SAKARYA ÜNİVERSİTESİ

Olasılık ve İstatistik

HAFTA 7

Özel Sürekli Olasılık Dağılımları:

Düzgün Dağılım
Üstel Dağılım
Normal Dağılım

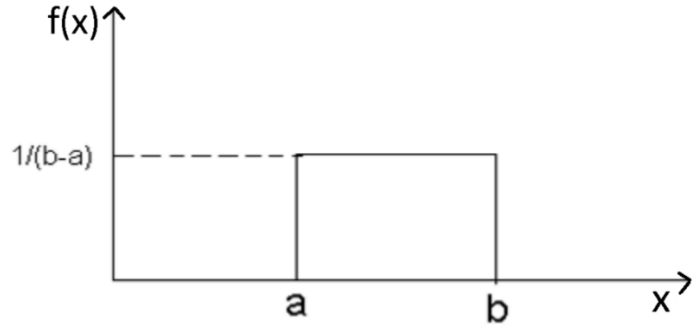
Dr. Öğretim Üyesi Burcu ÇARKLI YAVUZ

bcarkli@sakarya.edu.tr

Düzgün Dağılım

- İki değer arasında kalan bütün değerlerin olasılıklarının birbirine eşit olduğu dağılım türüdür. Yani tanımlı olduğu aralıkta olasılık yoğunluk fonksiyonunun değeri sabittir.
- Dikdörtgensel dağılım, uniform dağılım, tekdüze dağılım gibi isimlerle de anılabilir.
- Düzgün olasılık yoğunluk fonksiyonu altındaki toplam alan 1'e eşittir.
- a ve b değerleri arasında tanımlı, düzgün dağılım gösteren bir sürekli rastgele değişkenin olasılık dağılımı şu şekildedir:

$$f(x) = \begin{cases} \dfrac{1}{b-a} & a \leq x \leq b \\ 0 & dd \end{cases}$$



Düzgün dağılımın ortalaması ve varyansı

- Düzgün dağılımın ortalaması üst sınır ile alt sınırın aritmetik ortalamasına eşittir:

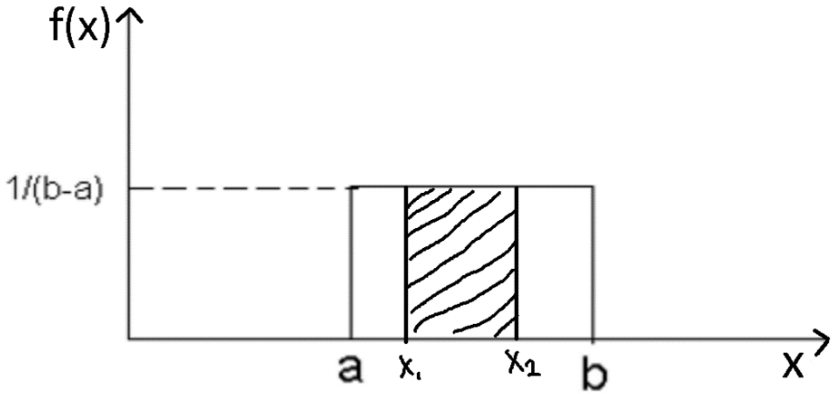
$$\mu = \frac{a+b}{2}$$

- Varyansı:

$$\sigma^2 = \frac{(b-a)^2}{12}$$

Düzgün Dağılım

- Düzgün dağılım gösteren bir rastgele değişkenin, iki değer arasında kalma olasılığı, düzgün dağılım grafiğindeki ilgili dikdörtgenin alanına eşittir.
- Bir X değişkeninin $x_1$ile x_2 değerleri arasında olma olasılığı:



$$P(x_1 < X < x_2) = P(x_1 \le X \le x_2) = \frac{x_2 - x_1}{b - a}$$

Örnek

- Bir benzin istasyonunda bir gün içinde satılan akaryakıtın 2000 ile 5000 litre arasında düzgün dağıldığı belirlenmiştir.

a) Günlük akaryakıt satışının 2500 ile 3000 litre arasında olma olasılığı

b) Günlük en az 4000 litre akaryakıt satılma olasılığı

c) Tam olarak 2500 litre akaryakıt satılması olasılığı nedir?

Çözüm

a=2000, b=5000

a) $P(2500 < X < 3000) = \frac{x2-x1}{b-a} = \frac{3000-2500}{5000-2000} = 0.1667$

b) $P(X > 4000) = \frac{x2-x1}{b-a} = \frac{5000-4000}{5000-2000}$=0.3333

c) P(X=2500)=0

Sürekli rastgele değişkenlerin bir noktadaki olasılığı sıfırdır.

Üstel Dağılım

- İki olayın oluşu arasındaki zamanın dağılımının genelde üstel olduğu kabul edilir.
- Örneğin bir boşaltma alanına kamyonların gelişleri arasındaki sürenin, bir ATM den para çekme işlemleri arasında geçen sürenin, bir firmanın müşteri hizmetlerine gelen telefonların arasındaki sürenin üstel dağılıma uyduğu farz edilir.
- Bir zaman aralığındaki olayların sayısı Poisson Dağılımına, bu olaylar arasında geçen süre üstel dağılıma uyar. Çünkü geliş sayısı kesikli, süre sürekli rassal değişkendir.

