarak istenen olasılık değeri,

Şekil 6.39

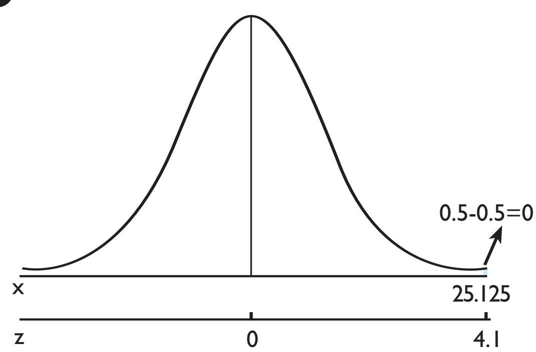


$P(X \geq 23.625)$ olasılık değeri.

$$P(X \geq 23.625) = P(Z \geq 2.9) = 0.5 - 0.4981 = 0.0019$$

olarak bulunur.

Şekil 6.40



P (X > 25.125) olasılık değeri.

b. Benzer olarak, verilen x = 25.125 değerinin z cinsinden değeri,

$$z = \frac{25.125 - 20}{1.25} = 4.1$$

elde edilir. Buna göre, X'in 25.125'ten büyük değer alma olasılığı aşağıdaki gibi hesaplanır.

$P(X \geq 25.125) = P(Z \geq 4.1)$
$= 0.5 - 0.5 = 0.$

Örnek 11: *Sürekli X rassal değişkeni, 16 ortalama ve 16* ***varyansla*** *normal dağılmaktadır. X rassal değişkeni için aşağıda verilen olasılıkların değerlerini bulunuz.*

a. $P(12 < X < 20)$
b. $P(8 < X < 24)$
c. $P(4 < X < 28)$

Çözüm: Burada verilen normal dağılımda μ = 16 ve σ = 4 (σ^2= 16)'tür.

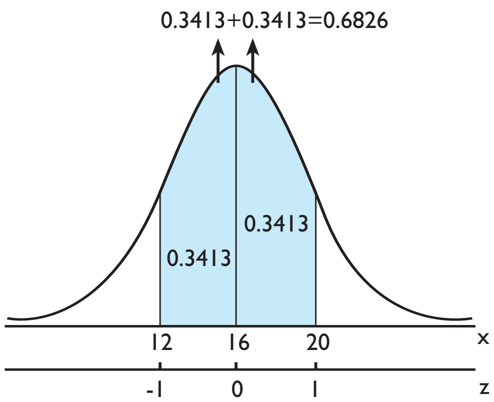
a. $P(12 < X < 20)$ olasılığının değeri, Ek 1'de verilmiş olan standart normal dağılım tablosu yardımıyla bulunacaktır. Bunun için x = 12 ve x = 20 noktalarının z cinsinden değerleri,

x = 12 için z: $z = \frac{12 - 16}{4} = -\frac{4}{4} = -1$

x = 20 için z: $z = \frac{20 - 16}{4} = \frac{4}{4} = 1$

biçiminde hesaplanır. Şekil 6.41'den görüldüğü gibi, z = -1 ile z = 1 noktaları arasındaki alan, z = 0'ın solunda ve sağındaki alanların toplamından oluşmaktadır. Bu durumda aranan olasılık değeri,

Şekil 6.41



Normal eğri altında ve ortalamanın bir standart sapma sınırları içerisindeki alan.

Ortalaması μ ve standart sapması σ olacak şekilde normal dağılıma sahip rassal değişkeni için, $P(\mu - \sigma < X < \mu + \sigma) = P(-1 < Z < 1)$ değeri toplam alanın **% 68.26**'sına eşittir.

$P(12 < X < 16) = P(-1 < Z < 0) = 0.3413$
$P(16 < X < 20) = P(0 < Z < 1) = 0.3413$
$P(12 < X < 20) = P(-1 < Z < 1) = 0.3413 + 0.3413 = 0.6826$

olarak elde edilir. Sonuç olarak, ortalamanın bir standart sapma sağında ve solunda kalan noktalar arasındaki alan ($P(12 < X < 20)$), toplam alanın **% 68.26**'sıdır.

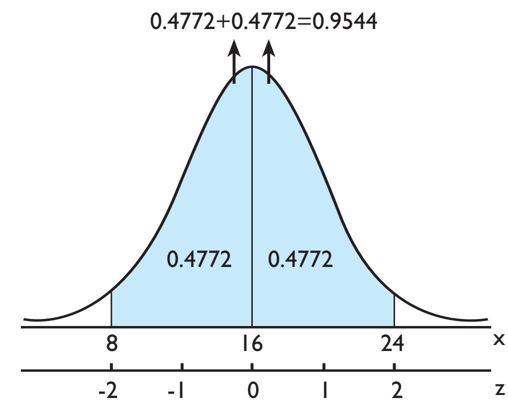
b. Yine $P(8 < X < 24)$ olasılığının değerini bulmak için, $x = 8$ ve $x = 24$ noktalarının z değerlerine dönüştürülmesinden,

$x = 8$ için z: $z = \dfrac{8-16}{4} = -\dfrac{8}{4} = -2$

$x = 24$ için z: $z = \dfrac{24-16}{4} = \dfrac{8}{4} = 2$

elde edilir. Bu noktalar ortalamanın solunda ve sağında olduğu için, tablodan elde edilecek alan değerleri toplanarak istenen olasılık değeri bulunur.

Şekil 6.42



Normal eğri altında ve ortalamanın iki standart sapma sınırları içerisindeki alan.

$P(8 < X < 16) = P(-2 < Z < 0) = 0.4772$
$P(16 < X < 24) = P(0 < Z < 2) = 0.4772$
$P(8 < X < 24) = P(-2 < Z < 2) = 0.4772 + 0.4772 = 0.9544.$

Bir başka ifadeyle, ortalamanın iki standart sapma sağında ve solunda kalan noktalar arasındaki alan ($P(8 < X < 24)$), toplam alanın **%95.44**'tür sonucuna ulaşılır.

Ortalaması μ ve standart sapması σ olacak şekilde normal dağılıma sahip X rassal değişkeni için, $P(\mu - 2\sigma < X < \mu + 2\sigma) = P(-2 < Z < 2)$ değeri toplam alanın **%95.44**'üne eşittir.

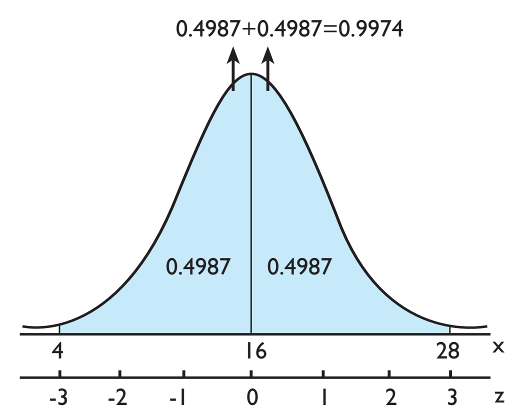
c. Burada yine, $x = 4$ ve $x = 28$ değerlerine karşılık gelen z değerleri,

$x = 4$ için z: $z = \dfrac{4-16}{4} = -\dfrac{12}{4} = -3$

$x = 28$ için z: $z = \dfrac{28-16}{4} = \dfrac{12}{4} = 3$

olarak dönüştürülür. Şekil 6.43'ten de görüldüğü gibi, ortalamanın sağında ve solunda kalan alanlar toplanarak,

Şekil 6.43



Normal eğri altında ve ortalamanın üç standart sapma sınırları içerisindeki alan.

$P(4 < X < 16) = P(-3 < Z < 0) = 0.4987$
$P(16 < X < 28) = P(0 < Z < 3) = 0.4987$
$P(4 < X < 28) = P(-3 < Z < 3) = 0.4987 + 0.4987 = 0.9974$

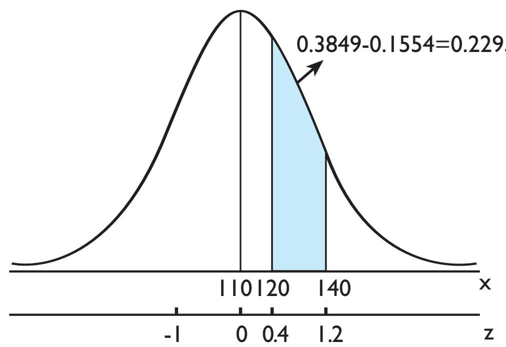
aranan olasılık değeri elde edilir. Bu sonuç, ortalamanın üç standart sapma sağındaki ve solundaki noktalar arasında kalan alan ($P(4 < X < 28)$), toplam alanın **%99.74**'tür olarak yorumlanır.

Ortalaması μ ve standart sapması σ olacak şekilde normal dağılıma sahip X rassal değişkeni için, $P(\mu - 3\sigma < X < \mu + 3\sigma) = P(-3 < Z < 3)$ değeri toplam alanın **%99.74**'üne eşittir.

Örnek 12: *Bir şirkette yapılan araştırmada, çalışanların iş yerinde bir yıl içinde interneti kullanma süresi 110 saat ortalama ve 25 saat standart sapmayla normal dağıldığı bulunmuştur. Bu şirketten rassal olarak seçilen bir çalışanın bir yıl içinde interneti kullanma süresinin 120 saat ile 140 saat arasında olması olasılığını bulunuz.*

Çözüm: X rassal değişkeni, bir yıl içinde interneti kullanma süresi olmak üzere, $\mu = 110$ saat ve $\sigma = 25$ saatle normal dağılım göstermektedir. Burada istenen $P(120 < X < 140)$ olasılık değeri, z dönüşümünden yararlanarak bulunacaktır. Buna göre, verilen x noktalarının z cinsinden değerleri,

Şekil 6.44



Normal eğri altında x = 120 ile x = 140 sınırları arasındaki alan.

$x = 120$ için z: $z = \dfrac{120 - 110}{25} = 0.4$

$x = 140$ için z: $z = \dfrac{140 - 110}{25} = 1.2$

olarak elde edilir. Şekil 6.44'ten de görüleceği gibi, ortalama ile bu iki nokta arasındaki alanlar Ek 1'deki tablodan bulunup, büyük alandan küçük alanın çıkartılmasıyla istenen olasılık (alan) değeri elde edilir.

$$P(110 < X < 120) = P(0 < Z < 0.4) = 0.1554$$
$$P(110 < X < 140) = P(0 < Z < 1.2) = 0.3849$$
$$P(120 < X < 140) = P(0.4 < Z < 1.2)$$
$$= 0.3849 - 0.1554 = 0.2295.$$

Diğer bir deyişle, rassal olarak seçilen bir çalışanın bir yıl içinde interneti kullanma süresinin 120 - 140 saat arasında olması olasılığı 0.2295'tir.

Örnek 13: *Bir kömür ocağı işletmesinde çıkarılan kömürler, otomatik bir makine yardımıyla çuvallara doldurulmaktadır. İşletme yetkililerinin yaptığı araştırma sonucunda, kömür doldurulan çuvalların ağırlıklarının 50 kg ortalama ve 2 kg standart sapmayla normal dağıldığı bulunmuştur. Otomatik makine yardımıyla doldurulan çuvallardan rassal olarak seçilen bir çuvalın ağırlığının,*

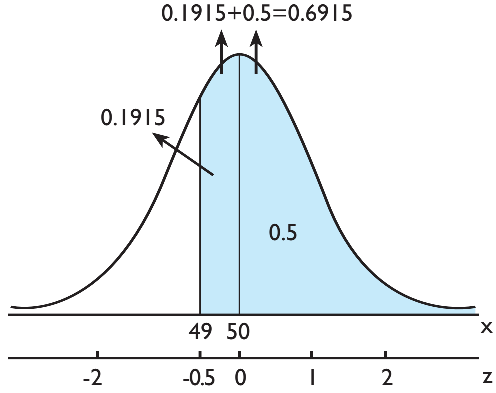
a. 49 kg'dan fazla olması,

b. 47 kg ile 52 kg arasında olması,

olasılıklarını bulunuz.

Şekil 6.45

P (X > 49) olasılık değeri.



