Rastgele Değişkenlerin Dağılımları

Kesikli Olasılık Dağılımları

- Kesikli düzgün dağılım
- Bernoulli dağılımı
- Binom dağılımı
- Poisson dağılımı
- Hipergeometrik dağılım
- Negatif binom dağılımı

Sürekli Olasılık Dağılımları

- Sürekli düzgün dağılım
- Normal dağılım
- Üstel dağılım
- Lognormaldağılım
- Gamma dağılımı
- Ki-kare dağılımı

Kesikli Olasılık Dağılımları

Bernoulli Dağılımı:

İki sonucu olan olayların olasılığının hesaplanmasında kullanılır.

Örneğin;

- Paranın atılması
- Kusurlu ve kusursuz parçaların bulunduğu kutudan bir parçanın çekilmesi
- Öğrencinin bir derste başarılı veya başarısız olması

p^x

Bernoulli Dağılımı

Tanım: X, rastgele değişkeni 0 ve 1 değerlerini alsın. X in olasılık fonksiyonu:

P(X=1)=p

P(X=0)=1-p=q

veya

f(x)=P(X=x)=p^x(1-p)^1-x   x=0,1

dir. Bu dağılıma Bernoulli dağılımı denir.

p^x

Bernoulli dağılımının ortalaması (X in beklenen değeri) ve varyansı ise aşağıdaki formüllerle hesaplanır:

$\mu=E(X)=p$

$\sigma^2=p.q$

Örnek: Ardışık üç Bernoulli denemesinde P(üçü de başarılı)=0,008 ise üçünün de başarısızlık olma olasılığını bulunuz.

Binom dağılımı

Bernoulli denemelerinden oluşan bir deneydir.

Binom dağılımında:

1. Deney n özdeş denemeden oluşur.
2. Her bir deneme için başarılı ve başarısızlık olmak üzere yalnız iki sonuç vardır.
3. Deneyler birbirinden bağımsızdır.
4. Başarı olasılığı (p) ve başarısızlık olasılığı q=1-p dir.
5. n deneyde elde edilen başarılı sonuçlar x değişkenine atanır.

Aşağıda verilen deneylerde tanımlanan X, binom rasgele değişkenidir.

- Bir para 10 kez atılsın. X, Tura gelme sayısı
- İçinde 8 siyah ve 4 beyaz top bulunan bir kavanozda tekrar yerine koyarak 3 top çekilsin. X, Çekilen siyah topların sayısı.
- İçinde 3 kusurlu ve 7 kusursuz parça bulunan bir kutudan tekrar yerine koyarak 4 parça seçilsin. X, Seçilen kusurlu parçaların sayısı.

Binom dağılımının olasılık fonksiyonu

$$P(n,X,p) = \left(\frac{n!}{(n-X)!\,X!}\right) p^{X}\, q^{n-X}$$

n: Deneyin tekrarlanma sayısı
X: İstenen sonuç sayısı
p: İstenen başarılı sonucun olasılığı
q: Başarısızlık olasılığı

Binom dağılımının ortalaması (X in beklenen değeri) ve varyansı ise aşağıdaki formüllerle hesaplanır:

$$\mathbf{E(X) = \mu = np}$$

$$\mathbf{V(X) = \sigma^2 = npq}$$

Örnek: a) Bir para 10 defa atılsın. 4 yazı gelme olasılığını hesaplayınız

b)Bir zarın 20 kez atılması durumunda tam 12 kez altı gelme olasılığını hesaplayınız.