Üstel Dağılım

- Üstel dağılımın olasılık yoğunluk fonksiyonu şu şekildedir:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \; için \\ 0, diğer\ durumlarda \end{cases}$$

- X'in [a,b] aralığında olma olasılığı:

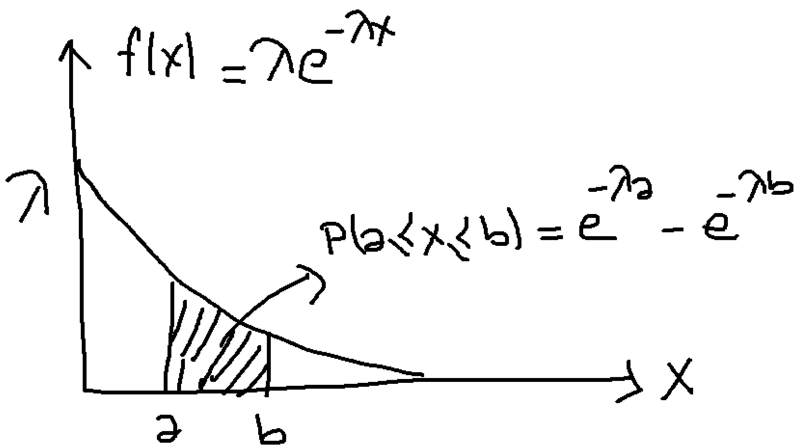
$$P(a \leq X \leq b) = e^{-\lambda a} - e^{-\lambda b}$$

- Özel olarak $e^0 = 1$ ve $e^{-\infty} = 0$ olduğu için

$P(X \leq b) = 1 - e^{-\lambda b}$ ve $P(X \geq a) = e^{-\lambda a}$ yazılabilir.

Üstel Dağılım

- Üstel ve Poisson dağılımlar bekleme yada kuyruk hatlarını analiz etmede kullanışlıdır.
- Kuyruk teorisi, hizmet eden kişi sayısını (acil serviste bulunan doktor sayısı gibi) belirlemek için kullanılır.



Üstel dağılımın ortalaması ve varyansı

- Üstel dağılıma sahip bir X rastgele değişkeninin ortalaması:

$$\mu = \frac{1}{\lambda}$$

- Varyansı:

$$\sigma^2 = \frac{1}{\lambda^2}$$

- Yani üstel dağılıma sahip bir X rastgele değişkeninin ortalaması ve standart sapması birbirine eşittir:

$$\mu = \sigma = \frac{1}{\lambda}$$

Örnek

- Bir pilin ömrünün $\lambda = 0.05$ olacak şekilde üstel dağıldığı belirlenmiştir.

a) Pilin ortalama ömrü ne kadardır?

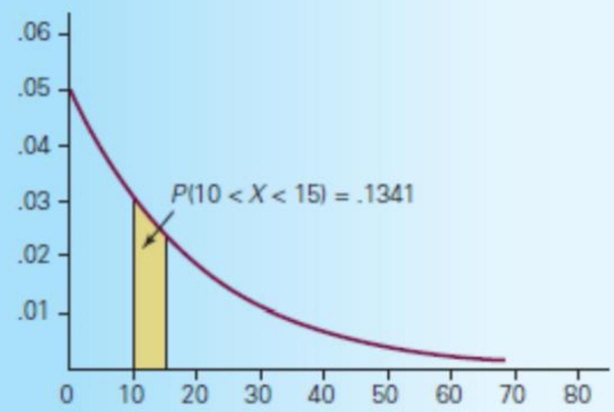
b) Pilin 10 ile 15 saat arasında bitmesi ihtimali nedir?

c) Bir pilin 20 saatten fazla dayanma ihtimali nedir?

Çözüm:

$a)\ \ \mu = \frac{1}{\lambda} = \frac{1}{0.05} = 20\ saat$

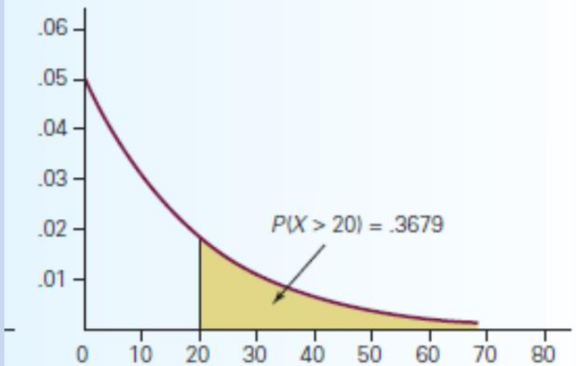
Çözüm



(b) $P(10 < X < 15)$

$$P(10 \leq X \leq 15) = e^{-0.05(10)} - e^{-0.05(15)}$$
$$= e^{-0.5} - e^{-0.75}$$
$$= 0.6065 - 0.4724$$
$$= 0.1341$$

$$P(X > 20) = e^{-0.05(20)} = e^{-1} = 0.3679$$



(c) $P(X > 20)$

Normal Dağılım

- Günlük yaşamda karşılaştığımız pek çok değişken normal dağılım göstermektedir.
- Normal rassal değişken için olasılık yoğunluk fonksiyonu şu şekilde ifade edilir:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$$