Çözüm: X rassal değişkeni otomatik makine yardımıyla doldurulan çuvalların ağırlığını göstermekte ve $\mu = 50$ kg ile $\sigma = 2$ kg değerleriyle normal dağılmaktadır ($X \sim N(50,4)$).

a. Buna göre rassal olarak seçilen bir çuvalın ağırlığının 49 kg'dan fazla olması olasılığı ile $P(X > 49)$'nin değerinin bulunması istenmektedir. Standart normal dağılım tablosunu kullanabilmek için gerekli z değeri,

$x = 49$ için z değeri: $z = \dfrac{49-50}{2} = -0.5$

olarak bulunur. Burada x değeriyle ilgili sadece alt sınır verilmiştir (Şekil 6.45). Bu nedenle, ortalamanın sağındaki tüm alan (0.5) ile $z = -0.5$ ve $z = 0$ arasındaki alan toplanarak gerekli sonuç,

$P(X > 49) = P(Z > -0.5) = 0.5 + 0.1915 = 0.6915$

elde edilir. Bu sonuç, rassal olarak seçilen bir çuvalın ağırlığının 49 kg'dan fazla olması olasılığının 0.6914 (% 69.14) olduğu anlamına gelir.

b. $P(47 < X < 52)$ olasılık değerinin bulunması için ilk olarak verilen x noktalarının z değerleri,

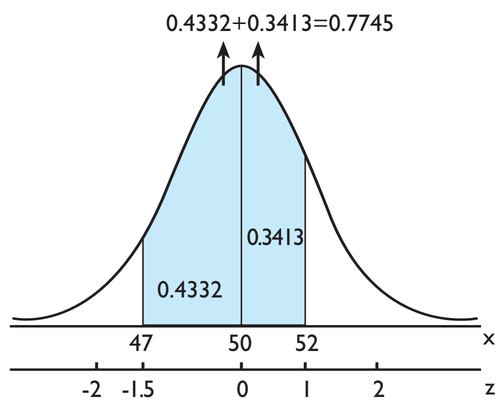
$x = 47$ için z değeri: $z = \dfrac{47-50}{2} = -1.5$

$x = 52$ için z değeri: $z = \dfrac{52-50}{2} = 1$

elde edilir. Şekil 6.46'dan da görüleceği gibi, sonra, ortalamanın sağında ve solunda bulunan alan değerleri toplanarak istenen olasılık değeri bulunur.

Şekil 6.46

$P(47 < X < 52)$ olasılık değeri.



$P(47 < X < 50) = P(-1.5 < Z < 0) = 0.4332$
$P(50 < X < 52) = P(0 < Z < 1) = 0.3413$
$P(47 < X < 52) = P(-1.5 < Z < 1) = 0.4332 + 0.3413 = 0.7745.$

Örnek 14: *Bir bankanın yeni çıkardığı kredi kartının 70000 müşterisi bulunmaktadır. Banka yöneticileri, ocak ayı boyunca bu kredi kartını kullanarak belli miktarlarda harcama yapan müşterilerine hediyeler dağıtacağını duyuracaktır. Banka yöneticileri yaptığı araştırma sonucunda, ocak ayı boyunca bu kredi kartı müşterilerin yapacağı harcamaların ortalama ₺525 ve ₺175 standart sapmayla normal dağılacağını düşünmektedirler. Bu durumda, banka yöneticilerinin yaptığı toplantıda ortaya çıkan,*

a. Kaç kredi kartı müşterisinin yapacağı harcamaların ₺800 ile ₺1000 arasında olacağı,

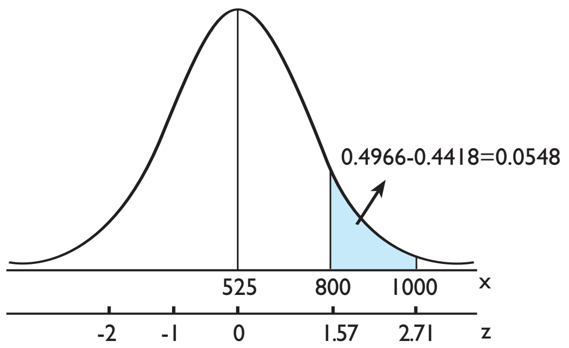
b. Kaç kredi kartı müşterisinin yapacağı harcamaların ₺1000 ve üzerinde olacağı,

sorularının yanıtlarını bulunuz.

Çözüm: Kredi kartı müşterilerinin ocak ayı boyunca yapacağı harcamalar X rassal değişkeni olmak üzere, μ = ₺525 ile σ = ₺175 değerleriyle normal dağılım göstermektedir.

Şekil 6.47

Normal eğri altında x = 800 ile x = 1000 sınırları arasındaki alan.



a. Bu bilgilere göre, sorunun cevabı için $P(800 < X < 1000)$ olasılık değerinin bulunması gerekir. Buna göre, verilen x değerlerinin z dönüşümü sonucunda,

$$x = 800 \text{ için } z = \frac{800 - 525}{175} = 1.57$$

$$x = 1000 \text{ için } z = \frac{1000 - 525}{175} = 2.71$$

elde edilir. Şekil 6.47'den görüldüğü gibi, ortalamanın sağındaki iki alanın değerlerinin farkı alınarak, istenen olasılık değeri,

$P(525 < X < 800) = P(0 < Z < 1.57) = 0.4418$
$P(525 < X < 1000) = P(0 < Z < 2.71) = 0.4966$
$P(800 < X < 1000) = P(1.57 < Z < 2.71) = 0.4966 - 0.4418 = 0.0548$

elde edilir. Buna göre, ₺800 ile ₺1000 arasında harcamalar yapacak müşterilerin oranı % 5.48'dir. Dolayısıyla müşteri sayısı, (70000)(0.0548) = 3836 işlemi sonucuyla bulunur. Bu sonuç, bankanın 3836 müşterisinin ₺800-1000 arasında harcamalar yapacağı biçiminde yorumlanır.

b. Bir önceki soruda x = 1000 için elde edilen z = 2.71 değeri, bu sorunun cevabı için de kullanılacaktır. Fakat burada P (X ≥ 1000)'in olasılık değeri arandığı için x = 1000 sağında kalan alanla ilgilenilmektedir.

Bu nedenle, ortalamanın sağındaki tüm alan değeri olan 0.5'ten ortalama ile z = 2.71 arasındaki alan çıkartılarak, gerekli olasılık

$P(X \geq 1000) = P(Z \geq 1000) = 0.5 - 0.4966 = 0.0034$

olarak bulunur. Bu durumda, ₺1000'nin üzerinde harcamalar yapacak müşterilerin oranı % 0.34'tür. Buna göre, müşteri sayısını bulmak için (70000)(0.0034) = 238 işlemi yapılır. Bir başka ifadeyle, 238 müşterinin ₺1000'nin üzerinde harcamalar yapacağı tahmin edilir.

Normal Dağılım İçin Olasılık Değeri Biliniyorken Uygun *z* ve *x* Değerlerinin Bulunması

Bundan önceki bölümlerde, normal dağılıma sahip sürekli bir rassal değişkenin verilen bir aralıkta bulunmasına ilişkin olasılık (alan) hesabı açıklandı. Şimdi ise normal dağılım için olasılık (alan) değeri biliniyorken, uygun z ve x değerlerinin bulunması ele alınacaktır. Bu değerin bulunması için aşağıda verilen adımlar izlenecektir.

Adım 1: Bilindiği gibi, μ ve σ değerleriyle normal dağılıma sahip bir rassal değişkenin belli bir aralıkta değer alması olasılığı, z dönüşümü kullanılarak standart normal dağılım tablosundan bulunmaktadır. Bu nedenle ilk olarak, normal dağılım için verilen olasılık (alan) değerine uygun z değerlerinin bulunması gerekir. Bu değerlerin bulunması için de Ek 1'de verilen standart normal dağılım tablosundan yararlanılır.

Adım 2: Bulunan z değerinden x değerine, z dönüşümünün **ters dönüşümü** olan,

$$x = \mu + z\sigma$$

formülü yardımıyla geçilir.

Yukarıda verilen adımlara dikkat edilirse, normal dağılımda verilen x değerlerine karşılık gelen olasılık değerinin bulunmasında izlenen adımların tam tersi izlenmektedir. Bununla ilgili örnekler aşağıda verilmiştir.

DİKKAT

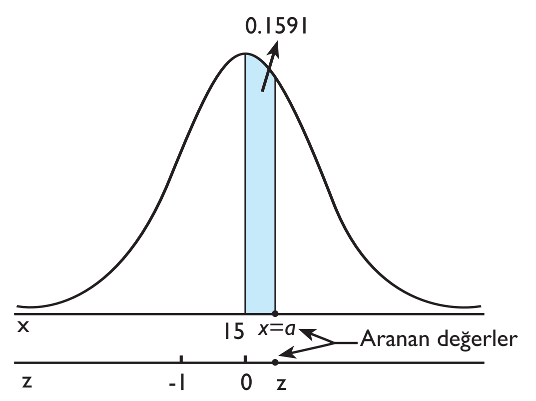
Ortalaması μ ve standart sapması σ olan normal dağılım için olasılık veya alan değeri verildiğinde, ilk olarak bu olasılığa uygun z değeri, Ek 1'de verilen standart normal dağılım tablosundan bulunur. Sonra bulunan z değerinden x değerine, $x = \mu + z\sigma$ formülü yardımıyla geçilir.

Örnek 15: *Sürekli X rassal değişkeni, 15 ortalama 3 standart sapmayla normal dağılıma sahiptir. $P(15 \leq X \leq a) = 0.1591$ olduğuna göre, uygun a değerini bulunuz.*

Çözüm: Bu örnekte normal dağılımın parametreleri $\mu = 15$ ve $\sigma = 3$'tür. Verilen $P(15 \leq X \leq a) = 0.1591$ olasılık değeri, Şekil 6.48'den görüleceği gibi normal eğri altında ortalama değeri olan 15 ile a arasında sınırlı olan alandır.

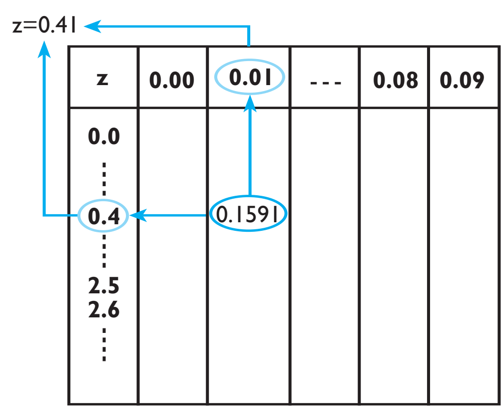
Bu a değerini bulmak için ilk olarak, verilen olasılık veya alan değerine uygun z değerinin standart normal dağılım tablosundan bulunması Şekil 6.49'da görsel olarak anlatılmıştır.

Şekil 6.48



Verilen olasılık değerine uygun z ve x değerlerinin gösterimi.

Şekil 6.49



Verilen olasılık değerine uygun z değerinin tablodan bulunması.

Şekil 6.49'dan görüldüğü gibi, 0.1591 olarak verilen olasılık (alan) değeri tablo içerisinde bulunduktan sonra, bu değerin bulunduğu satırın başındaki 0.4 ve sütunun başındaki 0.01 değerlerinin birleştirilmesi sonucunda aranan z değeri, 0.41 olarak bulunur. Bir başka ifadeyle, 0.1591 değeri, $z = 0$ ile $z = 0.41$ arasında kalan alanının değeridir.

0.1591 alan değerine uygun z değerinin bulunmasından sonra, ikinci adımda verilmiş olan eşitlik kullanılarak aranan a değeri,

$$x = \mu + z\sigma = 15 + (0.41)(3) = 16.23$$

olarak elde edilir.

Örnek 16: *μ = 28.6 ve σ = 8.1 değerleriyle normal dağılmakta olan X rassal değişkeni için,*

a. $P(X \geq a) = 0.10$ olduğuna göre, uygun a değerini bulunuz.

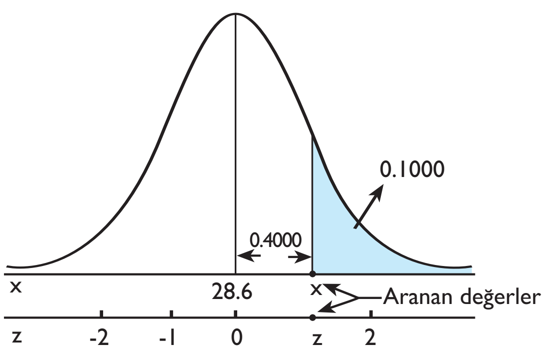
b. $P(X \leq b) = 0.15$ olduğuna göre, uygun b değerini bulunuz.