Çözüm:

a) Binom dağılımın uygun olduğu rastgele olaylarda başarılı ve başarısız olarak iki durumun olduğu olaylarla ilgilenildiğinden:

başarılı: yazı gelmesi (p=0.5)

başarısız: yazı gelmemesi (q=0.5)

olarak tanımlama yapılabilir. n=10;X=4 olduğundan istenilen olasılık:

$$P(n,X,p) = P(10,4,0.5) = \binom{n}{X} p^X q^{n-X}$$

$$\binom{10}{4} 0.5^4 0.5^6 = 0.205$$

b) başarılı: 6 gelmesi (p=1/6)

başarısız: 6 gelmemesi (q=5/6)

olarak tanımlama yapılabilir.

n=20;X=12 olduğundan istenilen olasılık:

$$P(n, X, p) = P(20,12,1/6) = \binom{20}{12}(1/6)^{12}(5/6)^{8} = 0.0000135$$

Poisson Dağılımı

Verilmiş bir zaman aralığında bir alanda ya da hacimde başarıların sayısı X rasgele değişkeni olsun. Aşağıdaki koşulları sağlayan X e Poisson rasgele değişkeni denir.

- Deney, verilmiş bir zaman, alan ya da hacimde bir olayın (başarıların) elde ediliş sayılarının sayılmasıyla oluşur.
- İki ayrık birim zamanda alanda ya da hacimde elde edilecek başarıların sayısı birbirinden bağımsızdır.
- Bir birim zaman aralığı, alan ya da hacimdeki başarı olasılığı tüm birimler için aynıdır.
- Çok küçük bir zaman aralığı, alan ya da hacimde iki ya da daha çok başarının olması hemen hemen imkansızdır. Yani, bu durumda birden çok başarı olması olasılığı sıfıra yaklaşır.
- Bir birim zaman aralığı, alan ya da hacimde bir sonucun ortalama elde ediliş sayısı μ dır.

Örneğin;

- Bir hava alanına her saat inen uçak sayısı,
- belirli bir trafik noktasında meydana gelen aylık trafik kazası sayısı,
- Bir üretim malındaki kusur sayısı,
- 1 cm^3 kandaki anormal hücre sayısı,......vb

Tanım: (Poisson dağılımı) X, 0,1,2,... değerlerini alabilen bir Poisson rasgele değişkeni olsun. X in olasılık fonksiyonu

$$P(X) = \frac{e^{-\mu}\,\mu^{X}}{X!}$$

Örnek: 200 sayfalık bir kitaba 200 yazım hatası rasgele dağıtılıyor. Rasgele seçilen bir sayfada

a) İki
b) İkiden az

yazım hatası bulunma olasılığı nedir?

Örnek: Saat 9.00 dan 9.05 e kadar bir operatörün aldığı telefon konuşmalarının sayısı $\mu=2$ olan Poisson dağılımına sahipse sonraki gün aynı zaman aralığında operatörün telefon konuşması almaması olasılığı nedir?

Sürekli Rasgele Değişkenlerin Dağılımları

- Sürekli düzgün dağılım
- Normal dağılım
- Üstel dağılım
- Lognormal dağılım
- Gamma dağılımı
- Ki-kare dağılımı

Sürekli düzgün dağılım

X sürekli rastgele değişken belirli bir aralıktaki her değerinin meydana gelme olasılığı eşit ise bu rastgele değişkenin dağılımı düzgün (Ünifrom) dağılımdır. Ünifrom dağılıma ait olasılık fonksiyonu:

$a<=x<=b$ olmak üzere

$$f(x)=f(x;a,b)=1/b-a$$

dır.

NORMAL DAĞILIM

Sürekli olasılık dağılımlarının en önemlisi ve en çok kullanılan dağılımdır. Normal dağılıma Gauss Laplace dağılımı, grafiğine normal eğri denir. Evrendeki birçok olay normal dağılıma uygunluk gösterdiğinden yapılan araştırmalarda elde edilen verilerin değerlendirilmesinde çok yaygın olarak kullanılmaktadır.

