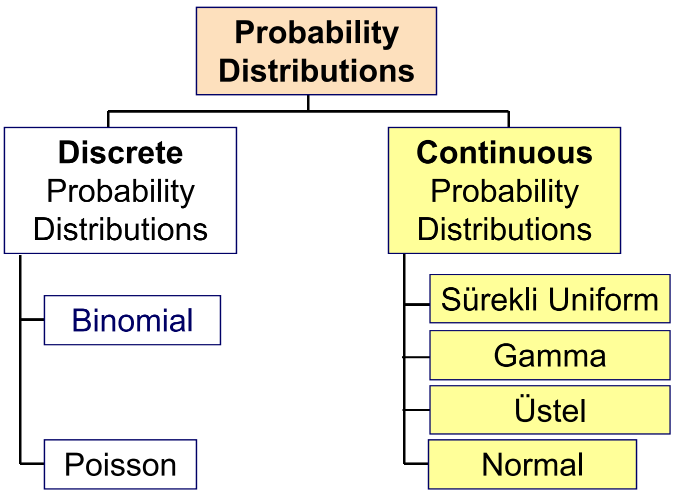
SÜREKLİ RASSAL DEĞİŞKENLER
ve
OLASILIK DAĞILIMLARI

Dr. Mehmet AKSARAYLI

Probability Distributions



Sürekli Şans Değişkenleri

- Sürekli bir aralıktaki tüm değerleri alabilen değişkenlerdir.
  - Bir yolun uzunluğu 25m < x < 50m olabilir.

x, D_x tanım aralığına sahip sürekli bir şans değişkeni olsun. f(x)'in x'e ait bir **olasılık yoğunluk fonksiyonu** (oyf) olabilmesi için;

Her x için olmalıdır.

$$\int_{D_x} f(x).dx = 1$$

$$\mu = E(x) = \int_{D_x} x.f(x).dx$$

$$V(x) = E[\,(X_i - \mu)^2\,] = E(X^2) - [E(X)]^2$$

3

Sürekli Uniform Dağılım

- a ve b gibi iki nokta arasından bir sayı seçmek istediğimizde herhangi bir değeri alabilecek x şans değişkeni sürekli uniform dağılışı göstermektedir.

- Sürekli üniform dağılımı ilgilenilen şans değişkeninin olasılık fonksiyonu hakkında bilgiye sahip olunmadığında ve verilen aralık içerisinde tanımlanan olayın eşit olasılıklarla ortaya çıkacağı varsayımı yapıldığında kullanışlıdır.

4

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{if } a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

f(x) = value of the density function at any x value

a = lower limit of the interval

b = upper limit of the interval

HESAPLAMA KOLAYLIĞI!! $$P(c \le x \le d) = \frac{d-c}{b-a}$$

The Mean and Standard Deviation

The mean (expected value) is:

$$E(x) = \mu = \frac{a+b}{2}$$

The standard deviation is

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

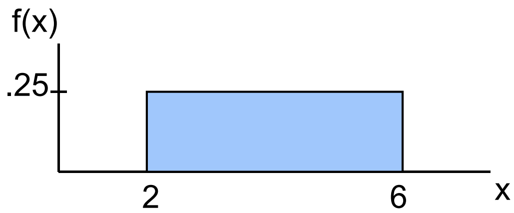
where

a = lower limit of the interval from a to b

b = upper limit of the interval from a to b

Example: Uniform Probability Distribution
Over the range $2 \le x \le 6$:

$$f(x) = \frac{1}{6-2} = .25 \quad \text{for } 2 \le x \le 6$$



Example: Uniform Probability Distribution
Over the range $2 \le x \le 6$:

$$E(x) = \mu = \frac{2+6}{2} = 4$$

$$\sigma = \sqrt{\frac{(b-a)^2}{12}} = \sqrt{\frac{(6-2)^2}{12}} = 1.1547$$

Örnek: Bir demir-çelik fabrikasında üretilen çelik levhaların kalınlıklarının 150 ile 200 mm arasında değiştiği ve bunların sürekli uniform şans değişkenine uygun olduğu bilinmektedir. Levha kalınlıkları 155 mm altında çıktığı zaman tekrar üretime gönderildiğine göre bu dağılımın beklenen değerini ve varyansını bulunuz
ve üretim sürecinde tekrar üretime gönderilen levhaların oranını bulunuz.