e=2.71828

π=3.14159

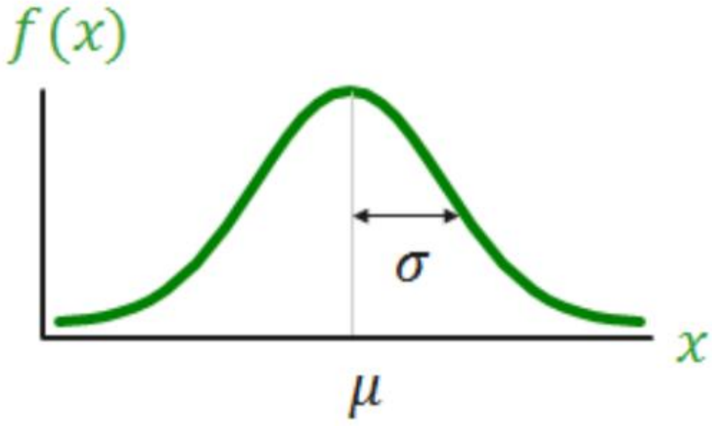
μ=Ana kütle ortalaması

σ=Ana kütle standart sapması

X=-∞<x<∞

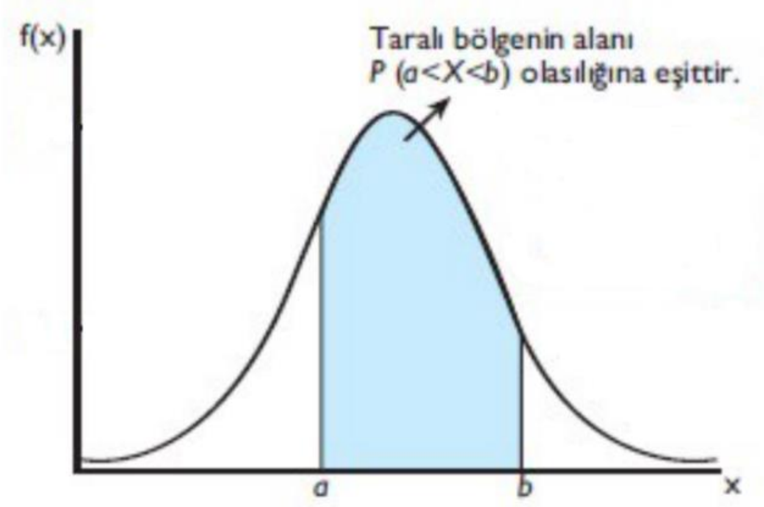
Normal Dağılım

- Normal dağılım, çan şeklinde bir dağılım gösterdiği için çan eğrisi olarak da karşımıza çıkabilir (yada Gaussian)
- Normal dağılım ortalama (μ) etrafında simetriktir.
- Eğrinin altında kalan toplam alan 1'e eşittir, ortalamanın sağında ve solunda kalan alanlar 0.5'tir (toplam alanın yarısı).
- Standart sapma küçüldükçe eğri sivrilir, standart sapma büyüdükçe eğri daha yayvan bir hal alır.



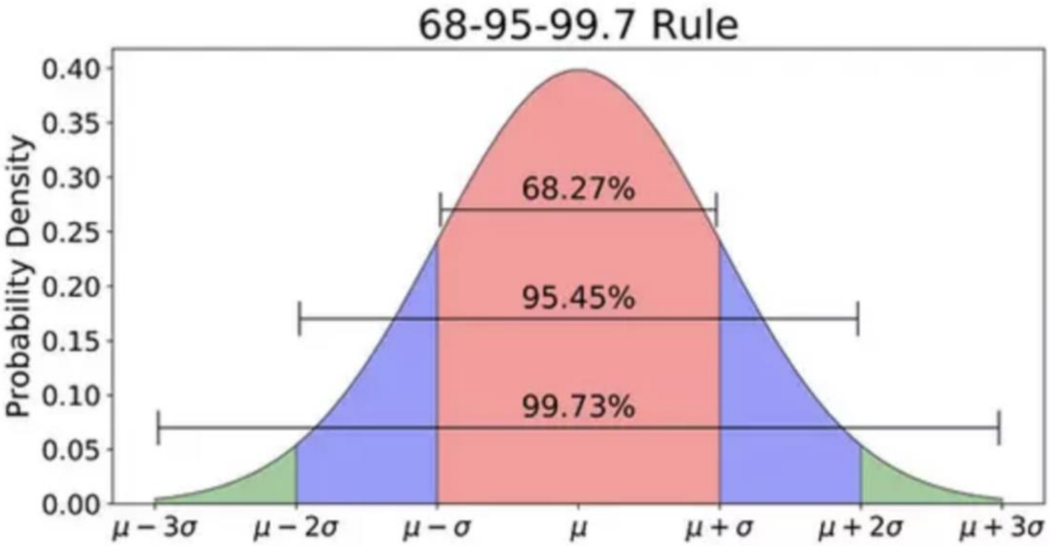
Normal Dağılım

- X'in a ile b arasında bulunma olasılığı P(a≤X≤b), [a,b] aralığında ortalaması μ ve standart sapması σ olan normal eğrinin altında kalan alana eşittir.
- Bu tür bir alanı formül ile hesaplamak daha zor olacağı için genelde tablolardan yararlanılır.