Normal dağılımın olasılık yoğunluk fonksiyonu

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

Normal yoğunluk fonksiyonu iki parametreye sahiptir:

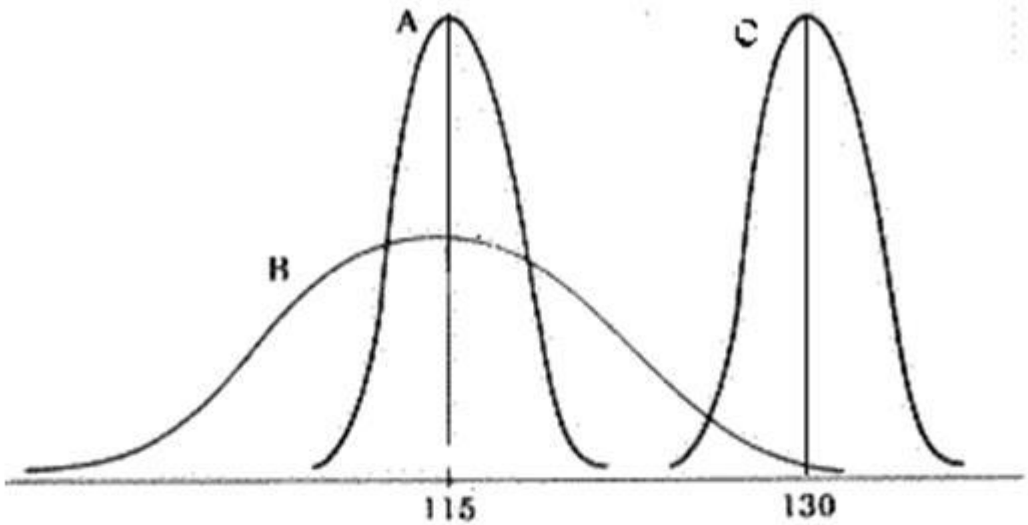
- ortalama μ
- standart sapma σ

Normal dağılım X ~N(μ ; σ^2) simgesiyle gösterilir.

Ortalama ve standart sapma değerlerine bağlı olarak Normal dağılımın yeri ve biçimi değişmektedir.

Örneğin:

Aşağıda şekilleri verilen A, B ve C normal dağılmış rastgele değişkenler arasında:



$$\mu_A = \mu_B < \mu_C$$

$$\sigma_A^2 = \sigma_C^2 < \sigma_B^2$$

Tanım: X rasgele değişkeni μ ve σ^2 varyanslı normal dağılıma sahipse, X in a ve b değerleri arasında bulunması olasılığı

$$P(a \le X \le b) = \int_a^b f(x)dx = \int_a^b \left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(X-\mu)^2}{\sigma^2}} \right) dx$$

Bu eşitlik, $f(x)$ eğrisi altında, x ekseni üzerinde ve $x=a$, $x=b$ doğruları arasında kalan alandır.

Normal dağılım aşağıdaki özelliklere sahiptir:

1. Normal dağılım $x=\mu$ doğrusuna göre simetriktir.

Yani normal dağılımın grafiği $x=\mu$ doğrusunun sağında ve solunda aynıdır.

2. Ortalama μ , ortadadır ve alanı iki eşit parçaya ayırır.

3. $f(x)$ eğrisi altında, x ekseni üzerinde kalan alan 1 e eşittir.

Standart normal dağılım

Standart normal dağılım: ortalaması μ=0 ve varyansı 1 ve olasılık fonksiyonu, $Z = \frac{X - \mu}{\sigma}$ olmak üzere

