Kaynak Kitap: TEMEL BİYOMETRİ (2013) Kocabaş Z., Özkan, MM ve Başpınar E.
Ankara Üniversitesi Ziraat Fakültesi Yayın No:1606 Ders Kitabı 558

İSTATİSTİK DAĞILIMLAR

Üzerinde çalışılan populasyonlardan tesadüfen seçilen örneklerden elde edilen değerlerin uygun analiz yöntemleriyle değerlendirildikten sonra sonuçlarının genelleştirilebilmesi ancak söz konusu populasyonların dağılımlarının şeklinin dolayısıyla fonksiyonlarının bilinmesiyle mümkündür. Bu dağılımlardan fonksiyonları bilinmekte olup, en çok kullanılanları klasik populasyonlar olarak adlandırılan **Binomiyal**, **Poisson** ve **Normal** dağılımlardır.

4.1. Binomiyal Dağılım

Binomiyal dağılım, kesikli bir dağılımdır ve bu dağılımı gösteren populasyonlarda değişkenlerin iki hali söz konusudur. Bu hallerden biri araştırıcının üzerinde durduğu (çalışılan, araştırılan, istenen) hal, diğeri ise o anda üzerinde durulmayan (çalışılmayan, araştırılmayan, istenmeyen) haldir. Binomiyal dağılım gösteren populasyonlarda istenen olayın oluş olasılığı $\boldsymbol{\pi}$ ile gösterilir. Üzerinde çalışılan değişkenin iki halinin söz konusu olduğu populasyonların dağılımlarının binomiyal dağılım olabilmesi için π'nin değişmediği, yani her bireyin istenen hali gösterme olasılığının sabit olduğu varsayılır. Değişkenin iki halinin olasılıklarının toplamı 1'e eşittir. İstenen olayın oluş olasılığı $\boldsymbol{\pi}$ ise istenmeyen olayın oluş olasılığı (**1-**$\boldsymbol{\pi}$)'dir. Eğer populasyonda istenen olayın oluş olasılığının ne olduğu bilinmiyorsa bu olasılık populasyondan seçilecek örnekten tahmin edilir. Bu durumda istenen olayın oluş olasılığı **p** ve istenmeyen olayın oluş olasılığı ise **q=(1-p)** ile gösterilir. Örneğin, doğan bebeklerin cinsiyeti, atılan bir madeni paranın yazı veya tura gelmesi, tesadüfen atılan bir zarın 5 veya farklı bir sayı gelmesi, iki renk top bulunan bir kutudan tesadüfen alınan topların rengi, binomiyal dağılım için en tipik örneklerdir.

Binomiyal dağılımın olasılık yoğunluk fonksiyonu (4.1) numaralı eşitlikte verilmiştir.

$$P(r) = C(n,r)\pi^{r}(1-\pi)^{(n-r)} \qquad \ldots(4.1)$$

(4.1) numaralı eşitlikte verilen olasılık yoğunluk fonksiyonunda, n deneme sayısını, r bu denemede istenen olayın görülme sayısını, P(r) istenen olayın n denemede r defa görülme olasılığını, C(n,r) ise n tane farklı şeyden r tanesi alınarak sıra gözetmeksizin yapılabilecek kombinasyonların (dizilişlerin) sayısını, yani "n" denemede istenen olayın "r" kere gözlenmesinin kaç farklı durumda söz konusu olduğunu gösterir. Kombinasyon sayısı **C(n,r)**, **nCr** veya $\boldsymbol{\binom{n}{r}}$ şeklinde gösterilir ve (4.2) numaralı eşitlik kullanılarak hesaplanır.

$$\binom{n}{r} = \frac{n!}{r!(n-r)!} \qquad \ldots(4.2)$$

4.1.1. İstenen Olayın Oluş Olasılığının Hesaplanması

Binomiyal dağılım gösteren populasyonlarda çalışılırken "n" denemede istenen olayın "r" kere oluş olasılığı (4.1) numaralı eşitlikte verilen binomiyal dağılımın olasılık fonksiyonu kullanılarak hesaplanabilir. Burada "n" denemeden kastedilen şudur: Eğer bir kişi bir madeni parayı 3 kere atmış ve paranın tesadüfen yazı geliş sayısı ile ilgileniyorsa deneme sayısı (n) 3'tür. Keza bir araştırıcı 4 çocuklu ailelerde kız çocuk sayısı ile ilgileniyorsa deneme sayısı (n) 4'tür. İki farklı renkte top bulunan bir kutudan eğer tesadüfen 8 tane top alınıyorsa deneme sayısı (n) 8'dir.

Eğer bir kişi madeni parayı 3 kere attığı zaman 2 keresinde yazı gelme olasılığını hesaplamak isterse, istenen olayın oluş sayısı (r) 2'dir. 4 çocuklu ailelerdeki kız çocuk sayısı ile ilgilenen bir araştırıcı, çocuklardan 3'ünün kız olma olasılığını hesaplamak isterse, istenen olayın oluş sayısı (r) 3'dür.

ÖRNEK 1:

Hilesiz bir para ile atış denemesinde, paranın yazı veya tura gelme olasılığı **π=1/2**'dir. Eğer para 3 kere atılmış ise yazı veya tura gelmesi bakımından aşağıdaki haller söz konusudur:

1. hal: 3 yazı 0 tura
2. hal: 2 yazı 1 tura
3. hal: 1 yazı 2 tura
4. hal: 0 yazı 3 tura

(4.1) numaralı eşitlikte verilen olasılık fonksiyonundan yararlanarak 4 halin olasılıkları aşağıdaki şekilde hesaplanır:

1. hal için: $P(3) = C(3,3)(\frac{1}{2})^3(1-\frac{1}{2})^{3-3} = \frac{1}{8}$=0.125 yani para 3 kere atıldığı zaman 3'ünde de yazı gelme olasılığı %12.5'tur.

2. hal için: $P(2) = C(3,2)(\frac{1}{2})^2(1-\frac{1}{2})^{3-2} = \frac{3}{8}$=0.375 yani para 3 kere atıldığı zaman 2'sinde yazı gelme olasılığı %37.5'tur.

3. hal için: $P(1) = C(3,1)(\frac{1}{2})^1(1-\frac{1}{2})^{3-1} = \frac{3}{8}$=0.375 yani para 3 kere atıldığı zaman 1'inde yazı gelme olasılığı %37.5'tur.

4. hal için: $P(0) = C(3,0)(\frac{1}{2})^0(1-\frac{1}{2})^{3-0} = \frac{1}{8}$=0.125 yani para 3 kere atıldığı zaman 3'ünde de tura (0'ında yazı) gelme olasılığı %12.5'tur.

Yukarıda (4.1) numaralı eşitlikte verilen binomiyal dağılım olasılık fonksiyonu kullanılarak bulunan olasılıklar aşağıda açıklandığı şekillerde de hesaplanabilir.

Hilesiz bir madeni para 3 kere atıldığı zaman paranın yazı gelmesi bakımından yukarıda açıklanan mümkün olan 4 hal Tablo 4.1'de verilen şekillerde gözlenir. Burada yazı ve tura gelme olasılığı eşit şansa sahip olup 1/2 'dir ve örneğin, madeni para 3 kere atıldığı zaman 3 kere yazı gelme olasılığı $[(1/2)(1/2)(1/2)] = (1/8)$'dir. Çünkü bağımsız olayların birlikte görülme olasılığı, olayların olasılıklarının çarpımına eşittir.

Hilesiz bir madeni para 3 kere atıldığı zaman, 3 keresinde de yazı gelmesi sadece bir durumda söz konusudur ve olasılığı 1/8'dir.

Madeni paranın 3 kere atılması durumunda 2 atışta yazı gelmesi, Tablo 4.1'den görüldüğü gibi 3 farklı şekilde söz konusudur. Fakat parayı atan kişi hangi atışta yazı geldiği ile değil hangi atışta olursa olsun 2 kere yazı gelmesi ile ilgilenmektedir. Bu sebeple 3 atıştan 2'sinde yazı gelme olasılığı $[(1/8) + (1/8) + (1/8)] = (3/8)$'dir.

Tablo 4.1. Bir paranın 4 kere atılması durumunda yazı gelmesi bakımından gözlenen durumlar ve olasılıkları

Atışlar			Yazı sayısı	Olasılık
1. atış	**2. atış**	**3. atış**		
Yazı	Yazı	Yazı	3	$[(1/2)(1/2)(1/2)]=(1/8)$
Yazı	Yazı	Tura	2	$[(1/2)(1/2)(1/2)]=(1/8)$
Yazı	Tura	Yazı	2	$[(1/2)(1/2)(1/2)]=(1/8)$
Tura	Yazı	Yazı	2	$[(1/2)(1/2)(1/2)]=(1/8)$
Yazı	Tura	Tura	1	$[(1/2)(1/2)(1/2)]=(1/8)$
Tura	Yazı	Tura	1	$[(1/2)(1/2)(1/2)]=(1/8)$
Tura	Tura	Yazı	1	$[(1/2)(1/2)(1/2)]=(1/8)$
Tura	Tura	Tura	0	$[(1/2)(1/2)(1/2)]=(1/8)$

Hilesiz bir madeni paranın 3 kere atılması durumunda, 1 atışta yazı gelmesi, Tablo 4.1'den görüldüğü gibi 3 farklı şekilde söz konusudur. Fakat parayı atan kişi hangi atışta yazı geldiği ile değil hangi atışta olursa olsun 1 kere yazı gelmesi ile ilgilenmektedir. Bu sebeple 3 atıştan 1'inde yazı gelme olasılığı da $[(1/8)+(1/8)+(1/8)]=(3/8)$'dir.

Hilesiz bir madeni para 3 kere atıldığı zaman, 0 keresinde de yazı (veya 3 keresinde tura) gelmesi sadece bir durumda söz konusudur ve bunun da olasılığı 1/8'dir.

Mümkün olan hallerin olasılıklarına dikkat edilirse, bu olasılıkların toplamı $[(1/8)+(3/8)+(3\,1/8)+(1/8)]=1.0$'e eşittir.

Yukarıda açıklanan olasılıklar binom açılımından yararlanılarak da hesaplanabilir. Bu durumda $[\pi+(1-\pi)]^3$ binomunun açılması gerekir. Burada $\boldsymbol{\pi=1/2}$ olup atılan hilesiz madeni paranın yazı gelme olasılığıdır. Açılan binomun terimleri yukarıda hesaplanan olasılıklara karşılık gelir:

$$[\pi+(1-\pi)]^3 = \pi^3 + 3\,\pi^2\,(1-\pi) + 3\,\pi(1-\pi)^2 + (1-\pi)^3$$

$$(1/2 + 1/2)^3 = (1/2)^3 + 3\,(1/2)^2\,(1/2) + 3\,(1/2)(1/2)^2 + (1/2)^3$$

$$(1/2 + 1/2)^3 = 1/8 + 3/8 + 3/8 + 1/8 = 1.0$$

Binom açılımındaki 1. terim, 3 atışta da yazı gelme olasılığını, 2. terim, 3 atıştan 2'sinde yazı gelme olasılığını, 3. terim, 3 atıştan 1'inde yazı gelme olasılığını ve son terim, atışların hepsinde tura gelme olasılığını verir.

(4.1) numaralı eşitlik kullanılarak hesaplanan olasılıklar, Tablo 4.1'den hesaplanan olasılıklar ile aynıdır. Bu sebeple binomiyal dağılım gösteren bir populasyon ile çalışılırken istenen olasılıklar bu yollardan her hangi biri izlenerek hesaplanabilir.

ÖRNEK 2:

Her hangi bir dersten yapılan sınav sonunda öğrencilerin %60'nın başarılı olduğu saptanmıştır. Bu dersi alan öğrencilerden tesadüfen 5 öğrenci seçilse bu öğrencilerden;

a. 5 tanesinin başarılı olma olasılığı nedir?

b. 4 tanesinin başarılı olma olasılığı nedir?

c. 3 tanesinin başarılı olma olasılığı nedir?
d. 2 tanesinin başarılı olma olasılığı nedir?
e. 1 tanesinin başarılı olma olasılığı nedir?
f. Hepsinin başarısız olma olasılığı nedir?
g. En az 2 tanesinin başarılı olma olasılığı nedir?
h. En fazla 4 tanesinin başarılı olma olasılığı nedir?