Çözüm: a. Burada $P(X \geq a) = 0.10$ şeklinde verilen olasılık değeri, normal eğrinin sağ ucunda bulunan bir alandır (Şekil 6.50). Dolayısıyla, bu verilen alana uygun olarak aranan a değeri, bir alt sınırdır.

Burada verilen olasılık değerine uygun z değerinin tablodan bulunması sırasında dikkat edilmelidir. Bunun nedeni, verilen alanın eğrinin sağ ucunda olmasıdır. Fakat tablodaki değerler, $z = 0$ (ortalama) ile verilen z değerleri arasındadır. Bu nedenle a değerini bulmak için, $z = 0$ ile $z = 0$'ın sağında aranan nokta arasındaki alanın dikkate alınması gerekir. Bu alan, normal dağılımda ortalamanın sağındaki tüm alanın değeri 0.5 olasından dolayı, 0.5000 - 0.1000 = 0.4000 olarak bulunur.

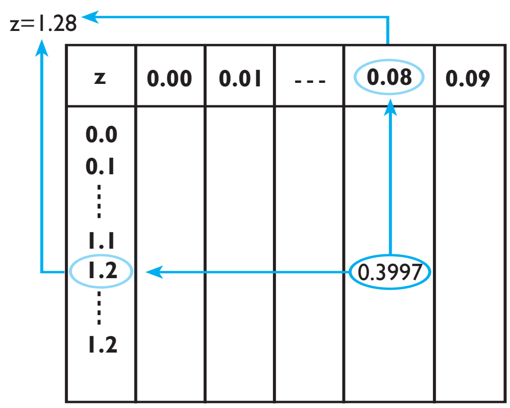
Şekil 6.51'den de görüleceği gibi, tablodaki alan değerleri arasından 0.4000 değeri yaklaşık olarak (0.3997) belirlenir ve bu alan için uygun z değeri 1.28 olarak bulunur. Diğer bir değişle, $z = 1.28$ noktasının sağında kalan alan değeri, verilen 0.10 değerine (yaklaşık) eşittir.

Şekil 6.50



Verilen 0.1 olasılık değerine uygun z ve x değerlerinin gösterimi.

Şekil 6.51



Verilen olasılık değerine uygun z değerinin tablodan bulunması.

Bulunan $z = 1.28$ değeri formülde yerine yazıldığında, aranan a değeri

$$x = \mu + z\sigma = 28.6 + (1.28)(8.1) = 38.968$$

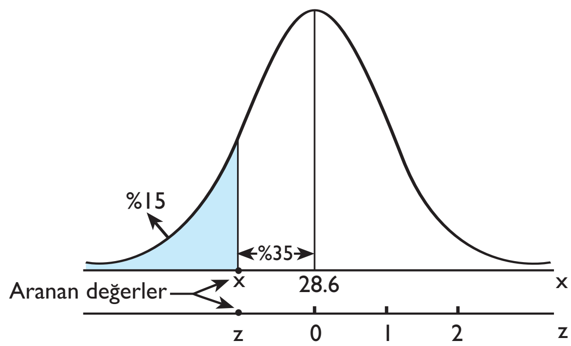
olarak elde edilir.

DİKKAT

Ortalaması μ ve standart sapması σ olan normal dağılım için verilen olasılık (alan) değeri normal eğrinin sağ veya sol ucundaysa, 0.5 alanından verilen alan çıkarılarak elde edilen alan için uygun *z* değerinin bulunması gerekir.

b. Şekil 6.52'den de görüleceği gibi, normal eğrinin sol ucunda bulunan alanın değeri, $P(X \leq b) = 0.15$ olarak verilmiştir. Bu nedenle, aranan b değeri, bir üst sınır oluşturur. Bu sınıra karşılık gelen z değeri, ortalama ile aranan nokta arasındaki alan bulunarak elde edilir. Bu alan, 0.5000 - 0.1500 = 0.3500'dür ve ortalamanın sol tarafında olduğu için uygun z değerinin negatif olması gerekir. Buna göre, bu alana uygun z değeri de tablo yardımıyla yaklaşık olarak -1.04 şeklinde elde edilir. Formülde z = -1.04 değeri yazıldığında, aranan b değeri

Şekil 6.52



Verilen 0.15 olasılık değerine uygun z ve x değerlerinin gösterimi.

$$x = \mu + z\sigma = 28.6 + (-1.04)(8.1) = 20.176$$

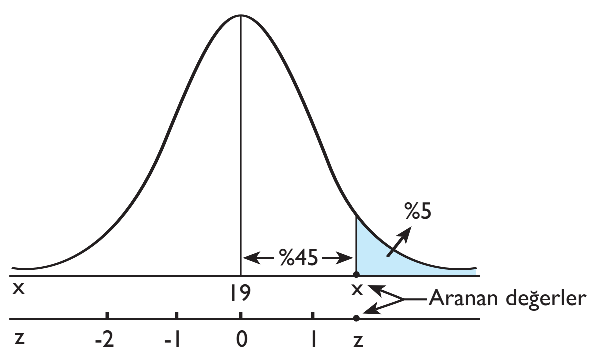
olarak bulunur.

Örnek 17: *Cep telefonu satan bir firma, cep telefonu kullanım sürelerinin 19 ay ortalama ve 4 ay standart sapmayla normal dağıldığı bilgisine sahiptir. Bu bilgi üzerine belli bir süreden fazla cep telefonunu kullanan müşterilerine, yeni bir cep telefonu almaları halinde indirim yapacağını duyuracaktır. Bu firma belirleyeceği sürenin, eski cep telefonunu en uzun süreli kullananların bulunduğu %5'lik kısma girmesini istemektedir. Buna göre duyurulması gereken sürenin alt sınırını belirleyiniz.*

Çözüm: X rassal değişkeni, eski cep telefonunu kullanma süresidir ve μ = 19 ay ile σ = 4 ay değerleriyle normal dağılmaktadır. Şekil 6.53'te x ekseni cep telefonunu kullanma süresini ifade etmektedir. Dolayısıyla, en uzun süreli cep telefonunu kullananlar eğrinin sağ ucunda, yani taralı alanda olacaktır. Taralı alanda x'in bulunduğu yer ise en uzun süreli cep telefonunu kullananların olduğu %5'lik kısım (alan) için alt sınırdır. Dolayısıyla aranan nokta burasıdır. Sonuç olarak belirlenen bu x değeri, firmanın indirim yapabilmesi için duyurması gereken süre olacaktır.

Bu sürenin bulunabilmesi için, ilk olarak ortalama ile istenen nokta arasındaki 0.5000 - 0.0500 = 0.4500 alanı dikkate alınır ve bu alana uygun z değeri Ek 1'de verilen tablodan, 1.65 olarak bulunur. Sonra, z = 1.65 değeri formülde yerine yazılarak, aranan x değeri,

Şekil 6.53



Verilen olasılık değerine uygun z ve x değerlerinin gösterimi.

$x = \mu + z\sigma = 19 + (1.65)(4) = 25.6$

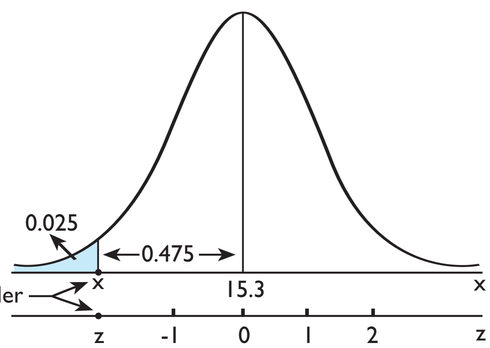
elde edilir. Sonuç olarak, firma 26 (yaklaşık olarak) aydan fazla cep telefonlarını kullananlara indirim yapacağını duyurursa bu 26 aylık süre, eski cep telefonunu en uzun süreli kullananların olduğu %5'lik kısım (alan) için alt sınır olacaktır.

Örnek 18: *Bir şirket ürettiği özel kimyasal maddeleri paketleyip satışa sunmaktadır. Fakat bilinmeyen bir nedenden dolayı paketlerin içine yabancı maddelerin karıştığı fark edilmiştir. Yapılan analizlerde, herhangi bir pakete karışan maddelerin miktarı 15.3 gram ortalama ve 3.2 gram standart sapmayla normal dağıldığı bulunmuştur. Şirket herhangi bir paket seçildiğinde, karışan yabancı maddelerin miktarının %2.5'ten az oranda olmasını istemektedir. Bu durumda gerekli olan miktarın üst sınırını bulunuz.*

Çözüm: Burada X rassal değişkeni, herhangi bir pakete karışan yabancı maddelerin miktarını göstermekte ve μ = 15.3 gram ile σ = 3.2 gram parametreleriyle normal dağılmaktadır. Burada karışan yabancı maddelerin miktarının %2.5'ten az oranda olması istendiği için eğrinin sol ucundaki taralı alan dikkate alınmalıdır. Dolayısıyla, Şekil 6.54'ten de görüleceği gibi, %2.5'lik taralı alan için uygun üst sınır, aranan x miktarı olacaktır.

Şekil 6.54

Verilen 0.025 olasılık değerine uygun z ve x değerlerinin gösterimi.



Bu üst sınırın z cinsinden değeri, ortalama ile sınır arasındaki 0.5000 - 0.0250 = 0.4750 alanına uygun olarak tablodan z = -1.96 şeklinde bulunur. Bu değer ters dönüşüm formülünde yerine yazılarak, aranan üst sınır değeri,

$x = \mu + z\sigma = 15.3 + (-1.96)(3.2) = 9.028$

olarak elde edilir. Sonuç olarak herhangi bir pakette yaklaşık olarak 9 gramdan az miktarda yabancı maddelerin karışması durumunda, bu miktar, şirketin istediği %2.5'lik alanda (oranda) olacaktır.

SIRA SİZDE 3

1. Bir süper markette, herhangi bir günde ihtiyaç duyulan bozuk para miktarı ₺1520 ortalama ve ₺125 standart sapmayla normal dağılmaktadır. Bu markette herhangi bir günde ihtiyaç duyulacak olan bozuk para miktarının,
a. ₺1300 -1600 arasında
b. ₺1800'den fazla
c. ₺1200'den az
olması olasılıklarını bulunuz.
2. Bir fabrikada üretilen LCD televizyonlarının ömürleri 10 yıl ortalama ve 2.7 yıl standart sapmayla normal dağılmaktadır. Fabrika yöneticilerinin yaptığı toplantıda gündeme gelen,
a. Rassal olarak seçilen bir LCD televizyonun ömrünün en az 12 yıl olması olasılığı,
b. Üretilen LCD televizyonların yüzde kaçının ömrünün 15 yıldan az olduğu,

c. Üretilen LCD televizyonlarının en fazla %0.5'inin bozulacağı varsayıldığına göre, üretilen LCD televizyonlar için uygun garanti süresinin belirlenmesi,
sorularını yanıtlayınız.
3. Sürekli X rassal değişkeni 47.62 ortalama ve σ standart sapmayla normal dağılmaktadır. Bu bilgiler göre,
a. X rassal değişkeninin x = 56.56 değeri için z değerinin 1 olduğu bilindiğine göre, σ değerini bulunuz.
b. P (38.74 < X < 79.21) olasılık değerini bulunuz.
c. Normal dağılım eğrisi altında, ortalama ile ortalamanın sağıda kalan yaklaşık 0.2357'lık alanı belirleyen x değerini bulunuz.

BİNOM DAĞILIMINA NORMAL DAĞILIM YAKLAŞIMI

Kesikli rassal değişkenlere uygun binom dağılımıyla ilgili anlatılanlar kısaca özetlenirse,

1. Binom deneyi n tane aynı (özdeş) denemeden oluşmaktadır.
2. Her bir deneme için başarı ve başarısızlık olarak isimlendirilen yalnız iki sonuç vardır.
3. Tek bir deneme için başarı olasılığı p olup her bir deneme için aynıdır. Başarısızlık olasılığı q = 1 - p'dir.
4. Denemeler birbirinden bağımsızdır.

koşullarını sağlayan bir binom deneyinde, n denemede x tane başarılı sonuç elde edilmesi olasılığı,

$$P\,(X=x)=\binom{n}{x}p^{x}\left(1-p\right)^{n-x},\ x=0,1,\ldots,n$$

binom olasılık fonksiyonu yardımıyla bulunmaktadır ve X kesikli rassal değişkenine **binom dağılımına** sahiptir denir.