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

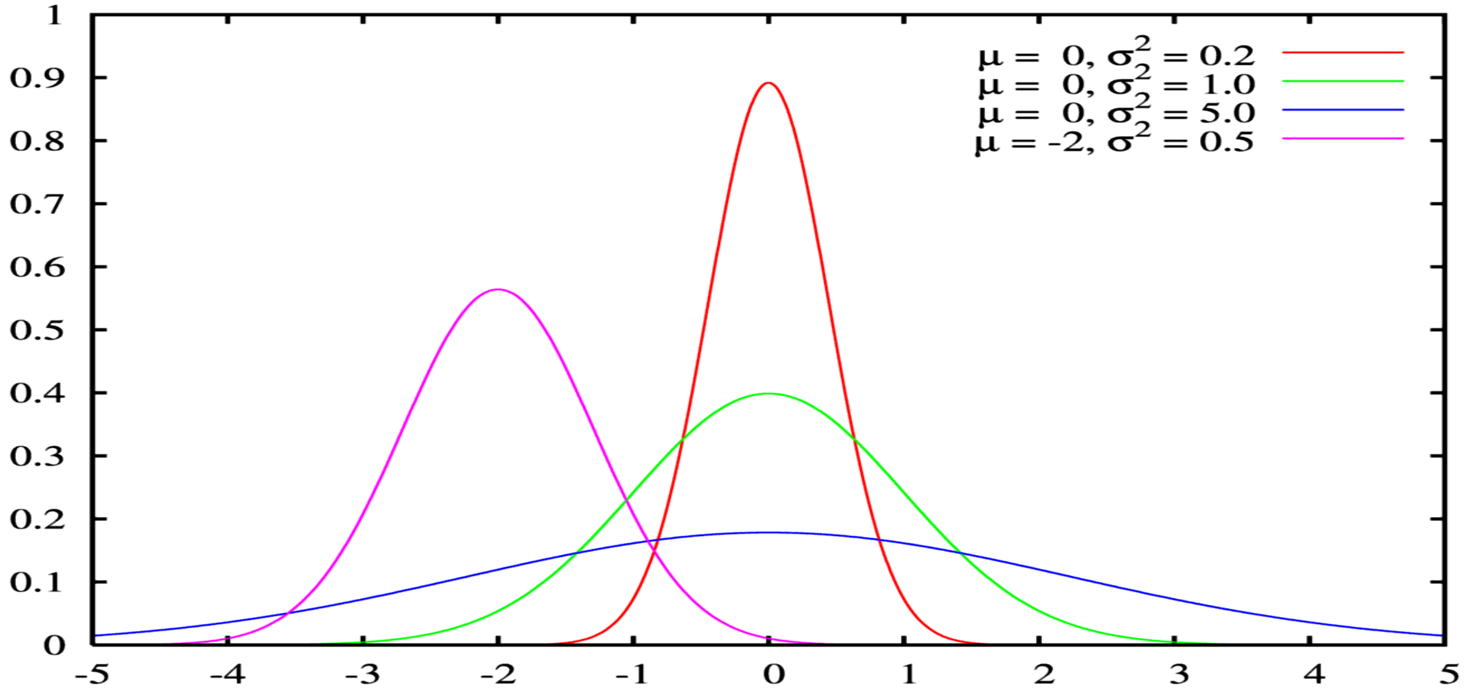
olan normal dağılıma standart normal dağılım denir ve bu dağılım X ~N(0;1) olarak gösterilir.

Tanım: X (μ=0 ve varyansı 1) normal dağılımına sahipse X in a ve b arasında bulunma olasılığı