Yukarıda belirtilen hallere ait olasılıkların hesaplanabilmesi için önce istenen olayın oluş olasılığının belirlenmesi gerekir. Herhangi bir dersten yapılan sınav sonunda öğrencilerin %60'nın başarılı olduğu belirtildiğine göre, istenen tipin olasılığı (başarılı olma olasılığı), $\pi = \frac{3}{5} =$ 0.60 ve istenmeyen tipin (başarısız olma) olasılığı ise $(1-\pi) = 1 - \frac{3}{5} = \frac{2}{5}$'e eşittir. Sınava giren öğrenciler arasından tamamen tesadüfen 5 öğrenci seçildiği için, deneme sayısı (n) 5'tir. Mümkün olan hallerin sayısı deneme sayısının bir fazlası, yani 6'dır.

a. Tesadüfen seçilen 5 öğrenciden 5'inin de başarılı olma olasılığı (4.1) numaralı eşitlik kullanılarak;

$$P(5) = C(5,5)(\frac{3}{5})^5(\frac{2}{5})^{5-5} = \frac{5!}{(5-5)!5!}(\frac{3}{5})^5(\frac{2}{5})^0$$

$$= (\frac{3}{5})^5 = \frac{243}{3125} = 0.07776$$

yani, %7.776'dır.

b. Tesadüfen seçilen 5 öğrenciden 4'ünün başarılı olma olasılığı (4.1) numaralı eşitlik kullanılarak;

$$P(4) = C(5,4)(\frac{3}{5})^4(\frac{2}{5})^{5-4} = \frac{5!}{(5-4)!4!}(\frac{3}{5})^4(\frac{2}{5})^1$$

$$= 5(\frac{3}{5})^4(\frac{2}{5}) = \frac{810}{3125} = 0.2592$$

yani, %25.92'dir.

c. Tesadüfen seçilen 5 öğrenciden 3'ünün başarılı olma olasılığı (4.1) numaralı eşitlik kullanılarak;

$$P(3) = C(5,3)(\frac{3}{5})^3(\frac{2}{5})^{5-3} = \frac{5!}{(5-3)!3!}(\frac{3}{5})^3(\frac{2}{5})^2$$

$$= 10(\frac{3}{5})^3(\frac{2}{5})^2 = \frac{1080}{3125} = 0.3456$$

yani, %34.56'dır.

d. Tesadüfen seçilen 5 öğrenciden 2'sinin başarılı olma olasılığı (4.1) numaralı eşitlik kullanılarak;

$$P(2)=C(5,2)(\frac{3}{5})^2(\frac{2}{5})^{5-2}=\frac{5!}{(5-2)!2!}(\frac{3}{5})^2(\frac{2}{5})^3$$
$$=10(\frac{3}{5})^2(\frac{2}{5})^3=\frac{720}{3125}=0.2304$$

yani, %23.04'tür.

e. Tesadüfen seçilen 5 öğrenciden 1'inin başarılı olma olasılığı (4.1) numaralı eşitlik kullanılarak;

$$P(1)=C(5,1)(\frac{3}{5})^1(\frac{2}{5})^{5-1}=\frac{5!}{(5-1)!1!}(\frac{3}{5})^1(\frac{2}{5})^4$$
$$=5(\frac{3}{5})(\frac{2}{5})^4=\frac{240}{3125}=0.0768$$

yani, %7.68'dir.

f. Tesadüfen seçilen 5 öğrenciden hiçbirinin başarılı olma olasılığı (hepsinin başarısız olma olasılığı) (4.1) numaralı eşitlik kullanılarak;

$$P(0)=C(5,0)(\frac{3}{5})^0(\frac{2}{5})^{5-0}=\frac{5!}{(5-0)!0!}(\frac{3}{5})^0(\frac{2}{5})^5$$
$$=(\frac{2}{5})^5=\frac{32}{3125}=0.01024$$

yani, %1.024'tür.

Daha önce de belirtildiği gibi mümkün olan bütün hallerin olasılıklarının toplamı da 1'e eşittir.

$$P(5)+P(4)+P(3)+P(2)+P(1)+P(0)=1.0$$

$$\frac{243}{3125}+\frac{810}{3125}+\frac{1080}{3125}+\frac{720}{3125}+\frac{240}{3125}+\frac{32}{3125}=\frac{3125}{3125}=1.0$$

(4.1) numaralı eşitlikte verilen binomiyal dağılımının olasılık yoğunluk fonksiyonu kullanılarak hesaplanan bu olasılıklar, $(\pi+(1-\pi))^5$ binomu açılarak da hesaplanabilir. Bu binomun açılımındaki terimler yukarıda hesaplanan olasılıklara karşılık gelir.

g. En az 2 öğrencinin başarılı olması demek 2, 3, 4 veya 5 öğrencinin başarılı olması demektir. Bu durumda hesaplanması gereken olasılık **P(r≥2)** şeklinde gösterilir ve **P(r≥2)=P(2)+P(3)+P(4)+P(5)**'e eşittir, yani en az iki öğrencinin başarılı olma olasılığı 2, 3, 4 ve 5 öğrencinin başarılı olma olasılıklarının toplamıdır ve

$$P(r\geq2)=P(2)+P(3)+P(4)+p(5)$$

$$P(r\geq2)=\frac{720}{3125}+\frac{1080}{3125}+\frac{810}{3125}+\frac{243}{3125}=\frac{2853}{3125}=0.91296$$

veya;

$$P(r\geq 2)=1-[P(0)+P(1)]$$

$$P(r\geq 2)=1-\frac{240}{3125}+\frac{32}{3125}=1-\frac{272}{3125}=\frac{2853}{3125}=0.91296$$

olarak hesaplanır. Hesaplanan bu olasılık tesadüfen seçilen 5 öğrenciden en az 2'sinin başarılı olma olasılığının 0.91296, yani %91.296 olduğunu gösterir.

h. En fazla 3 öğrencinin başarılı olma olasılığı 0, 1, 2 ve 3 öğrencinin başarılı olma olasılıklarının toplamıdır ve;

$$P(r\leq 3)=P(0)+P(1)+P(2)+P(3)$$

$$P(r\leq 3)=\frac{32}{3125}+\frac{240}{3125}+\frac{720}{3125}+\frac{1080}{3125}=\frac{2072}{3125}=0.66304$$

veya;

$$P(r\leq 3)=1-[P(4)+P(5)]$$

$$P(r\leq 3)=1-\frac{243}{3125}+\frac{810}{3125}=1-\frac{1053}{3125}=\frac{2072}{3125}=0.66304$$

olarak hesaplanır. Hesaplanan bu olasılık tesadüfen seçilen 5 öğrenciden en fazla 3'ünün başarılı olma olasılığının 0.66304, yani %66.304 olduğunu gösterir.

4.1.2. Binomiyal Dağılımına Göre Beklenen Frekanslar

Örnek 2'de, her hangi bir dersten yapılan sınav sonunda öğrencilerin %60'nın başarılı olduğu sınava giren öğrencilerden tesadüfen seçilen 5 öğrenciden, dersten başarılı olmaları bakımından mümkün olan bütün hallerin olasılıkları Tablo 4.2'deki gibi hesaplanmıştır.

Tablo 4.2. Öğrencilerin %60'nın başarılı olduğu bir sınavdan tesadüfen seçilen 5 öğrenciden dersten başarılı olmaları bakımından mümkün olan bütün hallerin olasılıkları

Başarılı öğrenci sayısı (r)	Olasılıklar [P(r)]
0	0.01024
1	0.07680
2	0.23040
3	0.34560
4	0.25920
5	0.07776
Toplam	1.00000

Tablo 4.2'de verilen olasılıklar, sınava giren öğrencilerden tamamen tesadüfen bir kere 5 öğrencinin seçildiği durumda hesaplanan olasılıklardır. Sınava giren öğrencilerden tesadüfen 5 öğrenci seçme işlemi 250 kere tekrarlanarak, 5'er öğrencilik 250 örnekte başarılı öğrenci sayısı bakımından dağılım Tablo 4.3'teki gibi gözlenmiş olabilir. Bu durumda, sınavda başarılı olma olasılığı %60 olarak biliniyorsa binomiyal dağılıma göre beklenen frekanslar hesaplanabilir. π=0.60 olan binomiyal

dağılımına göre beklenen frekanslar şu şekilde hesaplanır: Tesadüfen seçilen 5 öğrencilik bir örnekte 5 öğrenciden 0 tanesinin başarılı olma (yani hepsinin başarısız olma olasılığı) 0.01024 ise 250 kere 5 öğrenci seçme işlemi tekrarlanırsa 250(0.01024)=2.56 tanesinde 5 öğrenciden 0 tanesinin başarılı olması (yani hepsinin başarısız olması) beklenir. Bu şekilde, tesadüfen seçilen 5 öğrencilik bir örnekte, örneğin 5 öğrenciden 3 tanesinin başarılı olma olasılığı 0.3456 ise 250 kere 5 öğrenci seçme işlemi tekrarlanırsa 250(0.3456)=86.4 tanesinde 5 öğrenciden 3 tanesinin başarılı olması beklenir. Açıklandığı şekilde diğer sınıflar içinde hesaplanan beklenen frekanslar Tablo 4.3'te verilmiştir. Tablo 4.3'ten de görüldüğü gibi beklenen frekansların toplamı da gözlenen frekansların toplamına eşittir, yani 250'dir.

TABLO 4.3. 5'er öğrencilik 250 örnekte başarılı öğrenci sayısı bakımından gözlenen ve π=0.60 olan binomiyal dağılımına göre beklenen frekanslar

Başarılı öğrenci sayısı (r)	Gözlenen frekans (f)	Binomiyal dağılımına göre beklenen frekans (f')
0	4	250 x 0.01024=2.56
1	23	250 x 0.07680=19.20
2	54	250 x 0.23040=57.60
3	91	250 x 0.34560=86.40
4	61	250 x 0.25920=64.80
5	17	250 x 0.07776=19.44
Toplam	250	250

4.1.3. Binomiyal Dağılımının Parametreleri

Binomiyal dağılımının parametreleri "**n**" ve "**π**"'dir, yani binomiyal dağılımlar, deneme sayısı ve istenen olayın olasılığındaki farklılıklar ile birbirlerinden ayrılırlar.

Binomiyal dağılımın ortalaması, $\mu=n\pi$ ve varyansı, $\sigma^2=n\pi(1-\pi)$'dir. Öğrencilerin %60'ının başarılı olduğu bir sınavdan tesadüfen 5 öğrencinin her hangi bir dersten başarılı olmaları bakımından mümkün olan hallerin olasılıklarının Tablo 4.2'de ve π=0.60 olan binomiyal dağılımına göre beklenen frekansların da Tablo 4.3'te verildiği örneğin ortalaması $\mu=5(\frac{3}{5})=3$'tür. Binomiyal dağılımın ortalaması, örneğin 250 kere içlerinde 5'er öğrencinin bulunduğu tesadüf örnekleri oluşturulsa, başarılı öğrenci sayısı bakımından ortalamanın 3 olduğunu, yani ortalama 3 öğrencinin başarılı olacağını gösterir. Bu ortalama, Tablo 4.3'te verilen beklenen frekanslar kullanılarak aşağıdaki şekilde de hesaplanabilir, çünkü ortalama da beklenen değer olup bir parametredir ve beklenen frekansların ortalamasına eşittir.

$$\mu=\frac{0x2.56+1x19.2+2x57.6+3x86.4+4x64.8+5x19.44}{250}=\frac{750}{250}=3$$

Eğer ortalama Tablo 4.3'te verilen gözlenen frekanslar kullanılarak hesaplanırsa bir istatistiktir ve:

$$\overline{X} = \frac{(0)(4)+(1)(23)+(2)(54)+(3)(91)+(4)(61)+(5)(17)}{250} = \frac{733}{250} = 2.932$$

olarak hesaplanır. Populasyon ortalaması olan μ'ünün bir tahminidir.