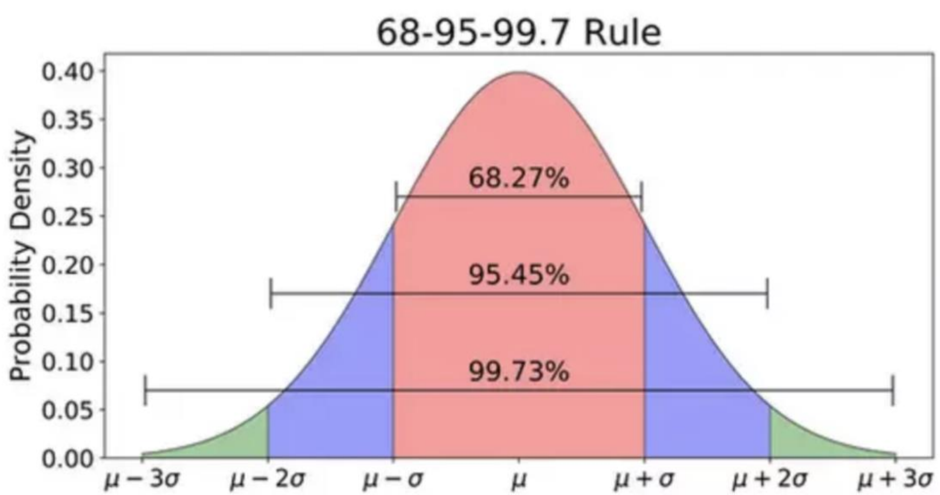
Normal Dağılım

- Eğri altında kalan alanın tablo ile hesaplanması için eğri altında kalan üç önemli bölgenin bilinmesi gerekmektedir.
- X'in gözlenen tüm değerlerinin %68.26'sı ortalamanın bir standart sapması (artı yada eksi) içinde yer alır, %95.45'i ortalamanın iki standart sapması içinde yer alır, %99.73'ü ortalamanın üç standart sapması içinde yer alır.



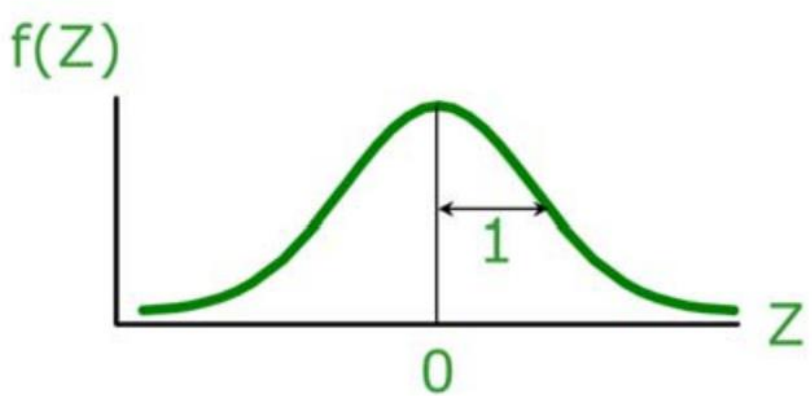
Normal Dağılım

- Standart sapma değerinden bağımsız olarak örneklerin dağılımı her zaman bu kurala uyar.
- Normal dağılım hesaplarının daha kolay yapılabilmesi için normal rassal değişken standartlaştırılır.
- $z = \frac{x-\mu}{\sigma}$ değeri, x'in ortalamanın kaç standart sapma uzağında olduğunu ifade eder.