Bir binom deneyinde denemelerin sayısı (n) çok büyük olduğunda, binom olasılık fonksiyonuyla gerekli hesaplamaları yapmak zordur. Bu zorluğu aşmak için, binom olasılık fonksiyonuyla kesin olasılık değerini bulmak yerine, normal dağılımdan yararlanarak yaklaşık olasılık değerini bulmak daha uygundur. Binom dağılımına normal dağılım yaklaşımının kullanılabilmesi için, $n\,p \geq 5$ ve $n\,q \geq 5$ koşullarının sağlanması gerekir. Bu koşulların sağlanması durumunda, binom dağılımında aranan olasılık değerinin normal dağılım yardımıyla elde edilmesinde aşağıda verilen adımlar izlenir.

Adım 1: Binom dağılımının μ ortalaması ve σ standart sapması hesaplanır. Binom dağılımı için bu parametreler,

$$\mu = n\ p,$$

$$\sigma = \sqrt{n\,p\,q}\ ,$$

formülleri yardımıyla bulunur. Bu durumda, normal dağılımın kullanılabilmesi için gerekli olan μ ve σ parametre değerleri elde edilmiş olur.

Adım 2: Kesikli rassal değişkenin sürekli rassal değişkene dönüştürülmesi için **"süreklilik düzeltmesi"** yapılmalıdır. Buna göre, binom dağılımı için $P\,(X = x)$ olasılığında x'e ± 0.5 değeri eklenerek, normal dağılımı için $P\,(x - 0.5 \leq X \leq x + 0.5)$ olasılık değeri hesaplanır. Sonuç olarak, kesikli rassal değişkenlere uygulanan binom dağılımının, sürekli rassal değişkenlere uygulanan normal dağılıma yaklaşımı sağlanır. Ayrıca, binom dağılımıda aranan olasılık eşitsizlik olabileceği gibi, bir arlıkta olabilmektedir. Örneğin, binom dağılımda $P(X \leq 6)$ olasılık değeri normal dağılım yaklaşımında aranırken $P(X \leq 6.5)$ şeklinde, $P(X \geq 18)$ aranırken $P(X \geq 17.5)$ şeklinde ve $P(19 \leq X \leq 28)$ aranırken de $P(18.5 \leq X \leq 28.5)$ şeklinde süreklilik düzeltmesi yapılır.

DİKKAT

Binom dağılımda $P(X \leq 6)$ olasılık değeri normal dağılım yaklaşımında aranırken $P(X \leq 6.5)$, $P(X \geq 18)$ aranırken $P(X \geq 17.5)$ ve $P(19 \leq X \leq 28)$ aranırken de $P(18.5 \leq X \leq 28.5)$ şeklinde süreklilik düzeltmesi yapılır.

Adım 3: İlk adımda bulunan μ ve σ değerlerine sahip normal dağılım kullanılarak $P(x - 0.5 \leq X \leq x + 0.5)$ olasılık değeri bulunur. Bunun için x - 0.5 ve x + 0.5 sınır değerlerine karşılık gelen z değerleri bulunduktan sonra standart normal dağılım tablosundan yararlanılır.

Örnek 19: *Kesikli X rassal değişkeni, n = 35 ve p = 0.55 olacak şekilde binom dağılımına sahiptir.*

a) Binom olasılık fonksiyonu kullanarak P (X ≥ 23) olasılık değerini bulunuz.

b) Normal dağılım yaklaşımından yararlanarak P (X ≥ 23) olasılık değerini bulunuz ve bulunan sonuçları karşılaştırınız.

Çözüm: a. Verilen bilgilere göre, binom dağılımının parametreleri,

$n = 35$, $p = 0.55$ ve $q = 1 - p = 1 - 0.55 = 0.45$

dir. $P(X \geq 23)$ olasılık değerini bulmak için, bu değerler binom olasılık fonksiyonunda yerine yazıldığında,

$$P(X \geq 23) = P(X = 23) + P(X = 24) + ... + P(X = 35)$$

$$P(X \geq 23) = \binom{35}{23}(0.55)^{23}(0.45)^{12} + \binom{35}{24}(0.55)^{24}(0.45)^{11} + ... + \binom{35}{35}(0.55)^{35}(0.45)^{0}$$

$$P(X \geq 23) = 0.0614 + 0.0375 + ... + 0.0000 = 0.1343$$

sonucu elde edilir. Yukarıdaki işlemlerden de görüldüğü gibi, binom olasılık fonksiyonunu kullanarak $P(X \geq 23)$ olasılık değerini bulmak oldukça zordur. Bu durumda istenen olasılık değerini bulmak için normal dağılım yaklaşımı kullanmak, işlem açısından büyük kolaylık sağlayacaktır.

b. Burada $n\,p = 35(0.55) = 19.25$ ve $n\,q = 35(0.45) = 15.75$ şeklinde gerekli koşullar sağlandığı için, istenen olasılık değeri normal dağılım yaklaşımıyla bulunabilir. Buna göre yukarıda belirtilen adımlardan ilk olarak,

$$\mu = n\,p = 35\,(0.55) = 19.25,$$

$$\sigma = \sqrt{n\,p\,q} = \sqrt{35\,(0.55)\,(0.45)} = 2.9432$$

biçiminde ortalama ve standart sapma değerleri elde edilir.

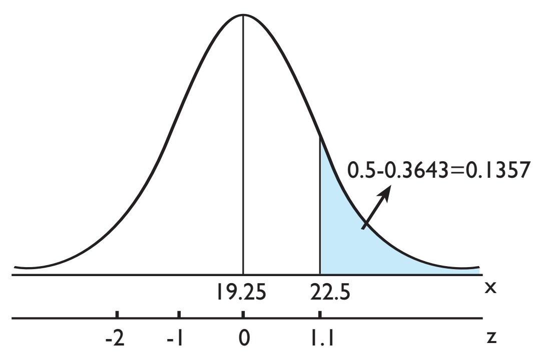
İkinci adımda, $x = 23$ değeri için, 22.5 biçiminde süreklilik düzelmesi yapılır. Bu durumda, binom dağılımında $P(X \geq 23)$ olasılığı yerine, normal dağılımda $P(X \geq 22.5)$ olasılığı bulunacaktır (Şekil 6.55).

Son adımda, Ek 1'deki tablo değerlerinden yararlanabilmek için gerekli z değeri,

x = 22.5 için z: $z = \dfrac{22.5 - 19.25}{2.9432} = 1.10$

olarak elde edilir ve istenen olasılık değeri bulunur. Bunun için, ortalamanın sağındaki tüm alan değerinden ortalama ile z = 1.10 arasındaki alan değeri çıkartılır ve

Şekil 6.55



Normal eğri altında x = 22.5'in sağındaki alan.

$P(X \geq 22.5) = P(Z \geq 1.10) = 0.5 - 0.3643 = 0.1357$

değeri bulunur. Sonuç olarak, binom olasılık fonksiyonu yardımıyla oldukça güç hesaplamalarla bulunan kesin olasılık değeri, normal dağılım yaklaşımıyla yaklaşık olarak da olsa, daha basit yolla bulunmuş olur. Ayrıca, bu iki olasılık değeri arasındaki fark (0.1343 - 0.1357 = -0.0014) ihmal edilebilecek kadar küçüktür.

Örnek 20: *Bir firma, bir arızadan dolayı ürettiği ürünlerinin %15'inin kusurlu olduğunu bilmektedir. Bu üretilen ürünler arasından 300 tanesi rassal olarak seçildiğinde,*

a. 50 veya daha az ürünün kusurlu,

b. 30 ile 40 arasında ürünün kusurlu,

olma olasılıklarını bulunuz.

Çözüm: Kusurlu ve kusursuz olarak adlandırılan iki sonuçlu binom deneyinde verilen parametreler,

$n = 300$, $p = 0.15$ ve $q = 1 - p = 1 - 0.15 = 0.85$

tir ve $np = 45$ ve $nq = 255$ olmasından dolayı istenen olasılık değerleri normal dağılım yaklaşımıyla bulunabilir.

a. Burada istenen olasılık $P(X \leq 50)$'dir. Dolayısıyla normal dağılım yaklaşımında aranacak olasılık ise $P(X \leq 50.5)$ biçiminde olur. Bu olasılığı bulmak için, ilk olarak μ ve σ değerleri,

$\mu = np = 300\,(0.15) = 45,$

$\sigma = \sqrt{npq} = \sqrt{300(0.15)(0.85)} = 6.1846$

olarak elde edilir. Sonra, x = 50.5 sınır değeri için z değeri,

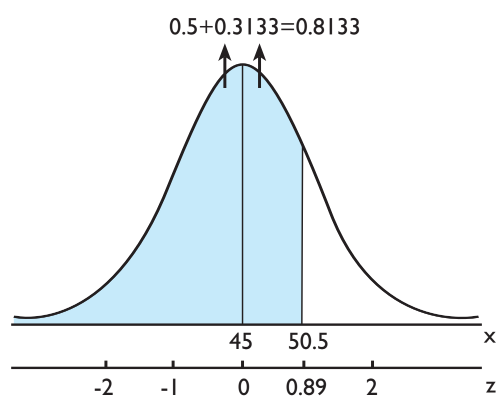
x = 50.5 için z: $z = \dfrac{50.5 - 45}{6.1846} = 0.89$

biçiminde bulunur. Şekil 6.56'dan da görüldüğü gibi, $P(X \leq 50.5)$ olasılık değeri,

$P(X \leq 50.5) = P(Z \leq 0.89) = 0.5 + 0.3133 = 0.8133$

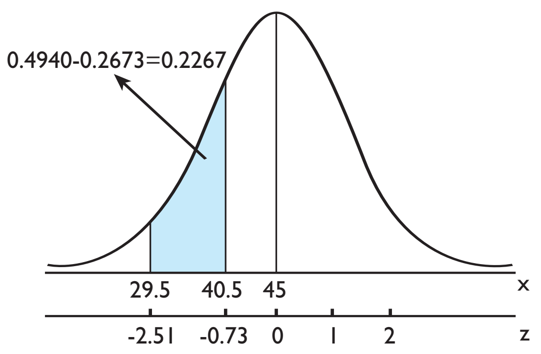
olarak elde edilir.

Şekil 6.56



$P(X \leq 50.5)$ olasılık değeri.

Şekil 6.57



P (29.5 ≤ X ≤ 40.5) olasılık değeri.

b. Bu soruda aranan $P(30 \leq X \leq 40)$ olasılığı, süreklilik düzletmesinden sonra $P(29.5 \leq X \leq 40.5)$ biçiminde ifade edilir. Bir önceki soruda hesaplanan $\mu = 45$ ve $\sigma = 6.1846$ değerlerinden yararlanıldığında, gerekli z değerleri,

$$x = 29.5 \text{ için } z: \quad z = \frac{29.5 - 45}{6.1846} = -2.51$$

$$x = 40.5 \text{ için } z: \quad z = \frac{40.5 - 45}{6.1846} = -0.73$$

olarak bulunur. Bu değerler kullanılarak, ortalamanın solunda bulunan iki alanın değeri elde edilir ve büyük alanın değerinden küçük alanın değeri çıkarılarak aranan olasılık değeri bulunur.

$P(29.5 \leq X \leq 45) = P(-2.51 \leq Z \leq 0) = 0.4940$
$P(40.5 \leq X \leq 45) = P(-0.73 \leq Z \leq 0) = 0.2673$
$P(29.5 \leq X \leq 40.5) = P(-2.51 \leq Z \leq -0.73) = 0.4940 - 0.2673 = 0.2267.$

Örnek 21: *Yapılan bir araştırmaya göre, bir A şehrinde çalışanların %61'inin işe giderken toplu taşıma araçlarını kullandığı %39'unun ise kullanılmadığı bilinmektedir. Bu gruptan rassal olarak seçilen 150 çalışandan,*

a. 75 çalışanın toplu taşıma araçlarını kullanması,

b. 85 - 105 çalışanın toplu taşıma araçlarını kullanması,

olasılıklarını bulunuz.