Her hangi bir dersten başarı oranının %60 olduğu populasyondan, 250 kere 5 öğrencilik tesadüf örnekleri oluşturulsa, her örnekte başarılı öğrenci sayısı aynı olmayacak ve birbirinden farklılık gösterecektir. Başarılı öğrenci sayısı bakımından değişimin ölçüsü varyanstır ve yukarıdaki ifadeye göre:

$$\sigma^2 = 5(\frac{3}{5})(1-\frac{3}{5}) = 1.2$$

olarak hesaplanır.

4.1.4. Binomiyal Dağılımının Şekli

Binomiyal dağılımların şekli "**n**" ve "**π**"'ye göre değişir. Üzerinde çalışılan binomiyal dağılımda π=1/2 ise deneme sayısı ne olursa olsun dağılımın şekli simetriktir. Eğer π > 1/2 ise binomiyal dağılım sol tarafa yatıktır. π < 1/2 olması durumunda ise dağılım sağ tarafa yatıktır. π > 1/2 veya π < 1/2 olması durumlarında deneme sayısı (n) arttıkça dağılımın şekli giderek simetrikleşir.

ÖRNEK:

İstenen olayın oluş olasılığına bağlı olarak binomiyal dağılımın şekli, aşağıda ifade edilen 3 örnek üzerinde gösterilerek Tablo 4.4'te özetlenmiştir.

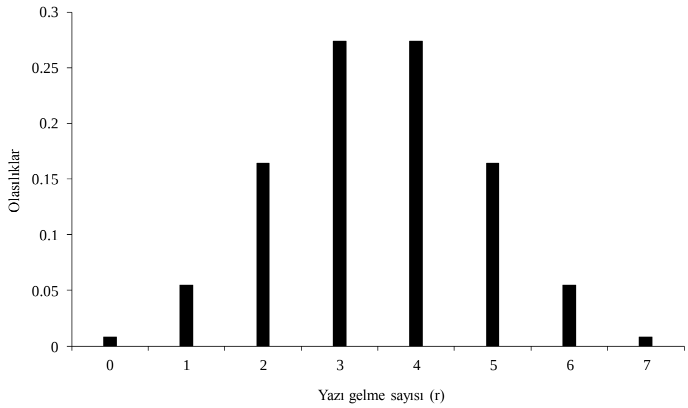
1. Hilesiz bir madeni para atıldığı zaman paranın yazı gelme olasılığı, π=1/2'dir. Hilesiz bir madeni para 7 kere atılmış ve 7 atışta paranın yazı gelmesi bakımından mümkün olan hallerin olasılıkları,

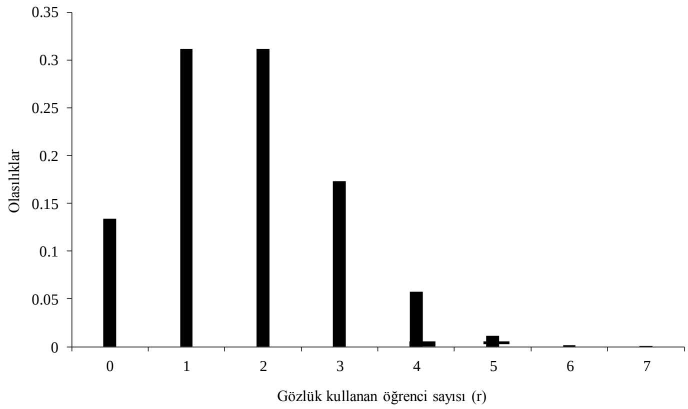
2. Öğrencilerin gözlük kullanma oranının, π=1/4 olduğu bir fakülteden tesadüfen seçilen 7 öğrencinin gözlük kullanması bakımından mümkün olan hallerin olasılıkları,

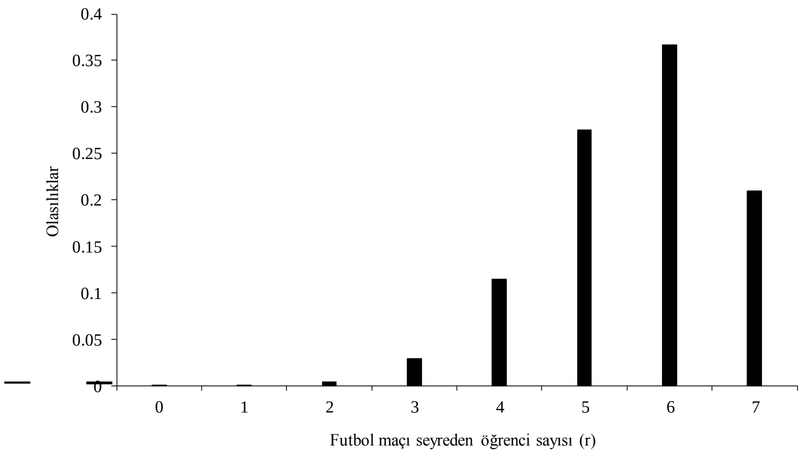
3. Erkek öğrencilerin futbol maçlarını takip etme olasılığının π=4/5 olduğu bir fakülteden tesadüfen seçilen 7 erkek öğrencinin futbol maçlarını seyretme alışkanlığı bakımından mümkün olan hallerin olasılıkları, Tablo 4.4'te verilmiştir. Bu olasılıklar doğrultusundaki binomiyal dağılımların dağılım şekilleri ise Şekil 4.1'de gösterilmiştir.

Tablo 4.4. n=7 için, paranın yazı gelmesi, öğrencinin gözlük kullanması ve öğrencinin futbol maçı seyretmesi bakımından mümkün olan bütün hallerin olasılıkları

r	P(r=paranın yazı gelmesi) (π=1/2)	P(r=öğrencinin gözlük kullanması) (π=1/4)	P(r=öğrencinin futbol maçı seyretmesi) (π=4/5)
0	0.007813	0.133484	0.000013
1	0.054688	0.311462	0.000358
2	0.164063	0.311462	0.004301
3	0.273438	0.173035	0.028672
4	0.273438	0.057678	0.114688
5	0.164063	0.011536	0.275251
6	0.054688	0.001282	0.367002
7	0.007813	0.000061	0.209715







Şekil 4.1. İstenen olayın oluş olasılığına bağlı olarak binomiyal dağılımların şekli

4.2. Poisson Dağılımı

Poisson dağılımı da binomiyal dağılım gibi kesikli bir dağılımdır. Bu dağılımda da iki tip olay vardır. Fakat bu dağılımda üzerinde çalışılan olay çok seyrek rastlanan, yani oluş olasılığı sabit fakat çok küçük olan olayların dağılımıdır. Örneğin, bir bölgede 115 yaşına kadar yaşayanların dağılımı, dünyada bir günde başlayan savaşların sayısı, bir hayat sigortası tarafından sigortalanan insanlardan belirli bir zaman aralığında ölenlerin sayısı ve bir kitaptaki yazım hatalarının sayısı gibi.

Araştırıcının üzerinde çalıştığı olayın (istenen tip olayın) oluş olasılığı küçüldükçe ve deneme sayısı büyüdükçe istenen olayın belirli sayıda olma olasılığı Poisson dağılımı olasılık yoğunluk fonksiyonundan yararlanılarak bulunabilir.

4.2.1. Poisson Dağılımının Parametresi ve Dağılım Şekli

(4.3) numaralı eşitlikte verilen Poisson dağılımının olasılık yoğunluk fonksiyonu, Fransız matematikçisi Poisson tarafından geliştirilmiş ve bu sebeple dağılıma bu matematikçinin adı verilmiştir.

$$P(r)=\frac{\mu^{r}}{r!}e^{-\mu} \qquad \ldots(4.3)$$

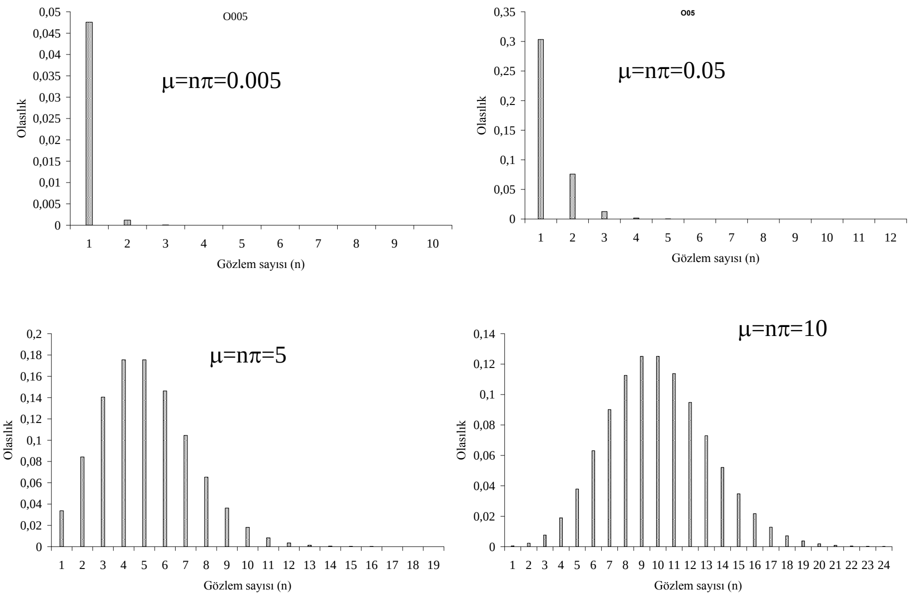
(4.3) numaralı eşitlikte, P(r): Görülme olasılığı çok düşük olayın r kere görülme olasılığını, μ: Dağılımın ortalamasını, r: Nadir olarak görülen olayın meydana geliş sayısını, e: doğal logaritma tabanını (e≅2.718) gösterir.

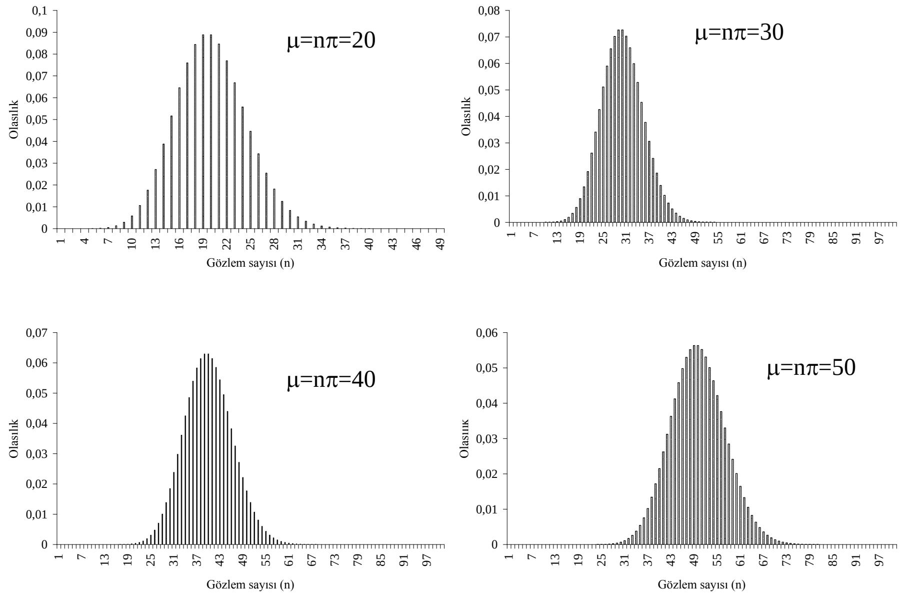
Poisson dağılımları (4.3) numaralı eşitlikte verilen olasılık yoğunluk fonksiyonundan da görülebileceği gibi birbirlerinden ortalamaları ile ayrılır, yani bu dağılımın tek parametresi dağılımın ortalaması (μ) olup, her Poisson dağılımında ortalama ile varyans birbirine eşittir ve aşağıdaki şekilde hesaplanır:

$$\mu=\sigma^2=n.\ \pi$$

eşitlikte, n birey (deneme) sayısı, π ise nadir görünen olayın oluş olasılığıdır.