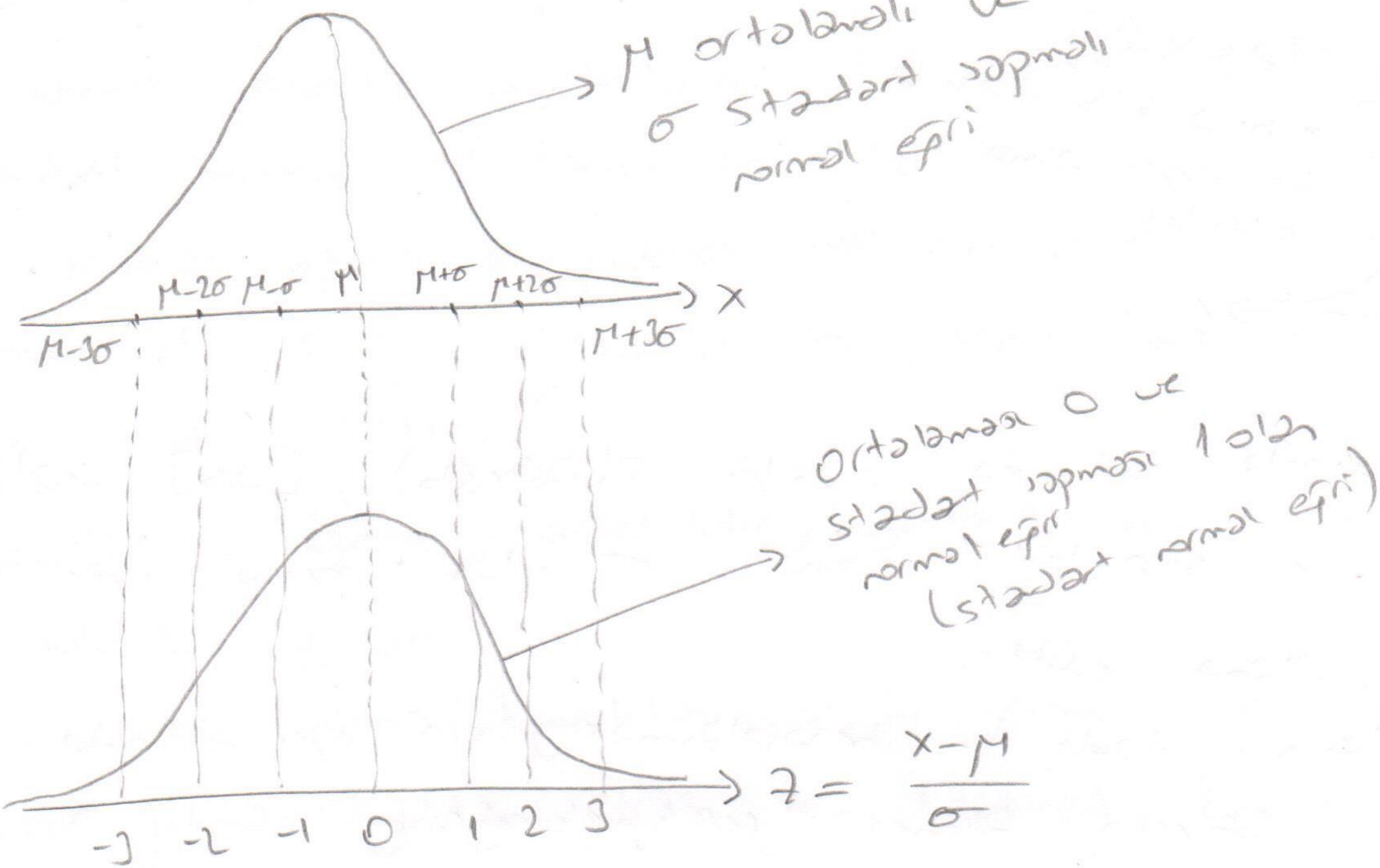


Standart Normal Dağılım

- X tesadüfi değişkeni ortalaması μ ve standart sapması σ olan bir normal dağılımdan geliyorsa, $z = \frac{x-\mu}{\sigma}$ tesadüfi değişkeni, ortalaması 0 ve standart sapması 1 olan normal dağılımı gösterir.
- Ortalaması 0 ve standart sapması 1 olan normal dağılım, standart normal dağılım olarak adlandırılır.
- Rassal değişken standartlaştırıldıktan sonra tablodan değeri okunur.



Standart Normal Dağılım

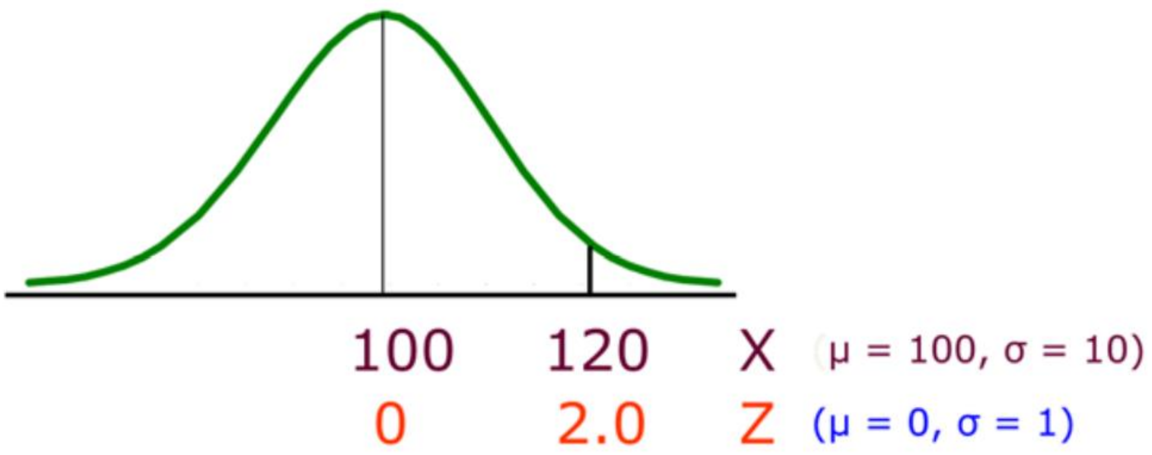


Z dönüşümü örnek

- Eğer bir X değişkeni 100 ortalama ve 10 standart sapma ile normal dağılıyor ise X=120 değeri için standart normal rassal değişken Z şu şekilde hesaplanır:

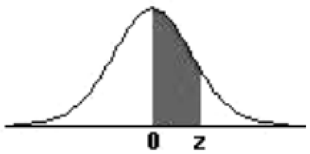
$$Z = \frac{x - \mu}{\sigma} = \frac{120 - 100}{10} = 2.0$$

- X=120 değeri, ortalamadan 2 standart sapma kadar sapmıştır.