Çözüm: İki sonuçlu binom deneyine uyan bu soruda,

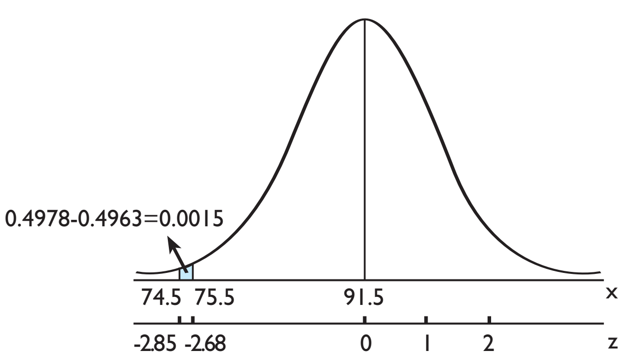
$n = 150$, $p = 0.61$ ve $q = 0.39$.

dur ve $n\,p = 91.5$ ve $n\,q = 58.5$ olması nedeniyle normal dağılım yaklaşımı kullanılabilir. Bunun için gerekli olan değerler aşağıda verilmiştir.

$$\mu = n\,p = 150\,(0.61) = 91.5,$$

$$\sigma = \sqrt{n\,p\,q} = \sqrt{150(0.61)\,(0.39)} = 5.9736.$$

Şekil 6.58



Normal eğri altında x = 74.5 ile x = 75.5 arasındaki alan.

a. Burada istenen olasılık $P(X = 75)$ dır. Ancak, bu olasılık, normal dağılım yardımında $P(74.5 \leq X \leq 75.5)$ biçiminde elde edilir ve gerekli z değerleri,

$$x = 74.5 \text{ için } z: \quad z = \frac{74.5 - 91.5}{5.9736} = -2.85$$

$$x = 75.5 \text{ için } z: \quad z = \frac{75.5 - 91.5}{5.9736} = -2.68$$

olarak bulunur. Bu değerler kullanılarak standart normal dağılım tablosundan gerekli olasılık değerleri bulunur ve sonuca ulaşılır.

$P(74.5 \leq X \leq 91.5) = P(-2.85 \leq Z \leq 0) = 0.4978$
$P(75.5 \leq X \leq 91.5) = P(-2.68 \leq Z \leq 0) = 0.4963$
$P(74.5 \leq X \leq 75.5) = P(-2.85 \leq Z \leq -2.68) = 0.4978 - 0.4963 = 0.0015.$

Burada $P(X = 75)$ olasılık değeri, binom olasılık fonksiyonu yardımıyla bulunmak istenilseydi,

$$P(X = 75) = \binom{150}{75}(0.61)^{75}(0.39)^{75} = 0.0016$$

sonucu elde edilirdi. Bulunan bu iki olasılık değeri arasındaki fark ise (0.0016 - 0.0015 = 0.0001) çok küçük olurdu.

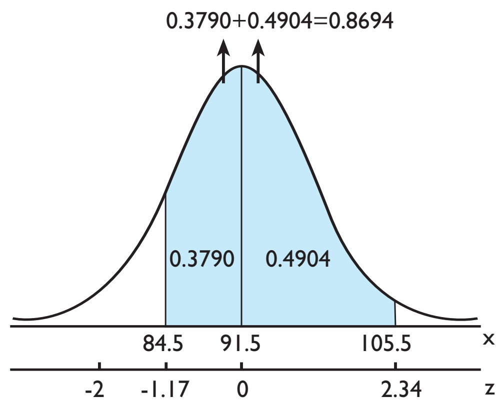
b. $P(84.5 \leq X \leq 105.5)$ olasılık değerini bulmak için gerekli z değerleri hesaplanacak olursa,

$x = 84.5$ için z: $z = \dfrac{84.5 - 91.5}{5.9736} = -1.17$

$x = 105.5$ için z: $z = \dfrac{105.5 - 91.5}{5.9736} = 2.34$

değerleri elde edilir. Bulunan z değerleri ortalamanın sağında ve solunda yer aldığı için, Ek 1'deki tablodan elde edilecek olasılık değerler toplanarak, istenen olasılık değeri bulunur.

Şekil 6.59



Normal eğri altında x = 84.5 ile x = 105.5 arasındaki alan.

$P(84.5 \leq X \leq 91.5) = P(-1.17 \leq Z \leq 0) = 0.3790$
$P(91.5 \leq X \leq 105.5) = P(0 \leq Z \leq 2.34) = 0.4904$
$P(84.5 \leq X \leq 105.5) = P(-1.17 \leq Z \leq 2.34) = 0.4904 + 0.3790 = 0.8694.$

SIRA SİZDE 4

1. Aşağıdaki soruların yanıtlarını bulunuz.
a. Binom dağılımına normal dağılım yaklaşımının kullanılabilmesi için gerekli koşulları yazınız.
b. Binom dağılımının ortalama ve standart sapma formülerini yazınız.
2. Kesikli X rassal değişkeni $n = 40$ ve $p = 0.40$ olacak şekilde binom dağılımına uymaktadır.
a. Binom olasılık fonksiyonunu kullanarak $P(18 \leq X \leq 24)$ olasılık değerini bulunuz.
b. Normal dağılım yaklaşımından yararlanarak $P(18 \leq X \leq 24)$ olasılık değerini bulunuz.
3. Bir alışveriş merkezine gelen müşterilere yapılan anket sonucunda, müşterilerin %82'sinin alışveriş merkezinde sunulan hizmetlerden memnun oldukları ve %18'inin memnun olmadıkları bulunmuştur. Bu müşterilerden rassal olarak seçilen 500 kişiden,
a. 400 veya daha az kişinin memnun olması
b. 250 veya daha fazla kişinin memnun olması
olasılıklarını bulunuz.

Özet

AMAÇ 1

Sürekli rassal değişken kavramını açıklamak.
Belli bir aralıkta her değeri alabilen rassal değişkene, sürekli rassal değişken denir. Bir başka ifadeyle, sürekli rassal değişken, alabileceği değerleri sayılamayacak kadar çok olan rassal değişkendir. Örneğin, bir hisse senedinin fiyatı, okul giderleri için yapılan harcama, bir elektronik eşyanın dayanma süresi, yağış miktarı, bir ürünün tamamlanma süresi, çocukların ağırlığı ve boy uzunluğu gibi değişkenler sürekli rassal değişkenlerdir.
Sürekli rassal değişkenlerle ilgili olasılıkların (alanların) belirlenebilmesi için $f(x)$ olasılık yoğunluk fonksiyonu kullanılır ve $f(x)$'in aşağıdaki özellikleri sağlaması gerekir.
i. Her x için $f(x) \geq 0$'dır.
ii. $f(x)$ eğrisi altında x-ekseniyle sınırlandırılmış alan veya olasılık 1'e eşittir
Sürekli X rassal değişkeninin herhangi a ve b ($a \leq b$) değerleri arasında olması olasılığı, $P(a < X < b)$ biçiminde gösterilir ve $f(x)$ eğrisi altında a ve b değerleri arasındaki alan değerine eşittir. Ayrıca, sürekli X rassal değişkeninin tek bir x değerini alması olasılığı (alanı) her zaman sıfırdır. Buna göre,

$$P(X = a) = 0 \text{ veya } P(X = b) = 0$$

olduğundan, sürekli X rassal değişkeni için

$$P(a < X < b) = P(a \leq X \leq b) =$$
$$= P(a < X \leq b) = P(a \leq X < b)$$

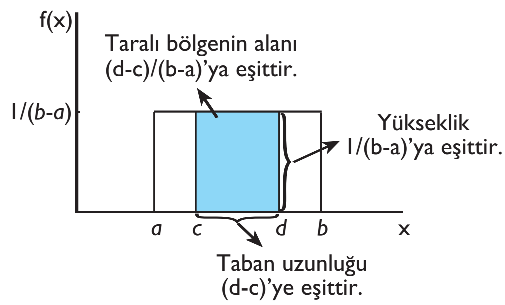
eşitliği yazılabilir.

AMAÇ 2

Düzgün dağılımı genel hatlarıyla incelemek, günlük yaşamda düzgün dağılımın yerini belirlemek ve düzgün dağılıma ilişkin olasılıkları hesaplamak.
Düzgün dağılım, sürekli bir rassal değişkenin tanımlı olduğu aralıkta belirlenen eşit uzunluktaki aralıkların olasılıklarının eşit olduğu bir dağılımdır. Örneğin; bir uçağın bir yerden başka bir yere olan uçuş süresi, belli uzunluktaki bir borunun arızalandığı noktadaki mesafesi gibi rassal değişkenler yaklaşık olarak düzgün dağılıma sahiptir.
Sürekli X rassal değişkenin olasılık yoğunluk fonksiyonu,

$$f(x) = \frac{1}{b-a}, \quad a \leq x \leq b$$

biçimde ise X rassal değişkenine düzgün dağılıma sahiptir denir ve kısaca $X \sim U(a,b)$ biçiminde gösterilir. Burada a ve b değerleri sırasıyla rassal değişkenin alabileceği değerlerin minimumunu (en küçük değeri) ve maksimumu (en büyük değeri) göstermektedir. Bu $a \leq x \leq b$ aralığının dışındaki x değerleri için $f(x)$ fonksiyonu sıfır değerini alır. $f(x)$ olasılık yoğunluk fonksiyonunun grafiği veya düzgün dağılım eğrisi belli bir $a \leq x \leq b$ aralığında düz bir doğru biçimindedir.



Şekil 6.60 *Düzgün dağılım eğrisi altında kalan alanın değeri.*

Düzgün dağılıma sahip X rassal değişkeninin c ve d ($a \leq c < d \leq b$) değerleri arasında olması olasılığı $P(c < X < d)$, düzgün dağılım eğrisi altında kalan c ve d değerleri arasındaki taralı bölgenin alanına eşittir. Bu bölgenin alanı,

Dikdörtgen alanı = (Yükseklik)(Taban uzunluğu)

formülü ile bulunur (Şekil 6.60). Diğer bir ifadeyle, taralı bölgenin alanı veya $P(c < X < d)$ olasılığı,

Taralı Bölgenin Alanı= $P(c < X < d) = (d - c)/(b - a)$

formülüyle hesaplanır. Düzgün dağılıma sahip sürekli X rassal değişkeninin μ ortalama ve σ standart sapması:

$$\mu = \frac{a+b}{2} \quad \text{ve} \quad \sigma = \frac{b-a}{\sqrt{12}}$$

eşitlikleri yardımıyla bulunur.

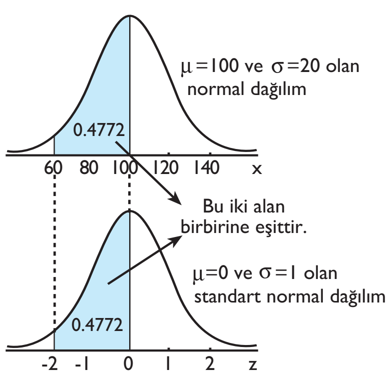
Normal dağılımı ana hatlarıyla öğrenmek, normal dağılımın gündelik hayattaki yerini belirlemek ve normal dağılımla ilgili olasılıkları hesaplamak.

Sürekli rassal değişkenler için en önemli dağılımlardan biri normal dağılımdır. Bunun nedeni, günlük yaşamımızda gözlenen sürekli rassal değişkenlerin büyük çoğunluğunun (yaklaşık olarak) normal dağılıma uymasıdır. Örneğin, bir ailenin yıllık geliri, üretilen ürünlerin ağırlıkları, zeka testi sonuçları, bebeklerin boy uzunlukları gibi rassal değişkenler yaklaşık olarak normal dağılır.

Sürekli X rassal değişkenin olasılık yoğunluk fonksiyonu,

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}, \quad -\infty < x < \infty$$

biçiminde ise, X rassal değişkenine normal dağılıma sahiptir denir ve kısaca $X \sim N(\mu, \sigma^2)$ biçiminde gösterilir. Burada μ ve σ^2 parametreleri, sırasıyla normal dağılıma sahip rassal değişkenin ortalamasını ve varyansını göstermektedir. Dolayısıyla, σ normal dağılıma sahip rassal değişkenin standart sapmasıdır. $f(x)$ olasılık yoğunluk fonksiyonun grafiği veya normal dağılım eğrisi çan eğrisi şeklindedir. $f(x)$ eğrisinin, $x = \mu$'ye göre simetrik olmasından dolayı ortalamanın sağındaki ve solundaki alanlar 0.5'e eşittir.