Poisson dağılımının şekli ortalamaya bağlıdır. Dağılımın ortalaması büyüdükçe dağılımın şekli giderek simetrikleşir. Farklı ortalamalara sahip Poisson dağılımları Şekil 4.2'de verilmiştir.





Şekil 4.2. Farklı ortalamalara sahip Poisson dağılımları

Şekil 4.2'de görüldüğü gibi Poisson dağılımının ortalaması büyüdükçe dağılımın şekli giderek simetrikleşmektedir.

4.2.2. Olayların Oluş Olasılıklarının Hesaplanması

İstenen olayın belirli sayıda görülme olasılığı (4.3) numaralı eşitlikte verilen Poisson dağılımı olasılık fonksiyonu kullanılarak hesaplanır.

İstenen olayın herhangi bir sayıda görülme olasılığı hesaplandıktan sonra (4.4) numaralı eşitlik kullanılarak birbirini izleyen sayıda görülme olasılıkları hesaplanabilir.

Araştırıcı istenen olayın görülme olasılığını (4.3) numaralı eşitliği kullanarak $P(0)=\frac{\mu^0}{0!}e^{-\mu}=e^{-\mu}$ şeklinde hesaplar. P(0) bulunduktan sonra P(1) ise $P(1)=\frac{\mu^1}{1!}e^{-\mu}=\mu e^{-\mu}$, yani $P(1)=\mu P(0)$ şeklinde hesaplanabilir. Hesaplanan olasılıklar arasındaki söz konusu ilişki;

$$P(2)=\frac{\mu^2}{2!}e^{-\mu}=\frac{\mu}{2}\frac{\mu}{1}e^{-\mu}=\frac{\mu}{2}P(1)$$

$$P(3)=\frac{\mu^3}{3!}e^{-\mu}=\frac{\mu}{3}\frac{\mu}{2!}e^{-\mu}=\frac{\mu}{3}P(2)$$

şeklinde devam eder. Bu ilişki (4.4) numaralı eşitlikte görüldüğü şekilde özetlenebilir.

$$P(n)=\frac{\mu^n}{n!}e^{-\mu}=\frac{\mu}{n}P(n-1) \qquad \ldots(4.4)$$

ÖRNEK 1:

Belirli bir hastalıktan ölüm oranının 1/10000 olduğu bilinmektedir. Tesadüfen seçilen 50000 hastadan;

a. İkisinin ölme olasılığı ne kadardır?

b. En fazla ikisinin ölme olasılığı ne kadardır?

c. En az birinin ölme olasılığı ne kadardır?

(4.3) numaralı eşitlikte verilen yoğunluk fonksiyonu kullanılarak istenen olasılıkların hesaplanabilmesi için önce Poisson dağılımının ortalamasının $\mu = 50000x\frac{1}{10000} = 5$ şeklinde hesaplanması gerekir.

a. Tesadüfen seçilen 50000 hastadan 2'sinin söz konusu hastalıktan ölme olasılığı (4.3) numaralı eşitlik kullanılarak;

$P(2) = \frac{\mu^2}{2!} e^{-\mu} = \frac{5^2}{2!} e^{-5} = \frac{25}{2}(2.718)^{-5} \cong 0.0842$ olarak bulunur, yani tesadüfen seçilen 50000 hastadan 2'sinin söz konusu hastalıktan ölme olasılığı %8.42'dir.

b. Tesadüfen seçilen 50000 hastadan en fazla 2'sinin söz konusu hastalıktan ölme olasılığı, hiç birinin ölme (P(0)), 1'inin ölme (P(1)) ve 2'sinin (P(2)) ölme olasılıklarının toplamıdır, yani **P(n≤2)=P(0)+P(1)+P(2)** dir. Bu durumda istenen olasılık;

$P(n \le 2) = \frac{5^0}{0!} e^{-5} + \frac{5^1}{1!} e^{-5} + \frac{5^2}{2!} e^{-5} = 0.00674 + 0.0337 + 0.0842 \cong 0.1246$ olarak hesaplanır, yani tesadüfen seçilen 50000 hastadan en fazla 2'sinin söz konusu hastalıktan ölme olasılığı %12.46'dır.

c. Tesadüfen seçilen 50000 hastadan en az 1'inin söz konusu hastalıktan ölme olasılığı P(n≥1)'dir ve P(n≥1)=1-P(0) şeklinde hesaplanır. Bir önceki şıkta hiç birinin ölmeme olasılığı, P(0)=0.00674 olarak hesaplanmıştı. Bu durumda, P(n≥1)=1-P(0)=1-0.00674 ≅ 0.9933 olarak bulunur, yani tesadüfen seçilen 50000 hastadan en az 1'inin söz konusu hastalıktan ölme olasılığı %99.33'dür.

ÖRNEK 2:

500 sayfalık bir kitabın hazırlanması sırasında 125 yazım hatası yapılmış ise sayfa başına ortalama hata μ=125/500=1/4'dür. Bir sayfada hiç hata yapılmamış olma ihtimali;

$$P(0) = \frac{0.25^0}{0!} e^{-0.25} \cong 0.779$$

olarak hesaplanır. Bu durumda 500 sayfalık bu kitabın 500(0.779)≅390 sayfasında hiç hata yapılmamış olması beklenir.

500 sayfalık bu kitabın 1 sayfasında hata yapılmış olma olasılığı (4.4) numaralı eşitlik kullanılarak;

$$P(1) = \frac{0.25^1}{1!} e^{-0.25} = \frac{0.25}{1} P(0) = 0.25P(0) \cong 0.1948$$

olarak bulunur. 500 sayfalık bu kitapta 1 sayfada 1 hata yapılmış olma olasılığı %19.48'dir.

4.2.3. Poisson Dağılımına Göre Beklenen Frekansların Hesaplanması

Sebze yetiştiricileri için hazırlanan ve içinde yaklaşık olarak 500 biber tohumu bulunan ambalajlardan 250 adet rastgele alınmış ve bu ambalajlardaki tohumlardan ölü veya canlı olanların sayısı belirlenmiştir. Bu ambalajlardaki tohumlar ya canlıdır ve ekildiği zaman çimlenir ya da ölüdür ve ekildiği zaman çimlenmez. Ambalaj içindeki tohumun ekildiği zaman çimlenip, çimlenmeme bakımından iki durumu söz konusudur. Fakat hazırlanan ambalajlardaki ölü tohum sayısı çok az ve ambalajdaki tohum sayısı fazladır. Ölü tohum sayısının binomiyal dağılım gösterdiği düşünülebilir. Fakat söz konusu tohumların ölü olma olasılığı düşük ve tohum sayısı fazla olduğu için Poisson dağılımı kullanılarak, istenen olasılıklar hesaplanabilir.

250 biber tohumu ambalajında ölü tohum sayıları belirlenmiş ve ölü tohum sayısına ait dağılım Tablo 4.5'te verilmiştir.

Tablo 4.5. İçlerinde yaklaşık olarak 500 adet biber tohumu bulunan 250 ambalajın ölü tohum sayısı bakımından dağılımı

Ölü tohum sayısı	f_i	f_iX_i	$f_iX^2_i$
0	31	0	0
1	50	50	50
2	56	112	224
3	44	132	396
4	38	152	608
5	20	100	500
6	5	30	180
7	4	28	196
8	2	16	128
Toplam	250	620	2282

Poisson dağılımında teorik olarak ortalama ve varyans birbirine eşittir. Fakat bir örnek üzerinde çalışırken ortalama ve varyans birbirine eşit olmayabilir. Ortalama ve varyansın birbirine eşit olması durumu sonsuz sayıda örnek üzerinde çalışıldığı zaman söz konusudur. Tablo 4.5 verilen bilgiler kullanılarak ölü tohum sayısına ait ortalama ve varyans (2.2) ve (2.16) numaralı eşitlikler kullanılarak aşağıdaki gibi hesaplanmıştır.

$$\overline{X} = \frac{620}{250} = 2.48$$

$$\sum d_x^2 = 2282 - \frac{(620)^2}{250} = 744.4$$

$$S_x^2 = \frac{744.4}{249} = 2.99$$

Örnekten hesaplanan ortalama ve varyans birbirine yakındır. İfade ettiği olayın niteliği de göz önüne alındığında örneğin, ortalaması 2.48 olan bir Poisson dağılımını temsil ettiği kabul edilebilir. 500 tohumluk ambalajda 0-8 ölü tohum olasılığı (4.3) numaralı eşitlikte verilen Poisson dağılımı olasılık yoğunluk fonksiyonu kullanılarak hesaplanmış ve Tablo 4.6'da verilmiştir.

Tablo 4.6. İçlerinde yaklaşık olarak 500 biber tohumu bulunan 250 ambalaj için gözlenen ve Poisson dağılımına göre beklenen frekanslar

Ölü tohum sayısı	f_i	Olasılık	f_i'
0	31	0.08374	20.94
1	50	0.20768	51.92
2	56	0.25753	64.38
3	44	0.21289	53.22

4	38	0.13199	33.00
5	20	0.06547	16.37
6	5	0.02706	6.76
7	4	0.00959	2.40
8	2	0.00405	1.01
Toplam	250	1.00000	250

500 tohumun bulunduğu bir ambalajdaki tüm tohumların canlı (yani ölü tohum olmama) olasılığı (4.3) numaralı eşitlik kullanılarak;

$$P(0) = \frac{2.48^0}{0!} e^{-2.48} = 0.08374$$

olarak bulunur. Bir ambalajdaki tüm tohumların canlı olma olasılığı 0.08374 ise 250 ambalajdan %8.374 tanesindeki tüm tohumların canlı olması beklenir. Bu durumda birinci sınıfın beklenen frekansı; $f_1^{'} = (250)(0.08374) = 20.94$ olarak bulunur. Diğer sınıfların da beklenen frekansları da açıklandığı şekilde bulunmuş ve Tablo 4.6'da verilmiştir. Tablo 4.6'da son sınıfın beklenen frekansı önceki sınıfların beklenen frekanslarının toplamı 250'den çıkarılarak, yani;

$$f^{'}_{sonsınıf} = 250 - (20.94 + 51.92 + 64.38 + 53.22 + 33.00 + 16.37 + 6.76 + 2.40) = 1.01$$

olarak bulunur.

4.3. Normal Dağılım

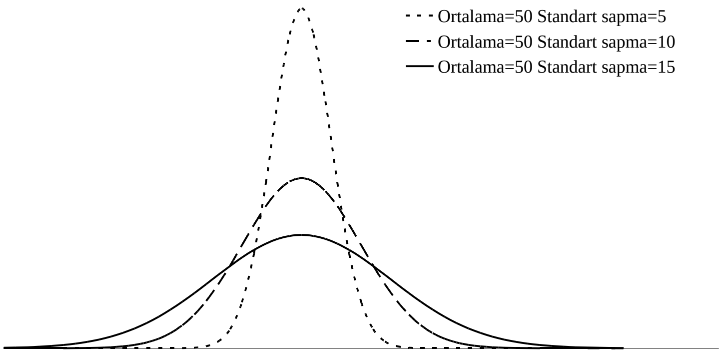
Normal dağılım sürekli bir dağılımdır. Aralarındaki farklılıklar tamamen tesadüften ileri gelen gözlem değerlerinin oluşturduğu populasyonlar normal dağılıma uygun dağılım gösterirler. Örneğin, bir ağılda aynı ırktan ve yaştan analardan yaklaşık olarak aynı tarihlerde doğan aynı cinsiyetteki kuzular aynı koşullar altında beslense bile belirli bir süre sonunda hepsinin canlı ağırlıkları birbirinin aynı olmayacaktır. Canlı ağırlıklar arasındaki bu farklılıklar, her kuzuya etkisi rastgele ve küçük olan çok sayıda etkenin (faktörün) üst üste eklenmesiyle oluşur. Bu etkenler araştırıcının kontrol edemediği faktörlerdir. Kontrol edilemeyen faktörlerden ileri gelen farklılıklar tesadüften ileri gelen farklılıklar olarak adlandırılır. Aralarındaki farklılık kontrol edilemeyen faktörlerden ileri gelen gözlem değerleri bir dağılım gösterir. Bu sebeple bir ağılda aynı ırktan ve yaştan analardan yaklaşık olarak aynı tarihlerde doğan aynı cinsiyetteki kuzular aynı koşullar altında beslense, belirli bir süre sonunda canlı ağırlıkları normal dağılım gösterir. Dolayısıyla söz konusu kuzular canlı ağırlık bakımından bir normal populasyonu temsil eder.