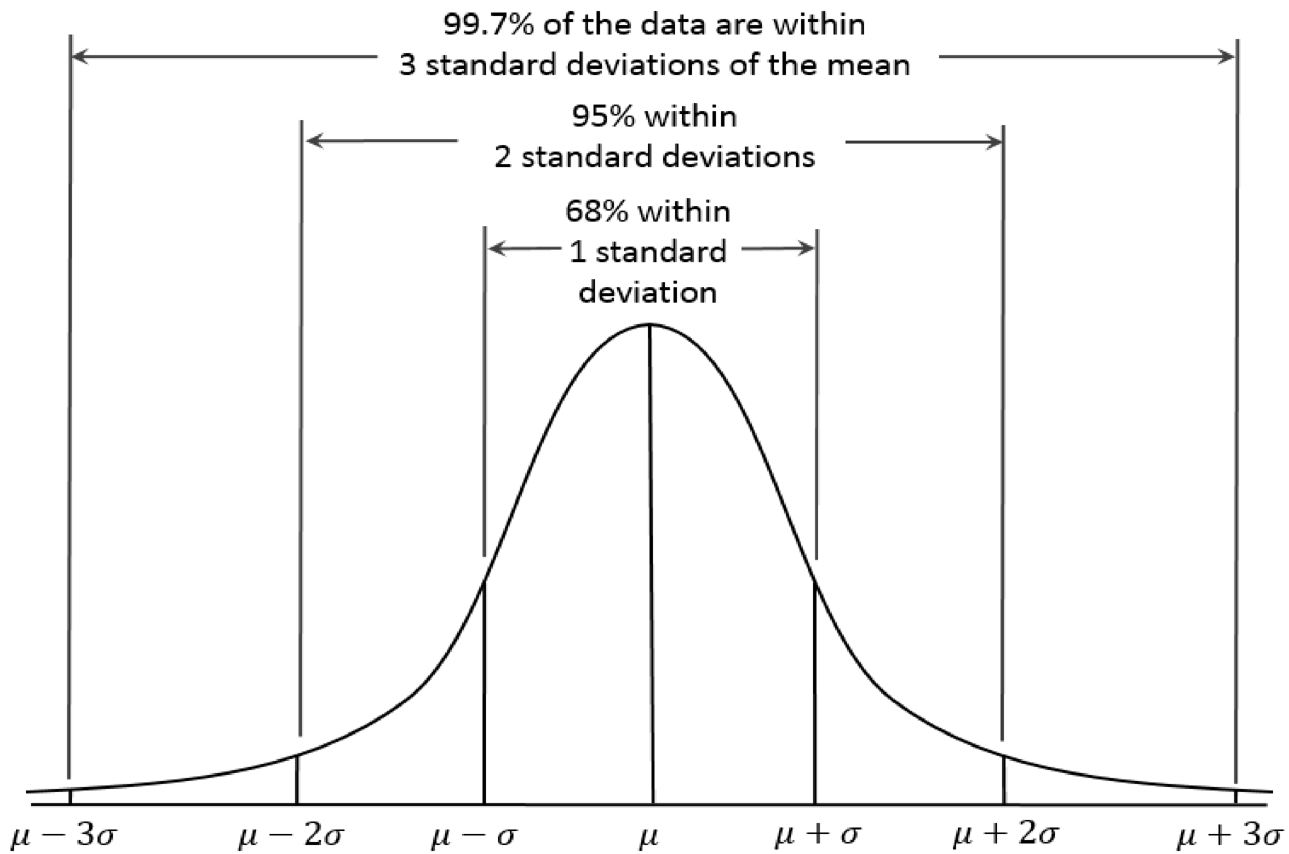
$$P(0 \leq Z \leq z_0) = \int_0^{z_0} f(z)dz = \int_k^{z_0} \left(\frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} \right) dz$$