Şekil 6.61 *Normal eğri altındaki alanın standart normal eğri yardımıyla gösterimi.*

μ ve σ parametre değerleriyle normal dağılıma sahip X rassal değişkeninin a ve b $(a \leq b)$ değerleri arasında olması olasılığı $P(a < X < b)$, normal dağılım eğrisi altında kalan a ve b arasındaki bölgenin alanına eşittir. Bu alanın bulunması için, daha önceden hazırlanmış olan Ek1'de verilen tablodan yararlanılır. Söz konusu tablo, $\mu = 0$ ve $\sigma = 1$ olan normal dağılım (standart normal dağılım) eğrisi altında kalan belli değerler arasındaki alanları vermektedir. Ek 1'deki standart normal dağılım tablosunun değerlerinden yaralanabilmek için, normal dağılıma sahip X rassal değişkeninin herhangi bir x değeri,

$$z = \frac{x-\mu}{\sigma}$$

formülü yardımıyla z değerine dönüştürülür ve aranan olasılık veya alan değeri bulunur (Şekil 6.61).

Ek 1'deki tabloda, standart normal dağılım eğrisi altında $z = 0$ ile z'nin 0.00'dan 3.09'a kadar olan değerleri arasında kalan alan (olasılık) değerleri yer almaktadır. Normal dağılım ortalamaya göre simetrik olması nedeniyle, $z = 0$ ile pozitif z değeri ($a > 0$ için $z = a$ gibi) arasındaki alan, negatif z değeri ile $z = 0$ arasındaki alana eşittir. Dolayısıyla, Ek 1'deki tablo değerlerinden, negatif z değerleri ile $z = 0$'a kadar olan alanlar için de yararlanılabilir.

Bir diğer önemli konu ise ortalaması μ ve standart sapması σ olan normal dağılım için olasılık veya alan değeri biliniyorken uygun z ve x değerlerinin bulunmasıdır. Bu değerin bulunması için aşağıda verilen adımlar izlenmelidir.

Adım 1: Verilen olasılık değeri için uygun z değeri, Ek 1'deki standart normal dağılım tablosu yardımıyla bulunur.

Adım 2: Bulunan z değerinden x değerlerine,

$$x = \mu + z\sigma$$

formülü yardımıyla geçilir.

Binom dağılımıyla normal dağılım arasındaki bağlantıyı ilişkilendirmek.

Bir binom deneyinde denemelerin sayısı (n) çok büyük olduğunda, binom olasılık fonksiyonuyla gerekli hesaplamaları yapmak zordur. Bu zorluğu aşmak için, binom olasılık fonksiyonuyla kesin olasılık değerini bulmak yerine, normal dağılımdan yararlanarak yaklaşık olasılık değerini bulmak daha uygundur. Binom deneyinde $np \geq 5$ ve $nq \geq 5$ olması durumunda, aranan olasılık değerinin normal dağılım yardımıyla elde edilmesinde, aşağıda verilen adımlar izlenir.

Adım 1: Binom dağılımının μ ortalaması ve σ standart sapması,

$$\mu = n\,p \text{ ve } \sigma = \sqrt{n\,p\,q}$$

formülleri yardımıyla bulunur.

Adım 2: Kesikli rassal değişkenin sürekli rassal değişkene dönüştürülmesi için "süreklilik düzeltmesi" yapılmalıdır. Buna göre, binom dağılımı için $P(X = x)$ olasılığında x'e ±0.5 değeri eklenerek, normal dağılımı için $P(x - 0.5 \leq X \leq x + 0.5)$ olasılık değeri aranmış olur. Ayrıca, binom dağılımıda aranan olasılık eşitsizlik olabileceği gibi, bir arlıkta olabilmektedir. Örneğin, binom dağılımda $P(X \leq 6)$ olasılık değeri normal dağılım yaklaşımında aranırken $P(X \leq 6.5)$ şeklinde, $P(X \geq 18)$ aranırken $P(X \geq 17.5)$ şeklinde ve $P(19 \leq X \leq 28)$ aranırken de $P(18.5 \leq X \leq 28.5)$ şeklinde süreklilik düzeltmesi yapılır.

Adım 3: İlk adımda bulunan μ ve σ değerlerine sahip normal dağılım kullanılarak $P(x - 0.5 \leq X \leq x + 0.5)$ olasılık değeri bulunur. Bunun için $x - 0.5$ ve $x + 0.5$ sınır değerlerine karşılık gelen z değerleri bulunduktan sonra standart normal dağılım tablosundan yararlanılır.

Kendimizi Sınayalım

1. Sürekli rassal değişkeninin tanım aralığı $-30 \leq x \leq 45$ olduğuna göre, aşağıda X rassal değişkeniyle ilgili verilenlerden hangisi **yanlıştır?**

a. $P(-15 \leq X \leq 20) = P(-15 < X < 20)$
b. $0 \leq P(-15 \leq X \leq 20) \leq 1$
c. $P(X = 40) = 0.6$
d. $P(-30 \leq X \leq 45) = 1$
e. $P(21 < X \leq 35) = P(21 \leq X < 35)$

2. Sürekli X rassal değişkeni, 7 ile 28 değerleri arasında düzgün dağılıma sahip olsun. Bu bilgilere göre $P(6 \leq X \leq 13)$ olasılık değeri aşağıdakilerden hangisidir?

a. 0.2500
b. 0.3333
c. 0.5834
d. 0.8451
e. 0.9165

3. Halı yıkama makinesi üreten bir şirket, ürettiği makinelerin herhangi bir halının bir m^2'sini temizlemek için harcadığı deterjan miktarının 105 ml ile 128 ml arasında düzgün dağıldığını bulmuştur. Bu makinelerden yüzde kaçının herhangi bir halının bir m^2'sini temizlemek için **en az** 115 ml deterjan harcadığını bulunuz?

a. % 13.01
b. % 43.48
c. % 56.52
d. % 75.65
e. % 89.84

4. Sürekli X rassal değişkeni, 28.14 ile 54.28 değerleri arasında düzgün dağılmaktadır. X rassal değişkeninin standart sapma değeri aşağıdakilerden hangisidir?

a. 7.54
b. 8.12
c. 15.67
d. 23.79
e. 41.21

5. Standart normal dağılıma sahip Z rassal değişkeni için $P(Z \geq -0.63)$ olasılık değeri aşağıdakilerden hangisidir?

a. 0.2357
b. 0.2643
c. 0.4670
d. 0.7357
e. 0.8745

6. Sürekli X rassal değişkeni, ortalaması 18 ve **varyansı** 9 olmak üzere normal dağılmaktadır. Bu bilgiler göre $P(22.5 < X < 24.3)$ olasılık aşağıdakilerden hangisidir?

a. 0.0489
b. 0.0666
c. 0.1243
d. 0.4494
e. 0.9152

7. Bir polis radarı, bir karayolundan geçen araçların hızlarını denetlemektedir. Araçların hızları 80 km/saat ortalama ve 5 km/saat standart sapmayla normal dağılım göstermektedir. Bu bilgilere göre, karayolundan geçen herhangi bir aracın hızının 68-88 km/saat arasında olması olasılığı kaçtır?

a. 0.0548
b. 0.3196
c. 0.4918
d. 0.7725
e. 0.9370

8. Bir üniversitedeki 15000 öğrencinin aylık yaptıkları harcamalar, μ = ₺525 ve σ = ₺70 değerleriyle normal dağılım göstermektedir. Bu üniversitede kaç öğrencinin yaptığı harcama miktarının ₺340'den az olması beklenir?

a. 61
b. 550
c. 1330
d. 3150
e. 7440

9. Bir özel okula başvuran adayların kayıt yaptırabilmesi için, bu okulun yaptığı giriş sınavı sonucunda ilk %16'ya giren adayların arasında yer alması gerekmektedir. Bu okulun yaptığı giriş sınavında alınan puanlar 60 ortalama ve 10 standart sapmayla normal dağıldığı bilindiğine göre, bir adayın bu okula kayıt yaptırabilmesi için alması gereken minimum puan nedir?

a. 50.1
b. 55.9
c. 64.1
d. 69.9
e. 76.7

10. Bir hastanede yapılan araştırmada, hastaların %85'inin verilen hizmetten memnun oldukları ve %15'inin memnun olmadıkları bulunmuştur. Araştırmanın yapıldığı gruptan rassal olarak seçilen 250 hastadan 200 veya daha fazlasının verilen hizmetten memnun olması olasılığı nedir?

a. 0.0063
b. 0.4864
c. 0.5892
d. 0.7351
e. 0.9893

Yaşamın İçinden

"Uluslararası bir teknoloji şirketi, dizüstü bilgisayarlar için çeşitli parçalar üretmektedir. Bu parçalar arasında, dizüstü bilgisayarların pili (bataryası) de yer almaktadır. Bu şirket son zamanlarda yaptığı araştırmalar sonucunda yeni bir pil üretmiştir. Bu pilin sağladığı birçok avantajın yanı sıra özelikle kullanım süresinin oldukça uzun olması dikkat çekmektedir. Bu yeni pil piyasaya sürülmeden önce ilk analizleri yapılmış ve elde edilen bulgular, şirket yöneticilerine bir toplantıda sunulmuştur. Toplantıda ortaya çıkan sonuçlar, pilin kullanım süresinin piyasada ses getirecek kadar uzun olduğu yönündedir. Şirket yöneticileri bu sonuçlara bağlı olarak verilecek kararların, satışlar ve şirketin geleceği açısından oldukça önemli olduğunu düşünmektedirler. Bu nedenle, pillerin kullanım süreleriyle ilgili doğru kararların verilebilmesi için istatistiksel analizlerin yoğun bir şekilde yapılmasını ve ayrıntılı bir raporun hazırlanmasını istemişlerdir. Bunun üzerine, şirket çalışanları, gerekli istatistiksel analizleri yaptıktan sonra bazı bulgular elde etmişlerdir: Bilindiği gibi, pillerin kullanım süresi belli bir aralıkta her değeri alabildiği için sürekli bir rassal değişkendir. Yapılan analizlerde, pillerin kullanım süresinin, 400 dakika ortalama ve 50 dakika standart sapma değerleriyle, normal dağılım gösterdiği bulunmuştur. Bu normal dağılım bulgusu, pillerin kullanım süresiyle ilgili çeşitli olasılıkların hesaplanabilmesini ve verilen olasılık değerlerine karşılık, kullanım süresinin (x değerinin) belli değerlerinin bulunabilmesini sağlamıştır. Örneğin; üretilen pillerin %97.7'sinin en az 300 dakika ve %5'inin ise 530 dakikadan fazla kullanım süresine sahip olduğuna dair çeşitli sonuçlar elde edilmiştir. Elde edilen bu sonuçlar, yapılan toplantıda şirket yöneticilerine sunulmuş ve pillerin kullanım süresi hakkında çeşitli çıkarsamalar yapılabilmiş ve doğru kararlar alınabilmiştir."

Kendimizi Sınayalım Yanıt Anahtarı

1. c	Yanıtınız yanlış ise "Sürekli Rassal Değişkenler" konusunu yeniden gözden geçirin.
2. b	Yanıtınız yanlış ise "Düzgün Dağılım" konusunu yeniden gözden geçirin.
3. c	Yanıtınız yanlış ise "Düzgün Dağılım" konusunu yeniden gözden geçirin.
4. a	Yanıtınız yanlış ise "Düzgün Dağılımın Ortalaması ve Standart Sapması" konusunu yeniden gözden geçirin.
5. d	Yanıtınız yanlış ise "Normal Dağılım" konusunu yeniden gözden geçirin.
6. a	Yanıtınız yanlış ise "Normal Dağılım Uygulamaları" konusunu yeniden gözden geçirin.
7. e	Yanıtınız yanlış ise "Normal Dağılım Uygulamaları" konusunu yeniden gözden geçirin.
8. a	Yanıtınız yanlış ise "Normal Dağılım Uygulamaları" konusunu yeniden gözden geçirin.
9. d	Yanıtınız yanlış ise "Normal Dağılım İçin Olasılık Değeri Biliniyorken Uygun z ve x Değerlerinin Bulunması" konusunu yeniden gözden geçirin.
10. e	Yanıtınız yanlış ise "Binom Dağılımına Normal Dağılım Yaklaşımı" konusunu yeniden gözden geçirin.

Sıra Sizde Yanıt Anahtarı

Sıra Sizde 1

1. Belli bir aralıkta her (sayılamayacak kadar çok) değeri alabilen rassal değişkene, sürekli rassal değişken denir. Bu tanıma bağlı olarak, bir ailenin geliri ve bebeklerin boy uzunluğu sürekli rassal değişkendir. Diğer şıklardaki satılan cep telefonu sayısı ve üretilen kusurlu ürünlerin sayısı ise, (0,1,...,50 gibi) sonlu sayıda değerler alabildikleri için kesikli rassal değişkendir.