Aralarındaki farklılık tamamen tesadüften ileri gelen özelliklerin temsil ettiği normal populasyonların dağılımı (4.5) numaralı eşitlikte verilen olasılık yoğunluk fonksiyonuna uygun bir dağılım gösterir.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad \ldots(4.5)$$

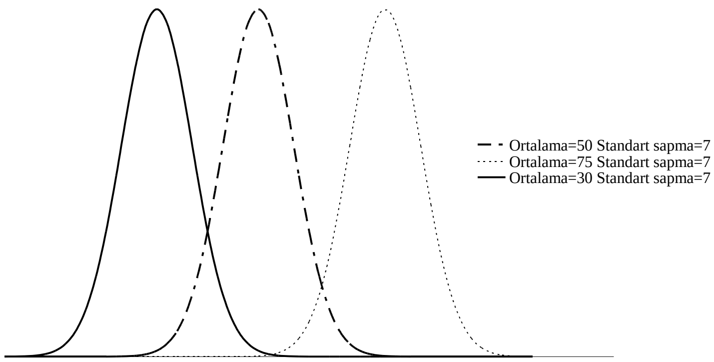
(4.5) numaralı eşitlikte verilen olasılık yoğunluk fonksiyonunda, μ, populasyonun ortalaması ve σ, populasyonun standart sapmasıdır. Fonksiyondaki e ve π sabit değerler olup yaklaşık olarak e≅2.718 ve π≅3.1416 alınabilir.

Normal dağılımın olasılık yoğunluk fonksiyonundan da görüldüğü gibi bu dağılımın ortalama ve varyans (veya standart sapma) olmak üzere iki parametresi vardır, yani normal dağılımlar birbirlerinden ortalama ve varyanslarındaki farklılıklar ile ayrılırlar. Şekil 4.3'te ortalamaları aynı fakat standart sapmaları farklı olan normal dağılımlar görülmektedir.



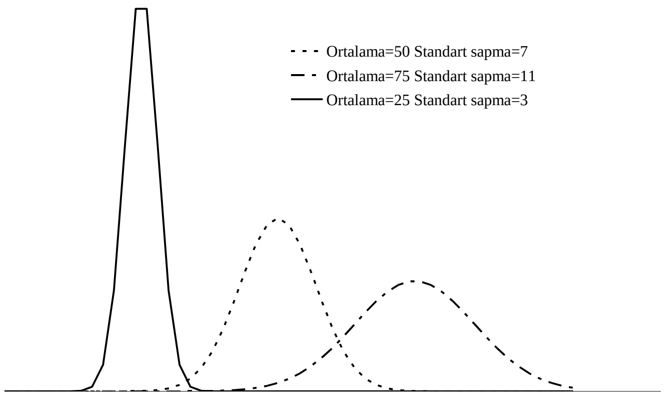
Şekil 4.3. Ortalamaları aynı standart sapmaları farklı olan normal dağılımlar

Normal dağılımlar, sadece ortalamalarının birbirinden farklı olması ile de ayrılabilirler, yani dağılımlar aynı standart sapmaya fakat farklı ortalamaya sahip olabilirler. Şekil 4.4'te standart sapmaları aynı fakat ortalamaları farklı normal dağılımlar gösterilmiştir.



Şekil 4.4. Ortalamaları farklı standart sapmaları aynı olan normal dağılımlar

Normal dağılımlar sadece bir parametrelerindeki farklılık bakımından birbirlerinden ayrılabilecekleri gibi hem ortalamalarındaki hem de standart sapmalarındaki farklılıklar bakımından da birbirlerinden ayrılabilirler. Şekil 4.5 ortalama ve standart sapmaları birbirinden farklı normal dağılımları göstermektedir.



ŞEKİL 4.5. Hem ortalaması hem de standart sapması birbirinden farklı olan normal dağılımlar

4.3.1. Normal Dağılımın Özellikleri

Aralarındaki farklılıklar tesadüften ileri gelen gözlem değerlerinin gösterdiği normal dağılım, Şekil 4.3, 4.4 ve 4.5'te görüldüğü gibi çan eğrisi şeklindedir.

Normal dağılımda aritmetik ortalama, ortanca değer ve tepe değeri birbirine eşittir. Aritmetik ortalama ve ortanca değerlerin birbirine eşit olmasından dolayı normal dağılım gösteren populasyondaki bireylerin %50'si ortalamadan büyük diğer %50'si ise ortalamadan küçüktür.

Normal dağılım ortalama etrafında simetriktir. Normal dağılımın simetrik olması sebebiyle, c mutlak değer olarak sabit olmak koşuluyla μ ve μ+c bulunan bireylerin % miktarları, μ ve μ-c arasında bulunan bireylerin % miktarına eşittir.

Normal dağılım -∞ ile +∞ arasındaki bütün değerleri kapsar ve üzerinde durulan özelliğin -∞ ile +∞ arasında bulunma olasılığı 1'e eşittir.

4.4. Standart Normal Dağılım

Ortalaması ve standart sapması bilinen her normal dağılımda, (4.6) numaralı eşitlik kullanılarak standardize edilebilir. Bunun için her bir gözlem değerinden (X_i) populasyon ortalaması çıkarılarak elde edilen fark, populasyonun standart sapmasına bölünür ve Standart Değer (Z_i) bulunur.

$$Z_i = \frac{X_i - \mu}{\sigma} \qquad \text{…(4.6)}$$

(4.6) numaralı eşitlik kullanılarak hesaplanan ve Z_i olarak gösterilen değerler, her bir gözlemin standardize edilmiş şeklidir. Z-değerlerinin gösterdiği dağılıma standart normal dağılım (veya Z-dağılımı) denir. Böylece ortalamaları ve standart sapmaları farklı olan sonsuz adet normal dağılım, (4.6) numaralı eşitlikle ortalaması 0, standart sapması 1 olan tek bir normal dağılıma (standart normal dağılım) dönüştürülmüş olur.

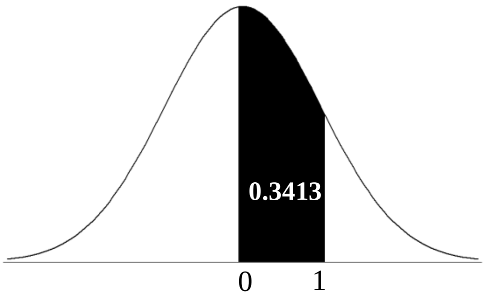
Standart normal dağılım da çan eğrisi şeklinde olup ortalama etrafında simetriktir. Yani $|-Z|=|+Z|$ olmak üzere, ortalama ile +Z arasında olma olasılığı, ortalama ile -Z değeri arasında olma olasılığı birbirine eşittir. Dolayısıyla tüm Z-değerlerinin %50'si ortalamadan büyük diğer %50'si ise ortalamadan küçüktür. Bu dağılıma dahil olan bütün Z değerlerinin oluş olasılıklarının toplamı 1'e eşittir. Standart normal dağılımın olasılık yoğunluk fonksiyonu;

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \qquad \ldots(4.7)$$

şeklindedir. Parametresi bilinen ve normal dağılım gösteren X değerleri standardize edilerek Z değerlerine çevrilebilirler. (4.7) numaralı eşitlikte verilen olasılık yoğunluk fonksiyonunun, 0 ile belirli Z değerleri arasındaki integrali alınarak tablo haline getirilmiştir (Tablo A). Tablo A kullanılarak ortalama (0) ile belirli Z-değerleri arasında kalan alanlar kolayca hesaplanabilir.

ÖRNEK 1:

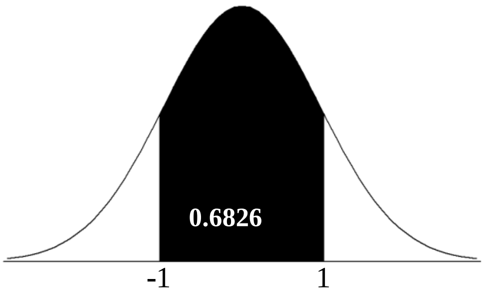
Standart normal dağılımda bütün Z-değerlerinin % kaçı 0 ile 1.0 değerleri arasındadır?



Tablo A'da görüldüğü gibi standart normal dağılımda 0 ile 1 arasındaki alan 0.3413'tür, yani standart normal dağılımı oluşturan bütün Z-değerlerinin %34.13'ü 0 ile 1 arasındadır. Başka bir deyişle; standart normal dağılımda herhangi bir Z-değerinin sıfır ile 1 arasında olma olasılığı %34.13'tür.

Standart normal dağılım simetrik olduğu için bütün Z-değerlerinin %34.13'ü de -1 ile 0 arasındadır.

Standart normal dağılımda bütün Z-değerlerinin -1 ile 1 arasında olma olasılığı ise aşağıdaki grafikte siyah taralı alandır ve P(-1<Z<0) ile P(0<Z<1) olasılıklarının toplamıdır.

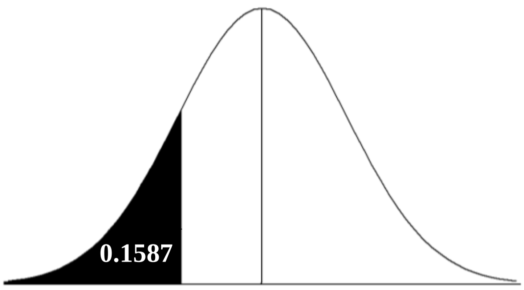


Bu durumda siyah taralı alan P(-1<Z<0)+P(0<Z<1) olarak hesaplanır ve;

0.3413+0.3413=0.6826

olarak bulunur, yani standart normal dağılımı oluşturan bütün Z-değerlerinin %68.26'sı -1 ile 1 arasındadır.

Standart normal dağılımda bütün Z-değerlerinin –1'den küçük olma olasılığı ise aşağıdaki grafikte siyah taralı alandır ve 0.5-P(-1<Z<0) olarak hesaplanır.



Bu durumda yandaki grafikte siyah taralı alan;

0.5-P(-1<Z<0)

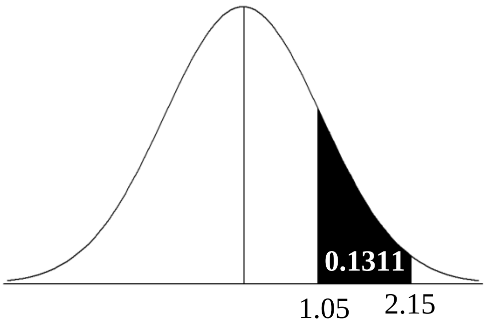
= 0.5-0.3413=0.1587

olarak hesaplanır, yani bütün Z-değerlerinin %15.87'si -1'den küçüktür.

ÖRNEK 2:

Standart normal dağılımda bütün Z-değerlerinin % ne kadarı 1.05 ile 2.15 arasındadır?

Standart normal dağılımda bütün Z-değerlerinin 1.05 ile 2.15 arasında olanlarının oluşturduğu alan aşağıdaki grafikte siyah taranmış alandır. Bu alanı bulmak için Z-değerlerinin 0 ile 2.15 arasında olma olasılığından 0 ile 1.05 arasında olma olasılığı çıkarılır, yani P(0<Z<2.15)-P(0<Z<1.05) hesaplanır.



Tablo A'dan P(0<Z<2.15)=0.4842 ve -P(0<Z<1.05)=0.3531 olarak bulunur. Dolayısıyla grafikteki taralı alan 0.4842-0.3531=0.1311 olarak hesaplanır. Bu, standart normal dağılımda bütün Z değerlerinin %13.11'inin 1.05 ile

2.15 arasında olduğunu gösterir.