ve

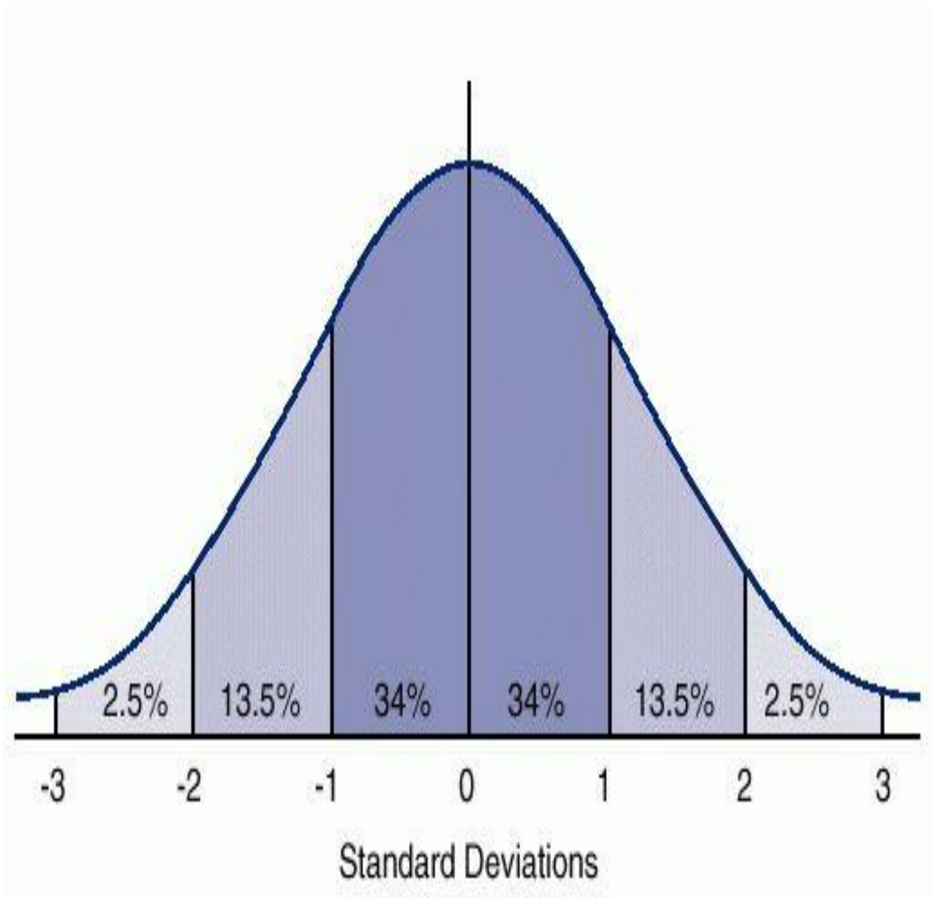
$$P(Z \leq z_0) = \int_{-\infty}^{z_0} f(z)dz = \int_{-\infty}^{z_0} \left(\frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} \right) dz$$



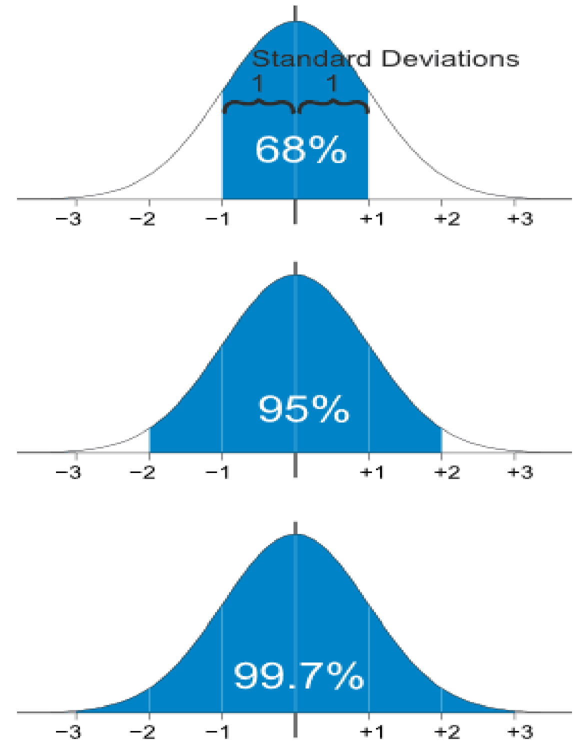
- Yeşil grafik standart normal dağılım



99.7% of the data are within
3 standard deviations of the mean
95% within
2 standard deviations
68% within
1 standard
deviation
$\mu - 3\sigma$
$\mu - 2\sigma$
$\mu - \sigma$
μ
$\mu + \sigma$
$\mu + 2\sigma$
$\mu + 3\sigma$



2.5%
13.5%
34%
34%
13.5%
2.5%
-3
-2
-1
0
1
2
3
Standard Deviations



Standard Deviations
1
1
68%
−3
−2
−1
+1
+2
+3
95%
−3
−2
−1
+1
+2
+3
99.7%
−3
−2
−1
+1
+2
+3

Örnek: Bir sınıftaki öğrencilerin %95 inin vize notu 50 ile 70 arasındadır. normal dağılıma sahip ise bu dağılımı bulunuz.