Standart Normal Dağılım Tablosu

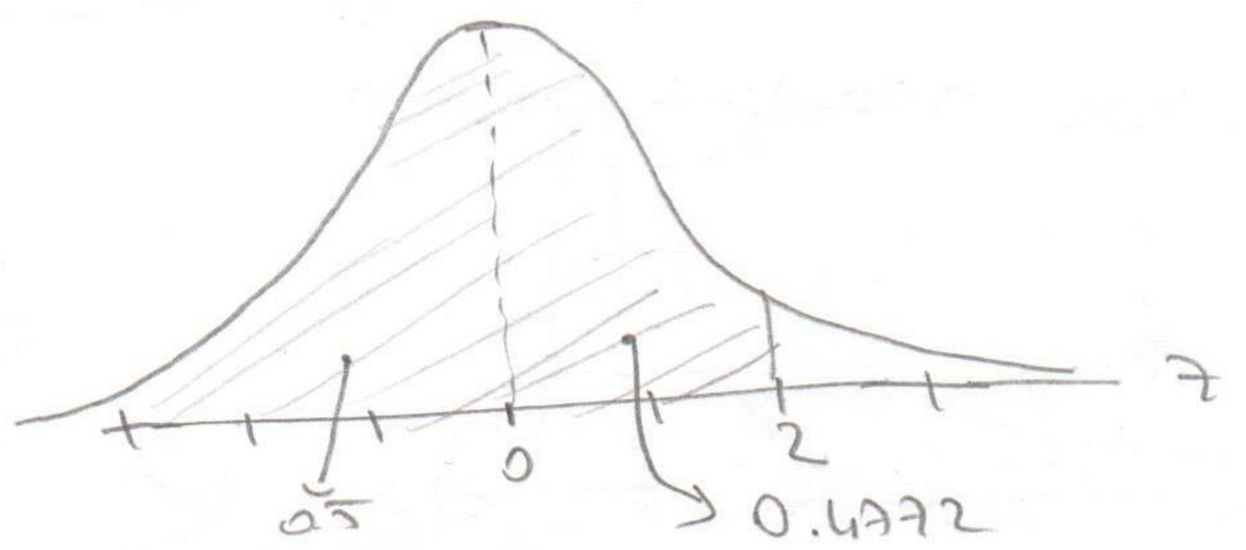
	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
⋮										
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990



Olasılıkların belirlenmesinde karşımıza çıkabilecek durumlar:

1. P(Z≤a) ve a≥0 ise tablodan okunan değere 0.5 eklenir

- P(Z ≤ 2)=?
- Tablodan Z=2.0 değerini oku, 0.5 ile topla

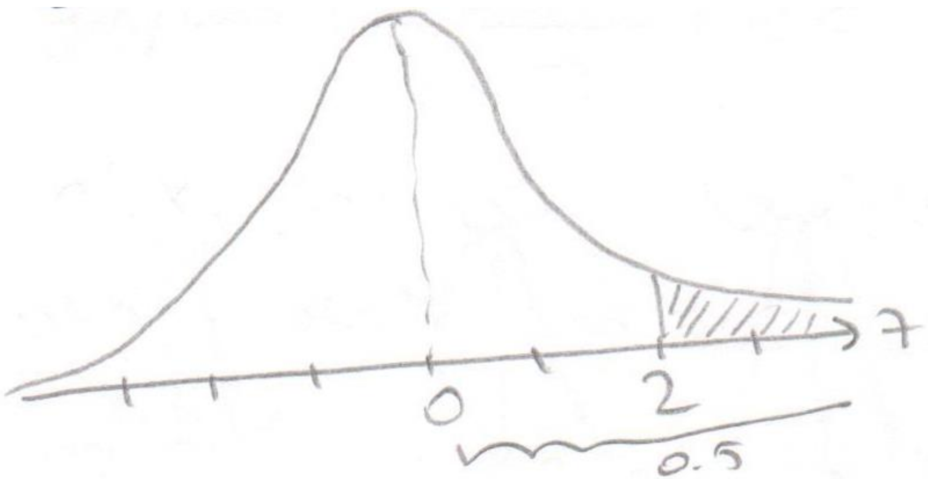


$$P(Z \leq 2) = 0.5 + 0.4772 = 0.9772$$

Olasılıkların belirlenmesinde karşımıza çıkabilecek durumlar:

2. P(Z≥a) ve a≥0 ise tablodan okunan değer 0.5 ten çıkartılır.

- P(Z ≥ 2)=?
- Tablodan Z=2.0 değerini oku, okuduğun değeri 0.5 ten çıkar

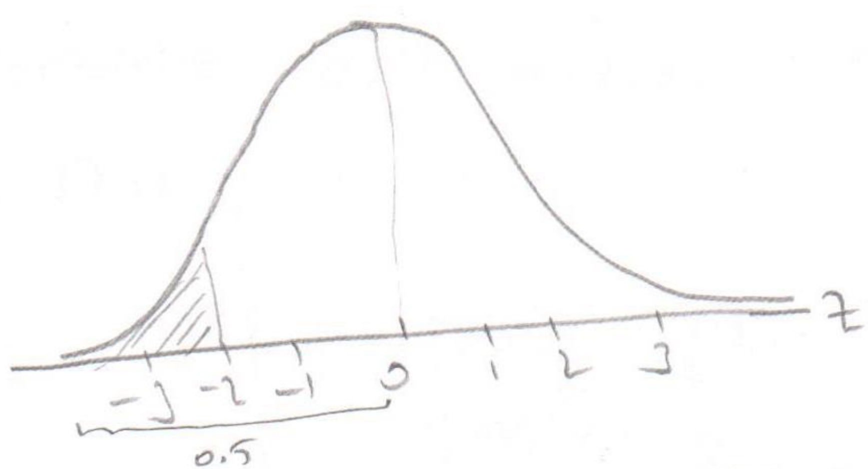


$P(Z \geq 2) = 0.5 - 0.4772 = 0.0228$

Olasılıkların belirlenmesinde karşımıza çıkabilecek durumlar:

3. P(Z≤a) ve a ≤ 0 ise simetri özelliğinden dolayı, negatif a'dan küçük olmak, pozitif a'dan büyük olmaya eşittir. Bu nedenle çözüm 2. maddedeki gibi yapılmalıdır.