ÖRNEK 3:

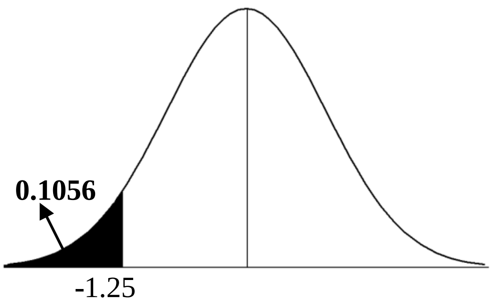
Sağlıklı yetişkinlerde toplam kolesterol değeri, ortalaması 160 mg/dL, standart sapması ise 20 mg/dL olan bir normal dağılım göstermektedir. Buna göre sağlıklı yetişkinlerin;

a. % kaçında toplam kolesterol değeri 135 mg/dL'den daha azdır?
b. % kaçında toplam kolesterol değeri 145 mg/dL ile 182 mg/dL arasındadır?
c. % kaçında toplam kolesterol değeri 115 mg/dL ile 135 mg/dL arasındadır?
d. % kaçında toplam kolesterol değeri 160 mg/dL ile 180 mg/dL arasındadır?
e. Toplam kolesterol miktarı en fazla olan %2.5' inin içinde en az kolesterol miktarı kaç mg/dL'dir?
f. Toplam kolesterol miktarı en az olan %5' inin içinde en fazla kolesterol miktarı kaç mg/dL'dir?

a. Burada sağlıklı yetişkinlerin % kaçının toplam kolesterol değerinin 135 mg/dL'den daha az olduğu, yani P(X<135 mg/dL) sorulmuştur. Yetişkinlerin toplam kolesterol değerlerinin ortalaması 160 mg/dL ve standart sapmasının da 20 mg/dL olan normal bir dağılım gösterdiği bildirilmiştir. Bunun için ilk olarak (4.6) numaralı eşitlik kullanılarak 135 değerinin $\frac{(135-160)}{20}$ şeklinde standardize edilmesi gerekir. Daha sonra istenilen alanın olasılığı:

$$P(X<135) = P(Z<\frac{(135-160)}{20}) = P(Z<-1.25)$$

şeklinde bulunur. Standart normal dağılımda, Z-değerlerinin -1.25'ten küçük olduğu alan, taralı alandır.



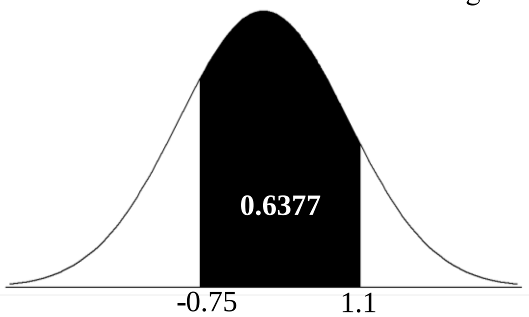
Bu alan 0.5-[P(-1.25<Z<0)] şeklinde hesaplanır. Tablo A'dan bakıldığı zaman Z-değerlerinin 0 ile 1.25 arasında olma olasılığı 0.3944'tür. Standart normal dağılım simetrik bir dağılım olduğu için -1.25 ile 0 arasında olma olasılığı da 0.3944'tür. Grafikteki taralı alan 0.5-0.3944=0.1056 olarak bulunur, yani sağlıklı yetişkinlerin %10.56'sının toplam kolesterol değeri 135 mg/dL'den daha azdır.

b. Sağlıklı bireylerin % kaçının toplam kolesterol değerinin 145 mg/dL ile 182 mg/dL arasında olduğunu bulmak için önce (4.6) numaralı eşitlik kullanılarak 145 ve 182 mg/dL değerlerinin standardize edilmesi gerekir.

$$P(145 < X < 175) = P(\frac{145-160}{20} < Z < \frac{182-160}{20})$$

$$=P(-0.75<Z<1.1)$$

Burada hesaplanan Z-değerlerinden biri negatif diğeri ise pozitif işaretlidir. Bu, hesaplanacak alanın ortalamanın iki tarafında olduğunu gösterir ve aşağıdaki grafikte siyah taralı alandır.



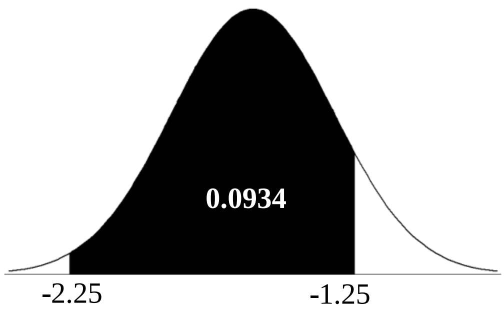
Bunun için P(-0.75<Z<0) ile P(0<Z<1.1) olasılıklarının toplanması gerekir. Tablo A'dan, P(-0.75<Z<0)=0.2734 ve P(0<Z<1.1)=0.3643 olarak bulunur. Bu durumda taralı alan 0.2734+0.3643=0.6377 olarak hesaplanır, yani sağlıklı yetişkinlerin %63.77'sinin toplam kolesterol değeri 145 mg/dL ile 182 mg/dL arasındadır.

c. Sağlıklı bireylerin % kaçının toplam kolesterol değerinin 115 mg/dL ile 135 mg/dL arasında olduğunu bulmak için önce (4.6) numaralı eşitlik kullanılarak 115 ve 135 mg/dL değerlerinin standardize edilmesi gerekir.

$$P(115 < X < 135) = P(\frac{115-160}{20} < Z < \frac{135-160}{20})$$

$$=P(-2.25<Z<-1.25)$$

Standart normal dağılımda P(-2.25<Z<-1.25), Z-değerlerinin -2.25 ile 0 arasında olma olasılığından Z-değerlerinin -1.25 ile 0 arasında olma olasılığı çıkarılarak hesaplanır, yani yandaki grafik üzerinde siyah taralı alan hesaplanacaktır. Sağlıklı yetişkinlerin toplam



kolesterol değerlerinin 115 ile 135 mg/dL arasında olma olasılığı, yani standart normal dağılımda Z-değerlerinin P(-2.25<Z<-1.25) olasılığı;

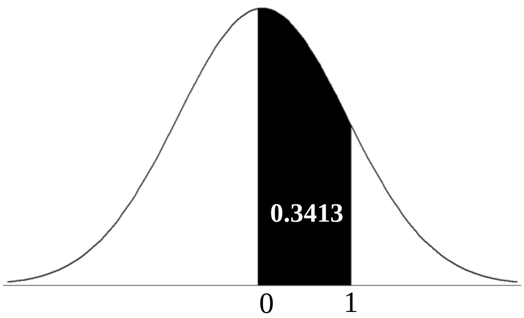
$$P(-2.25<Z<-1.25)= P(-2.25<Z<0)- P(-1.25<Z<0)$$

$$=0.4878-0.3944=0.0934$$

olarak hesaplanır. Bu durumda sağlıklı yetişkinlerin toplam kolesterol değeri 115 mg/dL ile 135 mg/dL arasında olanların oranı %9.34'tür.

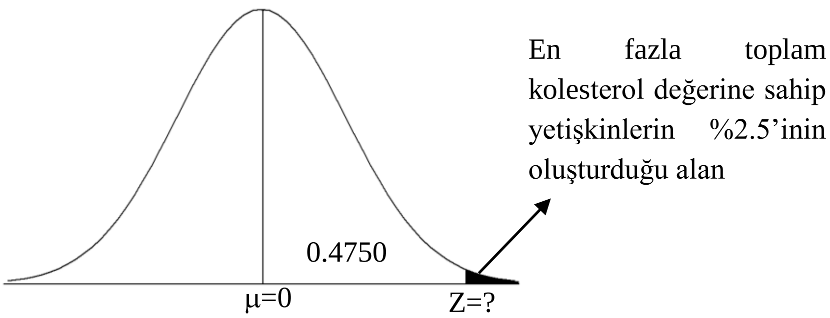
d. Sağlıklı bireylerin % kaçının toplam kolesterol değerinin 160 mg/dL ile 180 mg/dL arasında olduğunu bulmak için önce (4.6) numaralı eşitlik kullanılarak 160 ve 180 mg/dL değerlerinin standardize edilmesi gerekir.

$$P(160 < X < 180) = P\left(\frac{160-160}{20} < Z < \frac{180-160}{20}\right) = P(0<Z<1)$$



Tablo A'dan, standart normal dağılımda 0 ile 1 arasındaki alan 0.3413'tür, yani sağlıklı bireylerin toplam kolesterol değerinin 160 ile 180 mg/dL arasında olma olasılığı %34.13'tür.

e. Burada istenen, toplam kolesterol miktarı en fazla olan %2.5'uğun içindeki en az kolesterol miktarının kaç mg/dL olduğudur. Bu kolesterol değerinin bulunabilmesi için ilk olarak standart normal dağılımda %2.5'lik alanın hangi Z-değerinden başladığı belirlenmelidir.

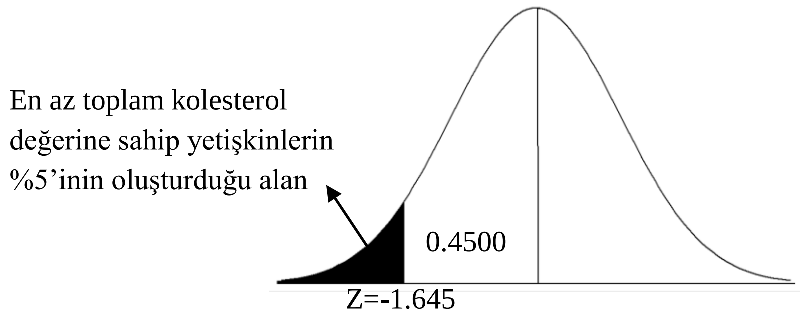


Kolesterol değeri en fazla %2.5 olanlar ile ortalama arasında kalan alan yukarıdaki grafikte görüldüğü gibi (%50 - %2.5) %47.50'dir. En fazla kolesterol değerine sahip %2.5'lik alanın başlangıç değerinin hesaplanabilmesi için standart normal dağılım tablosunun tersten kullanılması gerekmektedir, yani 0 ile hangi Z-değeri arasında kalan alanın 0.4750 olduğunun bulunması gerekir. Tablo A'dan Z-değeri bulunduktan sonra (4.6) numaralı eşitlik kullanılarak %2.5'lik alanın başlangıç noktası hesaplanabilir. Tablo A'dan, 0 ile Z=1.96 arasında kalan alanının 0.4750 olduğu kolaylıkla bulunur. Bu durumda;

$$1.96 = \frac{X-160}{20} \text{ ise}$$
$$X = 1.96(20) + 160 = 199.2 \text{ mg/dL}$$

olarak bulunur, yani yetişkin bireylerin en yüksek toplam kolesterol değerine sahip %2.5'inin kolesterol değeri 199.2 mg/dL'den daha yüksektir. Bir başka deyişle; en yüksek kolesterol değerine sahip %2.5 içinde en düşük kolesterol değeri 199.2 mg/dL'dir.

f. Burada istenen toplam kolesterol miktarı en az olan %5' in içindeki en fazla kolesterol miktarının kaç mg/dL olduğudur. Burada da yine bir önceki şıkta olduğu gibi standart normal dağılım tablosu tersinden kullanılmalıdır. En az kolesterol değerine sahip %5'inin başladığı nokta sorulduğuna göre ortalama ile bu nokta arasında kalan alan (%50 - % 5) %45'tir. Tablo A'ya bakıldığı zaman 0 ile Z=1.64 arasındaki alanın 0.4495 ve 0 ile Z=1.65 arasındaki alanın ise 0.4505 olduğu görülür. 0.4495 ile 0.4505'in ortalaması alınırsa tam 0.45 değeri elde edilir. 0.45 değerine karşılık gelen Z değerini bulmak için de benzer şekilde 1.64 ile 1.65'in ortalaması alınırsa 1.645 bulunur yani 0 ile Z=1.645 arasındaki alan 0.45'tir.