Çözüm:

Bu eşitlik, f(x) eğrisi altında, x ekseni üzerinde ve z=0, z=z0 doğruları arasında kalan alandır.
Bu integrallerin hesaplanmasıyla standart normal dağılımın altındaki olasılıkları veren tablo hazırlanmıştır.

- Verilen tablo yardımıyla normal dağılıma ait her türlü olasılık hesaplanabilmektedir.
- Ayrıca, dağılım simetrik olup dağılımın tepe noktasının yatay ekseni kestiği noktanın koordinatı sıfırdır (dağılımın ortalamasıdır) ve eğri altında kalan alanın değeri 1’e eşittir.
- Dağılım simetrik olduğu için $P(Z > 0) = P(Z < 0) = 0.5$ dir.
- Bu nedenle, ortalamanın sağında kalan kısmı tablolarda verilmekte, diğer yarısının aynı olduğu bilinmektedir.

Olasılık tablolarının okunuşu

Z tablosundan istenilen olasılık değeri bulunulurken verilen değer;

1. tam sayı kısmı ile birinci ondalık kısmı
2. ikinci ondalık kısmı

olmak üzere iki parçaya ayrılır.

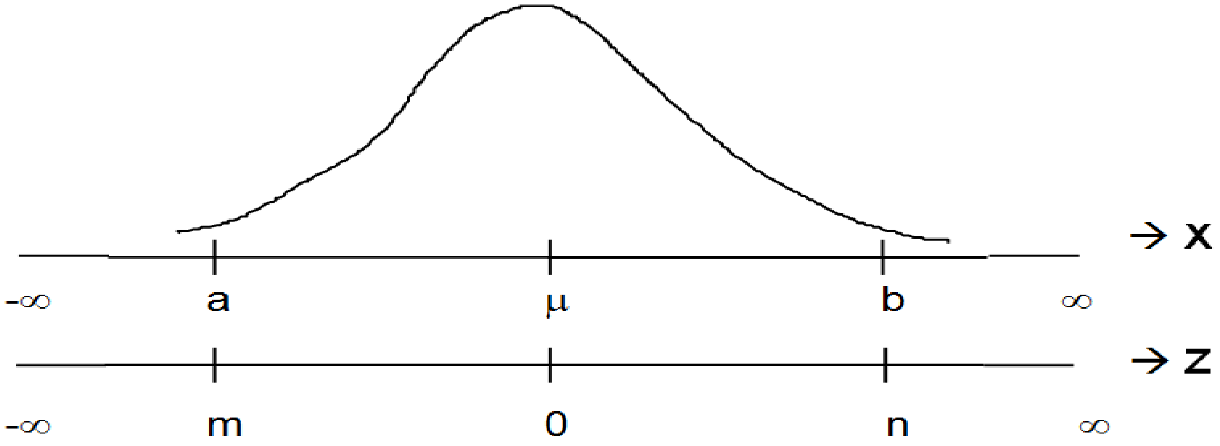
Z tablosundan bir olasılık değeri okumak için aşağıdaki adımlar takip edilir:

1. tamsayı kısmı ile birinci ondalık kısmı düşey eksende işaretlenir.
2. ikinci ondalık kısmı için yatay eksende eksende işaretlenir.
3. Bu değerlere yatay ve düşey eksende karşı gelen değerlerin kesiştiği hücredeki değer aranan olasılık değeridir.