- P(Z ≤ -2)=?
- Tablodan Z=2.0 değerini oku, okuduğun değeri 0.5 ten çıkar

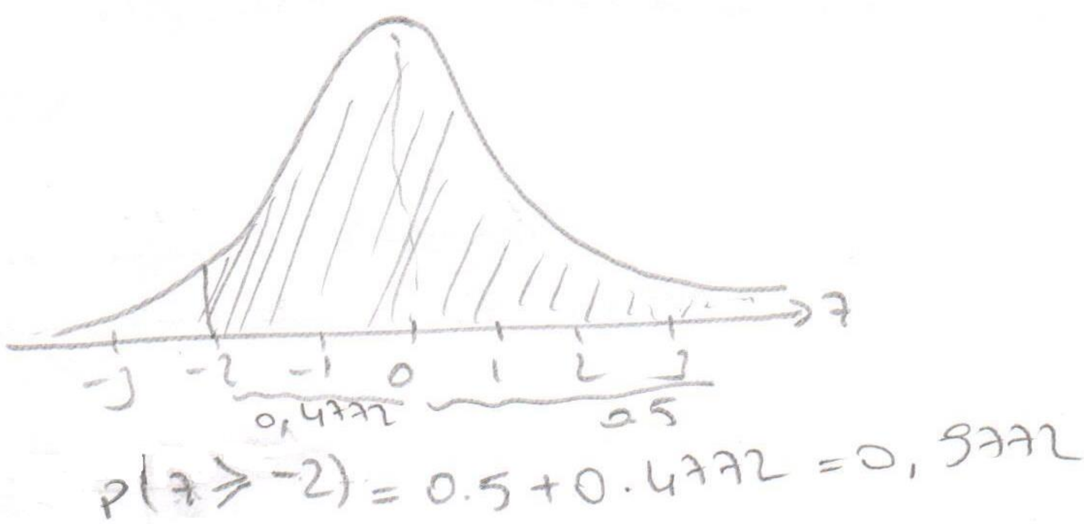


$$P(Z \leq -2) = 0.5 - 0.4772 = 0{,}0228$$

Olasılıkların belirlenmesinde karşımıza çıkabilecek durumlar:

4. P(Z≥a) ve a ≤ 0 ise simetri özelliğinden dolayı, negatif a'dan büyük olmak, pozitif a'dan küçük olmaya eşittir. Bu nedenle çözüm 1. maddedeki gibi yapılmalıdır.

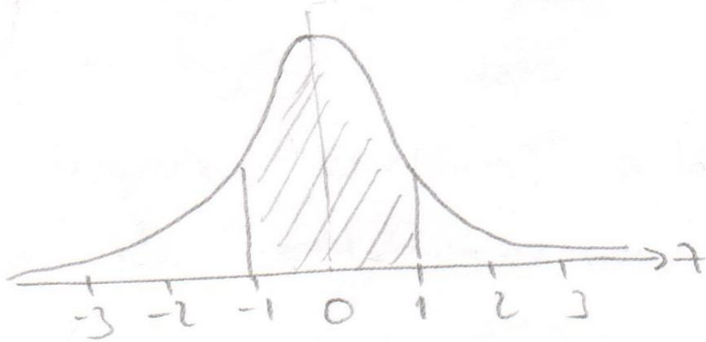
- P(Z ≥-2)=?
- Tablodan Z=2.0 değerini oku, okuduğun değere 0.5 ekle



Olasılıkların belirlenmesinde karşımıza çıkabilecek durumlar:

5. P(a≤Z≤b) olasılığı hesaplanırken P(Z≤b) olasılığı hesaplanıp bu değerden P(Z≤a) olasılığı çıkartılır. Hesaplamalar negatiflik ve pozitiflik durumuna göre ilk 4 seçenekteki gibi yapılır.

➢ P(-1 ≤ Z ≤ 1)=?



$$P(-1 \leq Z \leq 1) = 0.3413 + 0.3413 = 0,6826$$

ya da

$$P(Z \leq 1) - P(Z \leq -1) = \{0{,}5 + 0{,}3413\} - \{0{,}5 - 0{,}3413\}$$

$$= 0.8413 - 0.1587 = 0,6826$$

Örnek

Standart normal dağılıma uyan rassal bir değişkenin

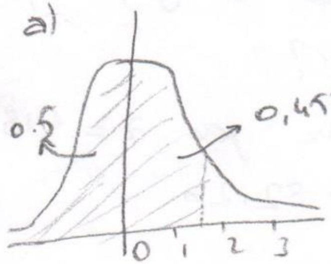
a) 1.72'den küçük

b) −0.88'den küçük

c) 1,30 ile 1.75 arasında

d) −0.25 ile 0.45 arasında bir değer alma olasılıklarını bulun

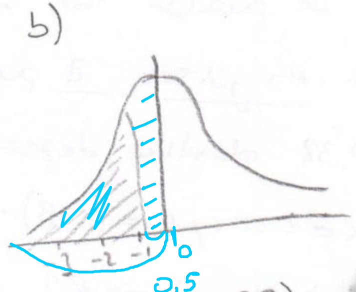
a)



$P(z \leq 1.72) = 0,5 + 0,4573$

$= 0,9573$

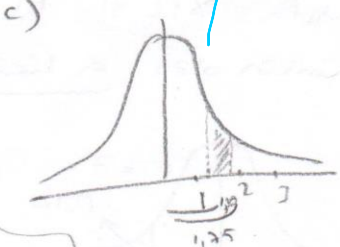
b)