Yukarıdaki grafikte görüldüğü gibi Z-değeri, ortalamanın sol tarafında olduğu için negatif işaretlidir. Standart normal dağılım simetrik olduğu için 0 ile 1.645 arasında kalan alan ile -1.645 ile 0 arasında kalan alan birbirine eşittir. Bu durumda kolesterol değeri;

$$-1.645=\frac{X-160}{20} \text{ ise}$$
$$X=-1.645(20)+160=127.1 \text{ mg/dL}$$

olarak hesaplanır, yani yetişkin bireylerin en düşük toplam kolesterol değerine sahip %5'inin kolesterol değeri 127.1 mg/dL'den daha düşüktür. Bir başka deyişle; en düşük kolesterol değerine sahip %5 içinde en yüksek kolesterol değeri 127.1 mg/dL'dir.

4.3.3. Frekans Dağılım Tablosunda Normal Dağılıma Göre Olması Beklenen Frekansların Hesaplanması

Yapılan bir çalışmada örnek oluşturularak veriler toplandıktan sonra Bölüm I'de açıklandığı gibi araştırıcı verilerini frekans dağılım tablosu yardımıyla özetleyebilir. Araştırıcı, frekans dağılım tablosunu oluşturduktan sonra elde edilen verilerin ortalama ve standart sapmasını hesaplayabilir. Frekans dağılım tablosu oluşturularak ortalama ve standart sapma hesaplandıktan sonra ortalaması ve standart sapması frekans dağılım tablosundan hesaplanan normal dağılıma göre frekans dağılım tablosundaki her sınıfın beklenen frekansı hesaplanabilir.

Bölüm I'de fasulye ağırlıkları ile ilgili yapılan araştırmada 60 fasulye ağırlığı için frekans dağılım tablosu Tablo 1.7'de verilmişti.

Tablo 1.7'de verilen 60 fasulye ağırlığının frekans dağılım tablosundan, ortalama 0.661 g ve standart sapmada 0.081619 ≅ 0.082 g olarak bulunmuştur. Örnekten hesaplanan ortalama ve standart sapma söz konusu normal dağılımın parametreleri kabul edilerek, ortalaması 0.661 ve standart sapması 0.082 olan normal dağılıma göre sınıfların beklenen frekansları hesaplanabilir.

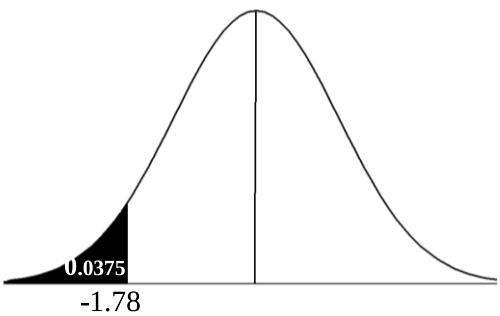
Normal dağılıma göre beklenen frekansların hesaplanabilmesi için ilk olarak Tablo 4.7'de görüldüğü gibi her sınıfın üst gerçek sınırlarının hesaplanması gerekir.

Tablo 4.7. 60 fasulye tanesinin ağırlığına ait gözlenen ve ortalaması 0.661 ve standart sapması 0.082 olan normal dağılıma göre beklenen frekanslar

Sınıf değeri	Alt gerçek sınır	Üst gerçek sınır	Frekans	Olasılık	Beklenen frekans
0.48 - 0.51	0.475	0.515	2	0.0375	2.250
0.52 - 0.55	0.515	0.555	5	0.0610	3.660
0.56 - 0.59	0.555	0.595	6	0.1105	6.630
0.60 - 0.63	0.595	0.635	9	0.1655	9.930
0.64 - 0.67	0.635	0.675	13	0.1930	11.580
0.68 - 0.71	0.675	0.715	8	0.1779	10.674
0.72 - 0.75	0.715	0.755	9	0.1295	7.770
0.76 - 0.79	0.755	0.795	5	0.0735	4.410
0.80 - 0.83	0.795	0.835	3	0.0516	3.096
		Toplam	60	1.0000	60

Tablo 4.7'de görüldüğü gibi **1. sınıf** 0.515 üst sınırında bitmektedir, yani bu sınıftaki fasulye tanelerinin ağırlığı 0.515 g'dan azdır. Bu durumda ortalaması 0.661 g ve standart sapması 0.082 g olan normal dağılımda tüm fasulye tanelerinin % kaçının ağırlığının 0.515'ten daha az olduğunun aşağıdaki şekilde hesaplanması gerekir.

$$P(X<0.515)=P(Z<\frac{(0.515-0.661)}{0.082})=P(Z<-1.78)$$



Ortalaması 0.661 g ve standart sapması 0.082 g olan normal dağılımda, fasulye ağırlıklarının 0.515'ten daha az veya standart normal dağılımda Z değerlerinin -1.78'ten daha küçük olma olasılığı yandaki dağılımda siyah taralı alandır ve;

$P(Z<-1.78)=0.5-P(0<Z<-1.78)=0.5-0.4625=0.0375$

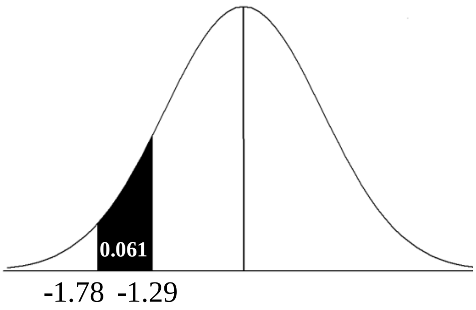
olarak hesaplanır. Bu durumda 60 fasulye tanesinden %3.75'sinin, yani 60x0.0375=2.25 tanesinin ağırlığı 0.515'ten azdır. Bu durumda Tablo 4.7'de de görüldüğü gibi 1. sınıfın beklenen frekansı 2.25'tir.

İkinci sınıftaki fasulye tanelerinin ağırlıkları 0.515 ile 0.555 g arasındadır, çünkü bu sınıf 0.515 sınırından başlayıp 0.555 üst sınırında bitmektedir. Bu sınıfın ortalaması 0.661 g ve standart sapması 0.082 g olan normal dağılıma göre beklenen frekansının hesaplanabilmesi için fasulyelerden

% kaçının ağırlığının 0.515 ile 0.555 g arasında olduğunun yani P(0.515<X<0.555) olasılığının aşağıdaki şekilde hesaplanması gerekir.

$$P(0.515 < X < 0.555) = P(\frac{0.515-0.661}{0.082} < Z < \frac{0.555-0.661}{0.082})$$

=P(-1.78<Z<-1.29)



Standart normal dağılımda P(-1.78 <Z< -1.29), Z-değerlerinin -1.78 ile 0 arasında olma olasılığından Z-değerlerinin -1.29 ile 0 arasında olma olasılığı çıkarılıp yandaki grafik üzerinde siyah taralı alan bulunur.

Fasulye ağırlıklarının 0.515 ile 0.555 g arasında olma olasılığı, yani standart normal dağılımda Z-değerlerinin P(-178<Z<-1.29) olma olasılığı;

P(-1.78<Z<-1.29)= P(-1.78<Z<0)- P(-178<Z<-1.29)
=0.4625-0.4015=0.061

olarak hesaplanır. Bu durumda fasulyelerin ağırlıklarının 0.515 g ile 0.555 g arasında olma olasılığı %6.1'dir, yani bu sınıfın beklenen frekansı 60x0.061=3.66'dır.

Üçüncü sınıftaki fasulyelerin ağırlıkları 0.555 g ile 0.595 g arasındadır. Fasulyelerin bu aralıkta olma olasılığı P(0.555<X<0.595)'tir, yani Z-değerlerinin -1.29 ile -0.81 arasında olma olasılığıdır ve 0.1105 olarak hesaplanır. Bu durumda üçüncü sınıfın beklenen frekansı 60x0.1105=6.63'tür.

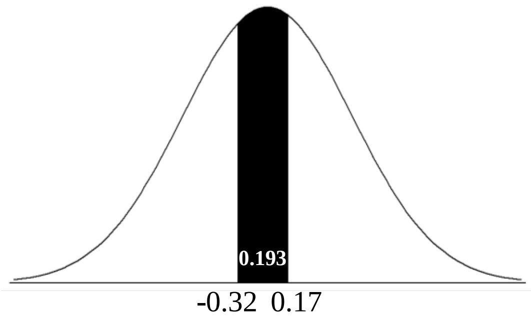
Dördüncü sınıftaki fasulyeler 0.595 g ile 0.635 g arasındadır. Fasulyelerin bu aralıkta olma olasılığı P(0.595<X<0.635)'tir, yani Z-değerlerinin -0.81 ile -0.32 arasında olma olasılığıdır ve 0.1655 olarak hesaplanır. Bu durumda dördüncü sınıfın beklenen frekansı 60x0.1655=9.93'tür.

Beşinci sınıftaki fasulyelerin ağırlıkları 0.635 g ile 0.675 g arasındadır, çünkü bu sınıf 0.635 sınırından başlayıp 0.675 üst sınırında bitmektedir. Bu sınıfın ortalaması 0.661 g ve standart sapması 0.082 g olan normal dağılıma göre beklenen frekansının hesaplanabilmesi için fasulyelerden % kaçının ağırlığının 0.635 ile 0.675 g arasında olduğunun yani P(0.635<X<0.675) olasılığının aşağıdaki şekilde hesaplanması gerekir.

$$P(0.635 < X < 0.675) = P(\frac{0.635-0.661}{0.0819} < Z < \frac{0.675-0.661}{0.0819})$$

=P(-0.32<Z<0.17)

olarak hesaplanır. Burada hesaplanan Z-değerlerinden biri negatif diğeri ise pozitif işaretlidir. Bu, hesaplanacak alanın ortalamanın iki tarafında olduğunu gösterir ve aşağıdaki grafikte siyah taralı alandır. Bu alanı bulmak



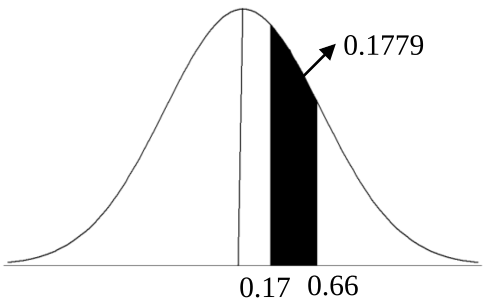
için P(-0.32<Z<0) ile P(0<Z<0.17) olasılıklarının toplanması gerekir. Bu durumda taralı alan 0.1255+0.0675=0.193 olarak hesaplanır. Bu, tüm fasulye tanelerinin %19.3'ünün ağrılığının 0.635 g ile 0.675 g arasında olduğunu gösterir.

Dolayısı ile bu sınıfın beklenen frekansı 60x0.193=11.58 olarak hesaplanır.

Altıncı sınıfta yer alan fasulyelerin ağırlıkları 0.675 g ile 0.715 g arasındadır. Bu sınıfın ortalaması 0.661 g ve standart sapması 0.082 g olan normal dağılıma göre beklenen frekansının hesaplanabilmesi için fasulyelerden % kaçının ağırlığının 0.675 ile 0. 715 g arasında olduğunun yani P(0.675<X<0.715) olasılığının aşağıdaki şekilde hesaplanması gerekir.

$$P(0.675 < X < 0.715) = P(\frac{0.675-0.661}{0.082} < Z < \frac{0.715-0.661}{0.082})$$
$$=P(0.17<Z<0.66)$$

olarak hesaplanır. Burada hesaplanan Z-değerlerinin ikisi de pozitif işaretlidir. Bu, hesaplanacak alanın ortalamanın sağ tarafında olduğunu gösterir ve aşağıdaki grafikte siyah taralı alandır. Bu alanı bulmak için;

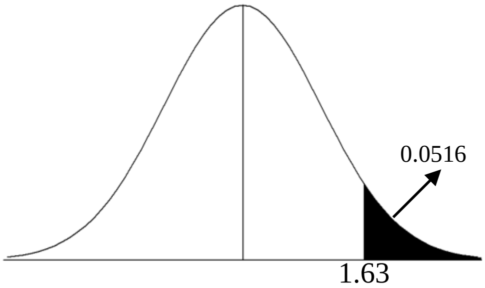


P(0<Z<0.66) ile P(0<Z<0.17) arasındaki fark bulunmalıdır. Bu durumda yandaki grafikte taralı alan 0.2454-0.0675=0.1779 olarak bulunur. Bu değer, tüm fasulye tanelerinin %17.79'unun ağırlığının 0.675 g ile 0.715 g arasında olduğunu gösterir. Dolayısı ile bu sınıfın beklenen frekansı 60x0.1779=10.674'tür.