- Verilen a < X < b aralığı m < Z < n aralığına dönüştürülür. Yani,

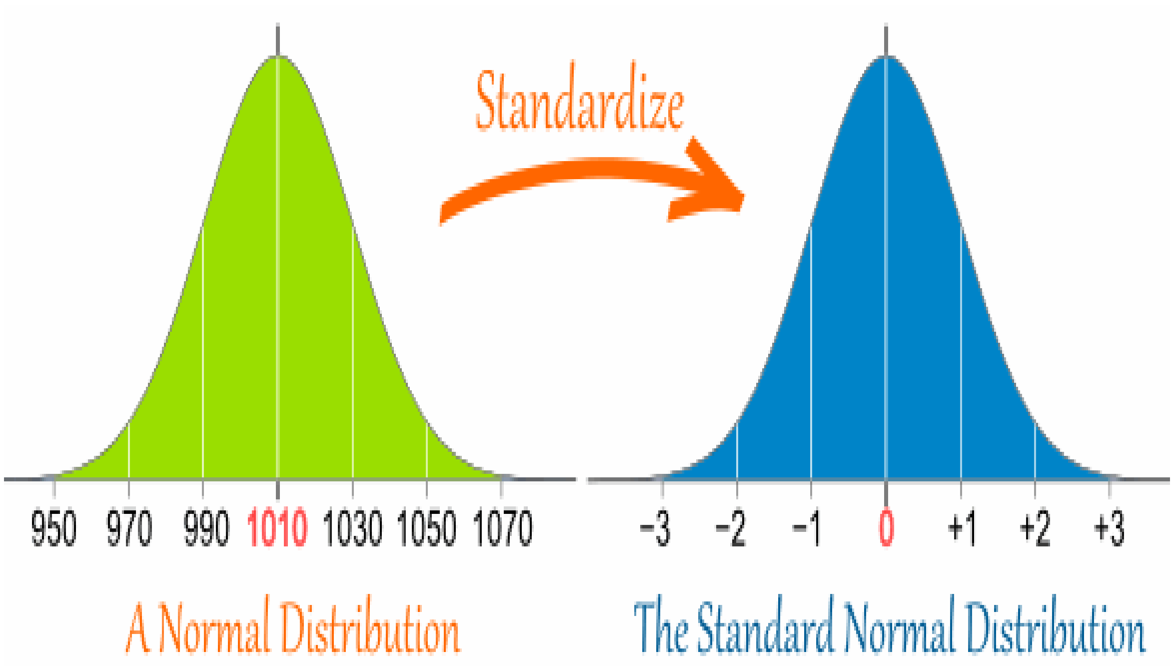
$$Z = \frac{X - \mu}{\sigma}$$

dönüşümü kullanılır.



Karşılık gelen P(m<Z<n) değeri tablo yardımıyla belirlenir. Öyle ise P(a<X<b):

$$P\left(\underbrace{\frac{a-\mu}{\sigma}}_{m} \leq \underbrace{\frac{x-\mu}{\sigma}}_{x} \leq \underbrace{\frac{b-\mu}{\sigma}}_{n}\right) \Rightarrow P(m \leq Z \leq n)$$


Standardize
950
970
990
1010
1030
1050
1070
A Normal Distribution
-3
-2
-1
0
+1
+2
+3
The Standard Normal Distribution

Örnek: X~N(μ=3 ve varyans 4) normal dağılımına sahipse, X in 3 ve 5 arasında bulunma olasılığı nedir?

Örnek: Eğer Z standart normal dağılmış bir rastgele değişken ise aşağıdaki olasılıkları grafiksel olarak gösterip hesaplayınız.

a) $P(0<=Z<=2)$

b) $P(-2<=Z<=2)$

c) $P(0<=Z<=1.50)$

d) $P(0.28 < Z < 1.28)$