$P(z \leq -0.88) = 0,5 - P(z \geq -0.88)$

$0,5 - 0,3106 = 0,1894$

c)



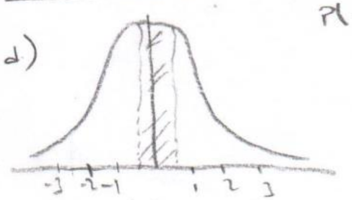
$P(1,30 \leq z \leq 1.75)$

$P(z \leq 1,75) - P(z \leq 1,30)$

$= 0,4599 - 0,4032$

$= 0,0567$

d)



$P(-0,25 \leq z \leq 0.45) = P(-0,25 \leq z) + P(z \leq 0,45)$

$= 0,0987 + 0,1736$

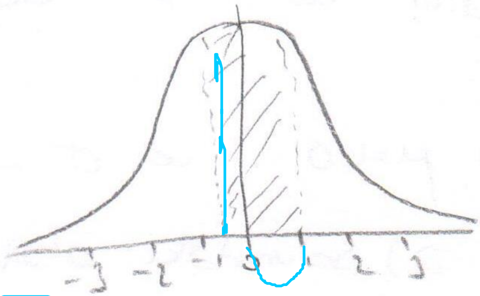
$= 0,2723$

Örnek

Bir bilgisayarın montajlaması, işi 50 dakika ortalamalı ve 10 dakika standart sapmalı normal dağılıma uymaktadır. Herhangi bir bilgisayarın 45 ile 60 dakika arasında monte edilmesi ihtimali nedir?

$$P(45 < X < 60) = ?$$

$$z = \frac{X-\mu}{\sigma}$$



$$P\left(\frac{45-50}{10} \leq \frac{X-\mu}{\sigma} \leq \frac{60-50}{10}\right) = P(-0{,}5 \leq z \leq 1)$$

$$P(-0.5 \leq z \leq 1) = P(0 \leq z \leq 0.5) + P(0 \leq z \leq 1)$$

$$= 0.1915 + 0.3413 = 0{,}5328$$

Normal eğriyi ters biçimde kullanma

Bazen tabloda istenen değerler arasına düşen belirli olasılıklara karşılık gelen z değerini bulmaya ihtiyaç duyarız.

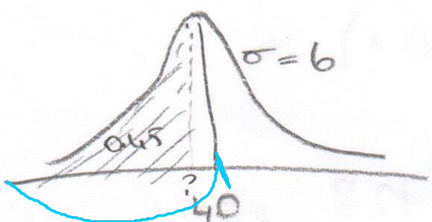
Bilinen bir alan yada olasılık değeri kullanılarak z değeri bulunur. $z = \frac{x - \mu}{\sigma}$ formülünden $x = z \cdot \sigma + \mu$ formülü elde edilir ve x değerine ulaşılır.

Örnek

$\mu=40$ ve $\sigma=6$ olan bir normal dağılım için $z=\frac{x-\mu}{\sigma}$

a) Solundaki alan %45

b) Sağındaki alanı %14 olan x değerini bulunuz.

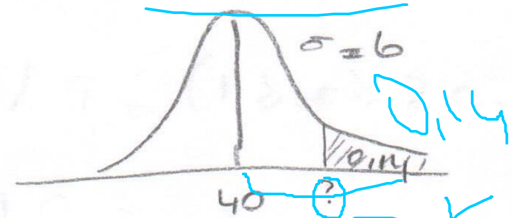


$0.5-0.45=0.05$ olasılık değerine sahip z değerini bulalım.

Tabloda 0.0517 değerine karşılık gelen z değeri 0,13 $(z=-0,13)$

$x=z\cdot\sigma+\mu=(-0,13)\cdot 6+40$

$x=39,22$



$0,5-0,14=0,36$ olasılık değerine karşılık gelen z'i tabloda bul

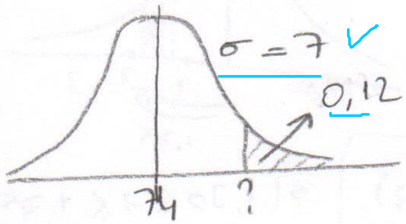
0.3599 olasılık değerine karşılık gelen $z=1,08$

$x=z\cdot\sigma+\mu=(1.08)\cdot 6+40$

$x=46,48$

Örnek

Bir sınav için ortalama puan 74 ve standart sapma 7 dir. Sınıfın %12 si A alıyorsa ve puanlar bir normal dağılıma uyuyorsa, mümkün olan en küçük A ve en yüksek B puanı nedir?



$0,5 - 0,12 = 0,38$

0,38 olasılık değerine karşılık gelen $z = 1,18$

$x = z \cdot \sigma + \mu = (1,18) \cdot 7 + 74 = 82,26$

en düşük A → 83

en yüksek B → 82

(Notların tamsayı olduğu kabul ediliyor)

Kaynaklar

- Dr. Halil İbrahim CEBECİ, İstatistik Ders Notları, Sakarya Üniversitesi
- Mühendisler ve Fen Bilimciler için Olasılık ve İstatistik, Ronald E. Walpole, Çeviri Editörü: Prof. Dr. M. Akif BAKIR, 9. baskıdan çeviri, Palme Yayıncılık, 2016.
- İstatistiksel Formüller ve Tablolar, Başkent Üniversitesi İktisadi ve İdari Bilimler Fakültesi, 2005