Yedinci sınıftaki fasulyeler 0.715 g ile 0.755 g arasındadır. Fasulyelerin bu aralıkta olma olasılığı P(0.715<X<0.755)'tir, yani Z-değerlerinin 0.66 ile 1.15 arasında olma olasılığıdır ve 0.1295 olarak hesaplanır. Bu durumda yedinci sınıfın beklenen frekansı 60x0.1295=7.77'dir.

Sekizinci sınıftaki fasulyeler 0.755 g ile 0.795 g arasındadır. Fasulyelerin bu aralıkta olma olasılığı P(0.755<X<0.795)'tir, yani Z-değerlerinin 1.15 ile 1.63 arasında olma olasılığıdır ve 0.0735 olarak hesaplanır. Bu durumda sekizinci sınıfın beklenen frekansı 60x0.0735=4.41'dir.

Son sınıfta yer alan fasulyeler ağırlıkları 0.795 g'dan daha ağır olan fasulyelerdir. Bu sebeple ortalaması 0.661 g ve standart sapması 0.082 g olan normal dağılımda tüm fasulyelerin % kaçının ağırlığının 0.795 g'dan daha fazla olduğunun, yani P(X>0.795)'in hesaplanması gerekir. 0.795 karşılık gelen Z değeri 1.63 olduğu için standart normal dağılımda 1.63'ten büyük olma olasılığı hesaplanmalıdır. Bu alan aşağıdaki grafikte taralı gösterilen alandır.



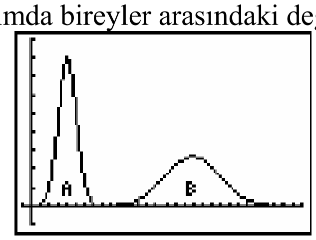
Bu alan P(Z>1.63)'tür ve 0.5- [P(0<Z<1.63)] = 0.0516 olarak hesaplanır. Bu, tüm fasulye tanelerinin %5.16'sının 0.795 g'dan daha ağır olduğunu gösterir.

Dolayısı ile bu sınıfın beklenen frekansı 60x0.0516=3.096'dır.

Her sınıf için normal dağılıma göre hesaplanan olasılıklar ve her sınıfın söz konusu normal dağılıma göre beklenen frekansı Tablo 4.7'de verilmiştir. Tablo 4.7'den de görüldüğü gibi frekans dağılım tablosunda her sınıf için hesaplanan olasılıkların toplamı 1'e eşit olmalıdır. Aynı şekilde her sınıf için hesaplanan beklenen frekansların toplamı da gözlenen frekansların toplamına eşit olmalıdır.

4.4. Sorular

1. Dağılım kavramı ne ifade etmektedir? Açıklayınız.
2. Sürekli ve kesikli dağılım ne demektir? Örnekler vererek açıklayınız.
3. Aşağıdaki grafiğe göre hangi dağılımda bireyler arasındaki değişim daha fazladır.



4. Dağılımın parametresi kavramı ne ifade etmektedir? Açıklayınız.
5. Binomiyal dağılımın parametreleri nelerdir? Açıklayınız?
6. Binomiyal dağılımın şekli nasıldır? Örnek vererek ve grafik çizerek açıklayınız.
7. Poisson dağılımının parametreleri nelerdir? Açıklayınız.
8. Poisson dağılımının şekli nasıldır? Örnek vererek ve grafik çizerek açıklayınız.
9. Normal dağılımın parametrelerini açıklayınız.
10. Normal dağılımın özellikleri nelerdir? Açıklayınız.
11. Standart normal dağılım nasıl elde edilir? Açıklayınız.
12. Standart normal dağılımın özellikleri nelerdir? Açıklayınız.
13. İstatistik dersinin ara sınavından aldığınız notun 60, istatistik dersi ara sınav not ortalamasının da μ = 65 olan normal dağılım gösterdiği bilinmektedir. Aldığınız not, sınıf notlarının en düşük % 2.5'unun başladığı değer olduğuna göre İstatistik dersi ara sınav notlarına ait normal dağılımın standart sapması kaçtır?
14. Ankara ilindeki 40-60 yaş arasındaki insanlarda yakını görememe probleminin ortaya çıkma olasılığının 0.60 olduğu bilinmektedir. Buna göre 40-60 yaş arası 3 kişiden;
    a. En az ikisinin yakını görememe olasılığını hesaplayınız.
    b. Üçünün de yakını görememe olasılığını hesaplayınız.
    c. En fazla birinde yakını görememe olasılığını hesaplayınız.
15. Sağlıklı yetişkinlerde total kolesterol değeri, ortalaması 160 mg/dL, standart sapması ise 20 mg/dL olan bir normal dağılım göstermektedir. Buna göre sağlıklı yetişkinlerin;

a. % kaçında total kolesterol değeri 135 mg/dL'den daha azdır?
b. % kaçında total kolesterol değeri 145 mg/dL ile 175 mg/dL arasındadır?
c. % kaçında total kolesterol değeri 135 mg/dL ile 155 mg/dL arasındadır?
d. Total kolesterol miktarı en fazla olan %5' inin içinde en az kolesterol miktarı kaç mg/dL'dir?

16. 67 bireyin oluşturduğu bir örnekte, ortalama 1.8 ve varyans 1.8 olarak hesaplanmıştır. Bu gözlem değerlerinin dağılımı nasıldır?
17. Her hangi bir dersten yapılan sınav sonunda öğrencilerin %40'nın başarılı olduğu saptanmıştır. Bu dersi alan öğrencilerden tesadüfen 3 öğrenci seçilse bu öğrencilerden en az 2 tanesinin başarılı olma olasılığı kaçtır?
18. Ortalaması μ =140 cm ve standart sapması σ = 5 cm olan normal dağılım gösteren bir populasyonda, en uzun gözlem değerlerinin %5'i hangi değerden daha uzundur?
19. Ortalaması μ =140 cm, varyansı da σ^2 = 25 cm olan normal dağılım gösteren bir populasyonda, en kısa gözlem değerlerinin %5'i hangi değerden daha kısadır?
20. Üniversitelerde sigara içme oranının 0.50 olduğu bilinmektedir. Tesadüfen alınan 4 kişilik bir öğrenci grubunda bu özellik bakımından varyans kaça eşittir? Hesaplayarak anlamını açıklayınız.
21. 100 kişilik bir sınıfta herhangi bir kişinin 1 Ocak'ta doğmuş olma olasılığına ilişkin varyans kaça eşittir? Hesaplayarak anlamını açıklayınız.
22. Ortalaması μ=762, standart sapması σ=75 olan normal bir dağılımdaki bireylerden 762 çıkarıldıktan sonra, elde edilen değerlerin hepsi 75'e bölünüp yeni değerler bulunsa, bu yeni değerler nasıl bir dağılım gösterir? Açıklayınız.
23. Ziraat Fakültesindeki öğrencilerin %40'ının kız, geri kalanının da erkek olduğu bilinmektedir. Ziraat Fakültesinden rasgele 2 öğrenci alınsa bunlardan 2'sinin de kız olma olasılığı kaçtır?
24. Ortalaması μ = 124.0 cm, standart sapması da σ = 2.0 cm olan, normal dağılım gösteren bir populasyonda, varyantların 126 cm ile 128 cm arasında olanlarının olasılığı kaçtır?
25. Merinos koyunu yapağı inceliklerinin, ortalaması μ = 20 μ standart sapması da σ = 2 μ olan normal bir dağılım gösterdiği bilinmektedir.

a. Piyasadan rasgele alınan merinos yapağılarında inceliğin 16 μ ile 24 μ arasında olma olasılığı ne kadardır?

b. Piyasadan rasgele alınan merinos yapağılarında inceliğin 14 μ ile 16 μ arasında olma olasılığı ne kadardır?

26. Bir ilaç firmasının ürettiği oksijenli suların kullanım ömrünün, ortalaması 24 ay ve standart sapması 2 ay olan normal dağılım gösterdiği bilinmektedir. Buna göre üretilen oksijenli suların;
a) % ne kadarının raf ömrü 24 aydan daha fazladır?
b) % ne kadarının raf ömrü 22 ay ile 25 ay arasındadır?
c) % ne kadarının raf ömrü 20 aydan daha kısadır?
d) En kısa raf ömürlü oksijenli suların %2.5' inin raf ömrü kaç aydan daha kısadır?
e) Bu firmanın bir parti de ürettiği 2500 oksijenli suyun kaç tanesinin raf ömrü 27.92 aydan daha kısadır?

27. Üniversite öğrencilerinin %30'unun sigara içtikleri bilinmektedir. Buna göre tesadüfen seçilen 3 öğrenciden 2'sinin sigara içme olasılığı ve bu dağılımın ortalaması kaçtır?
28. Bir eczaneye gelen hastalardan kas spazmı nedeniyle belirli bir kas gevşetici ilacı (X) kullananların oranı %68 olarak tespit edilmiştir. Tesadüfen seçilen 4 hastadan;

a. 3' ünün X kas gevşeticisini kullanma olasılığı nedir?
b. En az ikisinin X kas gevşeticisini kullanma olasılığı nedir?
c. Söz konusu dağılımın şekli nasıldır? Açıklayınız.

29. Ankara'daki üniversite öğrencilerinin %25' inin herhangi bir kütüphaneye üye olduğu bilinmektedir. Buna göre tesadüfen seçilen 4 öğrenciden;

a. En az 2' sinin herhangi bir kütüphaneye üye olma olasılığı nedir?
b. En fazla 2' sinin herhangi bir kütüphaneye üye olma olasılığı nedir?
c. 4' ünün de herhangi bir kütüphaneye üye olma olasılığı nedir?

30. Ankara'da 95 yaşından fazla yaşama şansı bir araştırma sonucunda 0.0005 olarak hesaplanmıştır. Buna göre, Ankara'da yaşayan insanlardan tesadüfen seçilecek 1000 kişiden;

a. Birinin 95 yaşından daha yaşlı olma olasılığı nedir?
b. İkisinin de 95 yaşından daha yaşlı olma olasılığı nedir?
c. En az üçünün 95 yaşından daha yaşlı olma olasılığı nedir?

31. 2000 kişinin yaşadığı bir şehirde, kalp krizi geçirmiş kişilerin oranının 0.0002 olduğu bildirilmiştir. Bu şehirden tesadüfen seçilen 3 kişiden;

a. En fazla 1 kişinin kalp krizi geçirmiş olma olasılığı nedir?
b. En az 2 kişinin kalp krizi geçirmiş olma olasılığı nedir?
c. 3 kişinin de kalp krizi geçirmiş olma olasılığı nedir?
d. Bu hastaların gösterdiği dağılımın parametrelerini hesaplayınız.

32. 120 aile üzerinde yapılan bir anket sonucunda bu ailelerden 80 tanesinin düzenli bir şekilde her gün gazete okudukları saptanmıştır. Buna göre;

a) Tesadüfen seçilen 5 aileden hepsinin düzenli olarak gazete okuyan aile olma olasılığı nedir?
b) Ortalama gazete okuma oranı nedir?
c) Bu dağılımın varyansını bulunuz.

33. 1000 kişinin yaşadığı bir şehirde, Lösemili (kan kanseri) hastaların oranının 0.0001 olduğu bildirilmiştir. Bu şehirden tesadüfen seçilen 3 kişiden;

a) En fazla 2 kişinin lösemili hasta olma olasılığı nedir?
b) En az 3 kişinin lösemili hasta olma olasılığı nedir?
c) Bu hastaların gösterdiği dağılımın parametrelerini hesaplayınız?