2. Sürekli bir rassal değişkenin $f(x)$ olasılık yoğunluk fonksiyonu aşağıdaki özellikleri sağlaması gerekir.

i. Her x için $f(x) \geq 0$'dır.

ii. $f(x)$ eğrisi altında kalan ve x-ekseniyle sınırlandırılmış alan veya olasılık 1'e eşittir

3. Sürekli bir X rassal değişkeninin tek bir x değerini alması olasılığı her zaman sıfırdır. Bu nedenle, a şıkkı doğrudur. Fakat b şıkkı yanlıştır. Olasılık değeri her zaman 0 ile 1 arasında olduğu için c şıkkı da yanlıştır.

Sıra Sizde 2

1. Bu soruda, düzgün dağılıma sahip X rassal değişkenin olasılık yoğunluk fonksiyonu,

$$f(x) = \frac{1}{b-a} = \frac{1}{9-4} = \frac{1}{5}, \qquad 4 \leq x \leq 9$$

biçimindedir. Bu bilgiye göre olasılık değerleri aşağıdaki gibi bulunur.

a) Burada sorulan $P(4 \leq X \leq 6)$ olasılık değeri, düzgün dağılım eğrisi altında $x = 4$ ile $x = 6$ arasında kalan alana eşittir (Şekil 6.62). Bu olasılık değeri,

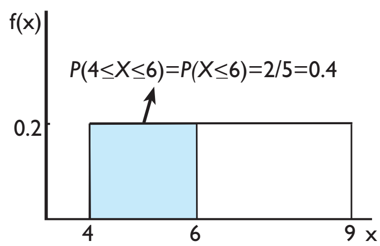
$P(c < X < d) = (d - c) / (b - a)$ formülü kullanıldığında,

$$P(4 \leq X \leq 6) = \frac{d-c}{b-a} = \frac{6-4}{9-5} = \frac{2}{5} = 0.4$$

olarak elde edilir. Bu sonuç aynı zamanda,

$$P(4 \leq X \leq 6) = P(X \leq 6) = 0.4$$

biçiminde de gösterilebilir.



Şekil 6.62 *Düzgün dağılım eğrisi altında x=4 ile x=6 arasındaki alan.*

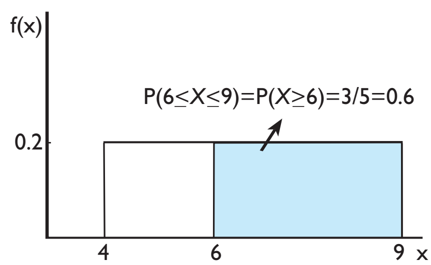
b) Bezer olarak $P(6 \leq X \leq 9)$ olasılık değeri, düzgün dağılım için olasılık formülüyle bulunur.

$$P(6 \leq X \leq 9) = P(X \geq 6) = \frac{9-6}{9-5} = \frac{3}{5} = 0.6$$
.

Ayrıca bu soruda bulunun olasılık değerlerinin toplamı, düzgün dağılım eğrisi altındaki toplam alan olan 1'e eşittir.

$P(4 \leq X \leq 9) = P(4 \leq X \leq 6) + P(6 \leq X \leq 9) = 0.4 + 0.6 = 1,$

$P(4 \leq X \leq 9) = P(X \leq 6) + P(X \geq 6) = 0.4 + 0.6 = 1.$



Şekil 6.63 *Düzgün eğilim eğrisi altında x=6 ile x=9 arasındaki alan.*

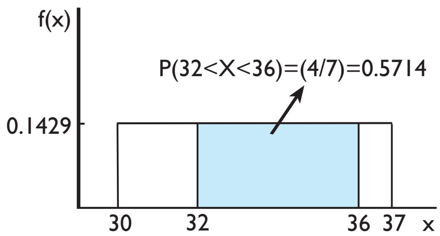
2. Bu soruda X rassal değişkeni, bir parçanın montaj süresini göstermekte ve tanım aralığında $30 \le x \le 37$ olacak şekilde düzgün dağılmaktadır. Buna göre olasılık yoğunluk fonksiyonu, aşağıdaki eşitlikle gösterilir.

$$f(x) = \frac{1}{37-30} = \frac{1}{7}, \quad 30 \le x \le 37$$

a) Soruda istenen X rassal değişkenin alabileceği değerlerin $x = 32$ ile $x = 36$ arasında olması olasılığı,

$$P(32 < X < 36) = \frac{36-32}{37-30} = \frac{4}{7} = 0.5714$$

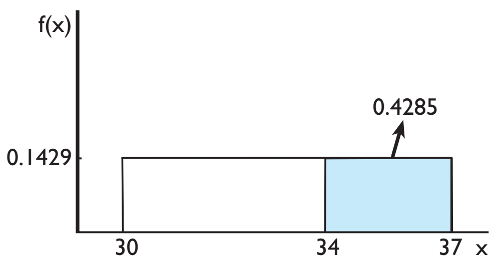
olarak elde edilir. Bir başka ifadeyle, üretim bandından rassal olarak seçilen bir parçanın montaj süresinin 32 ile 36 saniye arasında olması olasılığı 0.5714' tür.



Şekil 6.64 *P(32 < X < 36) olasılık değeri.*

b) Burada X rassal değişkenin $x = 34$'ten fazla (sonra) değerler alması veya $x = 34$'ün sağındaki değerleri alması olasılığı sorulmaktadır. Bu durumda, rassal değişkenin $30 \le x \le 37$ şeklindeki tanım aralığı dikkate alınarak, istenen olasılık değeri bulunur.

$$P(X > 34) = P(34 < X < 37) = \frac{37-34}{37-30} = \frac{3}{7} = 0.4285.$$



Şekil 6.65 *P(X ≥ 34) olasılık değeri.*

3. a) Soruda verilen $\mu = 42$ ve tanım aralığındaki ($28 \le x \le b$) $a = 28$ ve b değerleri ortalama formülünde yerine yazılarak, istenen sonuç bulunur.

$$\mu = \frac{a+b}{2} \Rightarrow 42 = \frac{28+b}{2} \Rightarrow 84 = 28 + b$$

$$\Rightarrow b = 84 - 28 = 56 \Rightarrow b = 56.$$

b) Bulunan $b = 56$ değeri ile $a = 28$ değeri standart sapma formülde yerine yazıldığında, istenen sonuç elde edilir.

$$\sigma = \frac{b-a}{\sqrt{12}} = \frac{56-28}{\sqrt{12}} = 8.0829$$

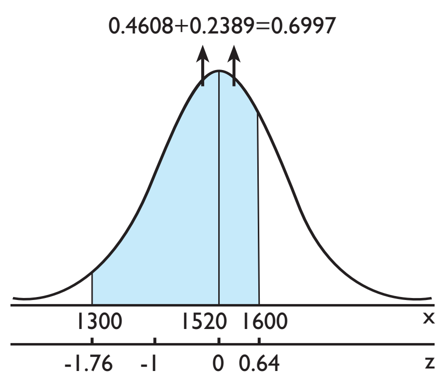
Sıra Sizde 3

1. Bu soruda X rassal değişkeni, herhangi bir günde ihtiyaç duyulan bozuk para miktarını göstermektedir. Ayrıca X, μ = ₺1520 ile σ = ₺125 değerleriyle normal dağılmaktadır.

a) Bu soruda istenen $P(1300 < X < 1600)$ olasılık değeri, standart normal dağılım tablosundan yararlanılarak bulunacaktır. Bunun için ilk olarak, verilen x değerlerinin z değerlerine dönüştürülmesi gerekir.

$$x = 1300 \text{ için } z: \ z = \frac{1300-1520}{125} = -1.76$$

$$x = 1600 \text{ için } z: \ z = \frac{1600-1520}{125} = 0.64$$



Şekil 6.66 *P(1300 < X < 1600) olasılık değeri.*

Sonra ortalamanın sağında ve solunda bulunan alan değerleri Ek 1'deki tablodan bulunur ve bu alanlar toplanarak istenen olasılık değeri elde edilir (Şekil 6.66).

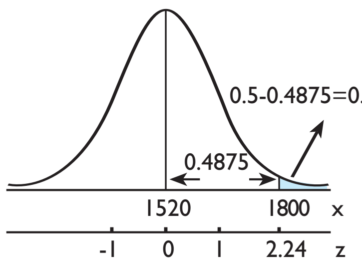
$$P(1300 < X < 1520) = P(-1.76 < Z < 0) = 0.4608$$
$$P(1520 < X < 1600) = P(0 < Z < 0.64) = 0.2389$$
$$P(1300 < X < 1600) = P(-1.76 < Z < 0.64)$$
$$= 0.4608 + 0.2389 = 0.6997.$$

b) Burada X rassal değişkeninin 1800'den fazla (büyük) değer alması olasılığıyla, $P(X > 1800)$ olasılık değeri sorulmaktadır. Bunun için gerekli z değeri,

$$x = 1800 \text{ için } z\text{: } z = \frac{1800 - 1520}{125} = 2.24$$

olarak elde edilir. Şekil 6.67'den de anlaşılacağı gibi, ortalamanın sağındaki alan (0.5) değerinden z = 0 ile z = 2.24 arasındaki alan değeri çıkarılarak, istenen olasılık bulunur.

$P(X > 1800) = P(Z > 2.24) = 0.5 - 0.4875 = 0.0125.$

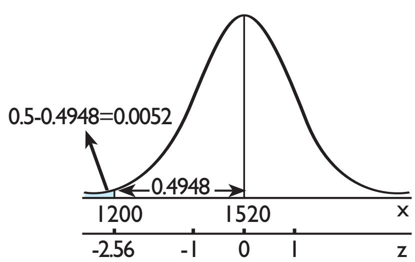


Şekil 6.67 P (X > 1800) olasılık değeri.

c) Bir önceki soruda izlenen süreç, benzer biçimde uygulanarak, $P(X < 1200)$ olasılık veya x = 1200'ün sol tarafındaki alan değeri bulunur (Şekil 6.68).

$$x = 1200 \text{ için } z\text{: } z = \frac{1200 - 1520}{125} = -2.56\ ,$$

$P(X < 1200) = P(Z < -2.56) = 0.5 - 0.4948 = 0.0052.$



Şekil 6.68 *P(X < 1200) olasılık değeri.*

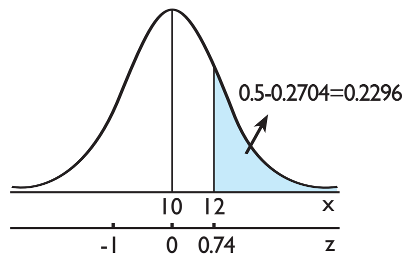
2. X rassal değişkeni, μ = 10 yıl ve σ = 2.7 yıl değerleriyle normal dağılmaktadır ve kısaca $X \sim N(10,\ 2.7^2)$ olarak ifade edilir.

a) Burada rassal olarak seçilen bir LCD televizyonun ömrünün en az 12 yıl olması olasılığıyla, $P(X > 12)$'in değeri sorulmaktadır (Şekil 6.69). Buna göre, verilen x değeri z değerine,

$$x = 12 \text{ için } z\text{: } z = \frac{12 - 10}{2.7} = 0.74$$

dönüştürdükten sonra, Ek 1'deki tablodan gerekli olasılık değeri bulunur ve aranan olasılık değeri elde edilir.

$P(X > 12) = P(Z > 0.74) = 0.5 - 0.2704 = 0.2296.$



Şekil 6.69 *P(X > 12) olasılık değeri.*

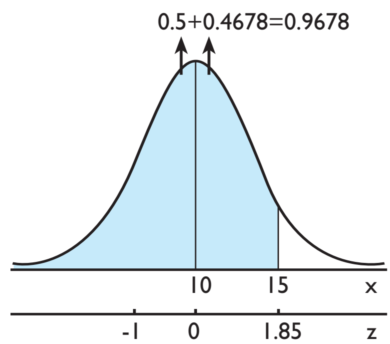
b) Bu sorunun çözümü için X rassal değişkeninin 15'ten az (küçük) değerler alması olasılığının bulunması gerekir. Bu durumda gerekli z değeri,

$$x = 15 \text{ için } z\text{: } z = \frac{15 - 10}{2.7} = 1.85$$

olarak elde edilir ve Şekil 6.70'ten de görüleceği gibi,

$P(X < 15) = P(Z < 1.85) = 0.5 + 0.4678 = 0.9678$

değeri bulunur. Sonuç olarak, üretilen LCD televizyonlarının %96.78'inin ömrünün 15 yıldan az olduğu yorumu yapılır.

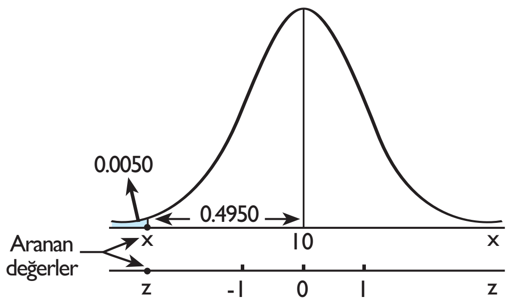


Şekil 6.70 *P(X < 15) olsılık değeri.*

c) Bu soruda, üretilen LCD televizyonları için uygulanacak garanti süresinin belirlenmesi istenmektedir. Diğer bir ifadeyle, normal dağılım için verilen olasılık değerine uygun x değeri sorulmaktadır (Şekil 6.71). Bu nedenle, ilk olarak ortalama ile aranan nokta arasındaki 0.5000 - 0.0050 = 0.4950 alanı elde edilir ve Ek 1'deki tablodan bu alan değerine karşılık gelen z = -2.58 değeri bulunur. Sonra z = -2.58 değeri ters dönüşüm formülde yerine yazılarak, aranan x değeri,

$x = \mu + z\sigma = 10 + (-2.58)(2.7) = 3.034$

elde edilir. Bu durumda, yöneticiler 3 yıldan önce arızalanan LCD televizyonları garanti kapsamına alırsa, bu süre bozulacak televizyonlar için varsayılan % 0.5'lik kısmı için üst sınır olacaktır.



Şekil 6.71 *Verilen olasılık değerine uygun z ve x değerlerinin gösterimi.*

3. Burada normal dağılımın parametreleri μ = 47.62'dir ve σ değeri ise bilinmemektedir.

a) Verilen μ = 47.62, x = 61.03 ve z = 1.5 değerleri,

$$z = \frac{x - \mu}{\sigma}$$

dönüşümde yerine yazıldığında,

$$1.5 = \frac{61.03 - 47.62}{\sigma} \Rightarrow (1.5)(\sigma) = 61.03 - 47.62$$

$$\Rightarrow \sigma = \frac{13.41}{1.5} = 8.94$$

değeri bulunur. Sorunun diğer şıkları σ = 8.94 değerine göre çözülür.

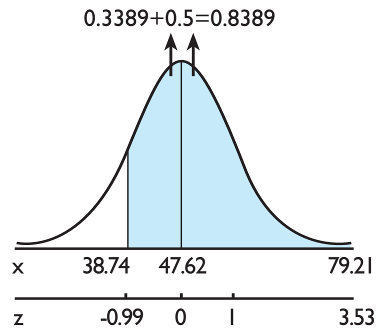
b) Verilen x = 38.74 ve x = 79.21 değerlerinin z cinsinden değeri,

$$x = 38.74 \text{ için } z\text{: } z = \frac{38.74 - 47.62}{8.94} = -0.99$$

$$x = 79.21 \text{ için } z\text{: } z = \frac{79.21 - 47.62}{8.94} = 3.53$$

olarak elde edilir. Bu değerler ortalamanın iki tarafında yer aldığı için, Ek 1'deki tablodan elde edilecek olasılık değerleri toplanarak, istenen sonuç bulunur (Şekil 6.72).

$P(38.74 < X < 47.62) = P(-0.99 < Z < 0) = 0.3389$
$P(47.62 < X < 79.21) = P(0 < Z < 3.53) = 0.5$
$P(38.74 < X < 79.21) = P(-0.99 < Z < 3.53)$
$= 0.3389 + 0.5 = 0.8389.$

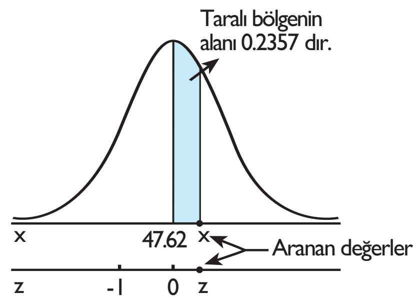


Şekil 6.72 *P(38.74<X<79.21) olasılık değeri.*

c) Verilen bilgilere göre burada aranan x değeri, P (47.62 < X < x) = 0.2357 olasılık (alan) değerine uygun olması gerekir (Şekil 6.73). Bu olasılığa uygun z değeri, standart normal dağılım tablosundan, 0.63 olarak bulunur ve z = 0.63 değeri formülde yerine yazıldığında,

$$x = \mu + z\sigma = 47.62 + (0.63)(8.94) = 53.252$$

sonucu elde edilir.



Şekil 6.73 *Verilen 0.2357 olasılık değerine uygun z ve x değerlerinin gösterimi.*

Sıra Sizde 4

1. a) Binom dağılımına normal dağılım yaklaşımının kullanılabilmesi için, $np \geq 5$ *ve* $nq \geq 5$ koşullarının sağlanması gerekir.

b) Binom dağılımının ortalaması (μ) ve standart sapması (σ), aşağıdaki formüller yardımıyla bulunur.

$$\mu = n\,p \text{ ve } \sigma = \sqrt{n\,p\,q}$$

2. a) Burada binom dağılımının parametreleri.

n = 40, p = 0.4 ve $q = 1 - p = 1 - 0.4 = 0.6$

dir ve $P(18 \leq X \leq 24)$ olasılık değerini bulmak için binom olasılık fonksiyonundan yararlanıldığında,

$$P(18 \leq X \leq 24) = 0.1025 + 0.0791 + \ldots + 0.0049 = 0.3081$$

sonucu elde edilir. Sonuç olarak binom olasılık fonksiyonunu kullanarak $P(18 \leq X \leq 24)$ olasılık değerini bulmak oldukça zahmetlidir. Bu durumda normal dağılım yaklaşımını kullanmak, işlem açısından büyük kolaylık sağlayacaktır.

b) Burada $n\,p$ = 40 (0.4) = 16 ve $n\,q$ = 40 (0.6) = 24 olmasından dolayı normal dağılım yaklaşımı kullanılabilir. Bunun için ilk olarak μ ve σ değerleri,

$$\mu = n\,p = 40\ (0.4) = 16,$$

$$\sigma = \sqrt{n\,p\,q} = \sqrt{40(0.4)\ (0.6)} = 3.0983$$

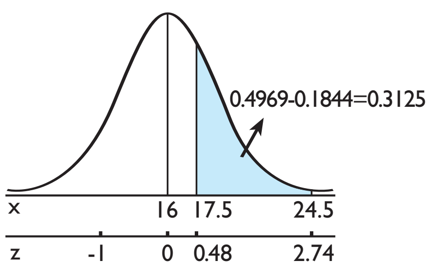
olarak bulunur ve uygun süreklilik düzeltmesiyle, normal dağılım için aranacak olasılık $P(17.5 \leq X \leq 24.5)$ biçiminde belirlenir. Daha sonra gerekli z değerleri,

$$x = 17.5 \text{ için } z\text{: } z = \frac{17.5 - 16}{3.0983} = 0.48$$

$$x = 24.5 \text{ için } z\text{: } z = \frac{24.5 - 16}{3.0983} = 2.74$$

elde edilir. Bu değerler kullanılarak, ortalamanın sağındaki iki alanın değeri bulunur ve büyük değerden küçük değer çıkartılarak istenen olasılık değeri hesaplanır.

$P(16 \leq X \leq 17.5) = P(0 \leq Z \leq 0.48) = 0.1844$
$P(16 \leq X \leq 24.5) = P(0 \leq Z \leq 2.74) = 0.4969$
$P(17.5 \leq X \leq 24.5) = P(0.48 \leq Z \leq 2.74)$
$= 0.4969 - 0.1844 = 0.3125.$



Şekil 6.74 *P (17.5 ≤ X ≤ 24.5) olasılık değeri.*

3. Binom deneyine uyan bu sorunun çözümünde,

n = 500, p = 0.82 ve q = 1 - p = 1 - 0.82 = 0.18,

dır ve $n\,p$ = 410, $n\,q$ = 90 olması nedeniyle normal dağılım yaklaşımından yararlanılır.

$$\mu = n\,p = 500\ (0.82) = 410 \text{ ve}$$

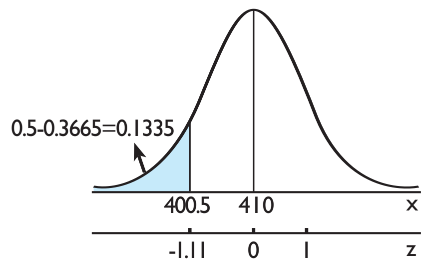
$$\sigma = \sqrt{n\,p\,q} = \sqrt{500(0.82)\ (0.18)} = 8.5906.$$

a) Burada aranan $P(X \leq 400)$ için uygun süreklilik düzeltmesi sonucunda, aranacak olan olasılık $P(X \leq 400.5)$ biçimindedir ve gerekli z değeri,

$$x = 400.5 \text{ için } z\text{: } z = \frac{400.5 - 410}{8.5906} = -1.11$$

olarak elde edilir. Bu değer kullanılarak bulunan 0.3665 alan değeri, ortalamanın solundaki tüm alan değeri olan 0.5'ten çıkarılarak sonuca ulaşılır.

$P(X \leq 400.5) = P(Z \leq -1.11) = 0.5 - 0.3665 = 0.1335.$



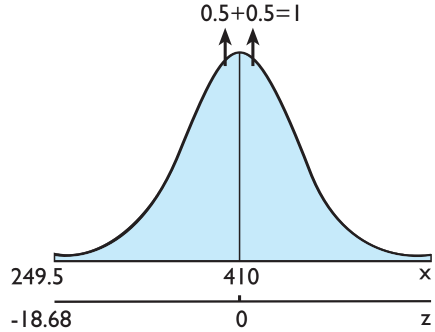
Şekil 6.75 *Normal eğri altında x= 400.5'in solundaki alan.*

b) $P(X \geq 249.5)$ olasılık değerini bulmak için gerekli z değerleri,

$$x = 400.5 \text{ için } z\text{: } z = \frac{249.5 - 410}{8.5906} = -18.68$$

biçiminde elde edilir. Şekil 6.76'dan da görüleceği gibi, bulunan z değerlerinin sağındaki tüm alan aranan olasılık değeridir.

$P(X \geq 249.5) = P(Z \geq -18.68) = 0.5 + 0.5 = 1.$



Şekil 6.76 *Normal eğri altında x = 249.5'in sağındaki alan.*

Yararlanılan ve Başvurulabilecek Kaynaklar

Newbold, P. (1995), **Statistics for Business and Economics,** Prentice Hall, New Jersey.

Yüzer, A.F., Ağaoğlu E., Tatlıdil, H., Özmen, A., Şıklar, E. (2010), **İstatistik,** Anadolu Üniversitesi Yayınları, Eskişehir.

Anderson, D.R., Sweeney, D.J., Williamas, T.A. (2009), **Statistics for Business and Economics,** 9th Edition, South-Western College, USA.

Mann P.S. (2010), **Introductory Statistics,** 7th Edition, Wiley, New York.

Akdeniz, F. (2007), **Olasılık ve İstatistik,** 13. Baskı, Nobel Yayınevi, Adana.

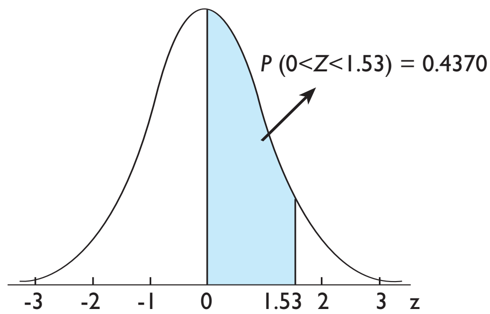
EK1: Standart Normal Dağılım Eğrisi Altındaki Alanlar Tablosu

Örnek:

$P(0 < Z < 1.53) = 0.4370$

$P(Z > 1.53) = 0.5\text{-}0.4370 = 0.0630$

$P(Z < 1.53) = 0.5+0.4370 = 0.9370$



Z	0.00	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	0.4641	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