a) Bu dağılışın ortalama ve varyansı;
$E(x)=(150+200)/2 = 175$ mm
$Var(x)=(200-150)^2/12 = 208.33$ mm^2 bulunur.

b) Üretime geri döndürülen ürünlerin oranı ise;
$P(150 < x < 155) = (155-150) / (200-150) = 0,1$
Ürünlerin %10'u üretime geri gönderilmektedir.

ÜSTEL (EXPONENTIAL) DAĞILIM


t (zaman)

Zaman ekseninde belirli bir zaman aralığındaki olay sayısı Poisson, iki olay arasında geçen süre ise ÜSTEL dağılış gösterir.

Bir(ilk) olayın (r = 1) meydana gelmesine kadar geçen zamanın olasılığı ile ilgili dağılış üslü dağılıştır.

$$f(x) = \begin{cases} \alpha.e^{-\alpha x} & x \geq 0 \\ 0 & diger \end{cases}$$

α: <u>birim zamandaki</u> olay sayısı

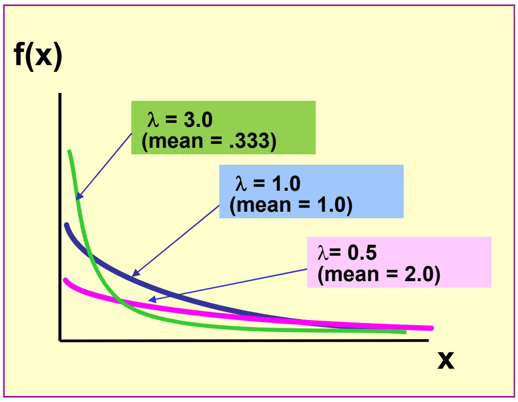
$$P(X \leq x) = \int_0^x \alpha.e^{-\alpha x} dx = 1 - e^{-\alpha x}$$

$$E(X) = \frac{1}{\alpha} \quad V(X) = \frac{1}{\alpha^2}$$

$$P(X > x) = e^{-\alpha x}$$

Exponential Distribution

- Shape of the exponential distribution



ÖRNEK

A marka televizyonun ömrü yıl olarak X şans değişkeni ile gösterilsin. X'in oyf:

$$f(x)=\begin{cases}\frac{1}{6}.e^{-\frac{1}{6}x} & x\geq 0\\ 0 & diger\end{cases}$$

$$E(X)=\frac{1}{\alpha}=\frac{1}{1/6}=6\ yil$$

$$V(X)=\frac{1}{\alpha^2}=\frac{1}{\left(1/6\right)^2}=36$$

Televizyonun ömrünün en az 6 yıl olması olasılığı nedir?

$$P(X>x)=\int_6^\infty \frac{1}{6}.e^{-\frac{1}{6}x}dx=e^{-\frac{1}{6}.6}=e^{-1}=0.367$$

Example

Example: Customers arrive at the claims counter at the rate of 15 per hour (Poisson distributed). What is the probability that the arrival time between consecutive customers is less than five minutes?

- Time between arrivals is exponentially distributed with mean time between arrivals of 4 minutes (15 per 60 minutes, on average)
- $1/\lambda = 4.0$, so $\lambda = .25$
- $P(x < 5) = 1 - e^{-\lambda a} = 1 - e^{-(.25)(5)} = \boxed{.7135}$

GAMMA DAĞILIŞI

- Bu olasılık dağılımının temel süreci poisson dağılımıdır. (r>1) tane olayın meydana gelmesine kadar geçen zamanın olasılığı, eğer olaylar poisson olayları ise, gamma dağılışına uyar. Meteorolojik verilerde yağış miktarının dağılışı gibi konularda gamma dağılışı kullanılır.

Dağılışın olasılık yoğunluk fonksiyonu

$$f(x) = \begin{cases} \frac{\alpha}{\Gamma(r)} . (\alpha . x)^{r-1} . e^{-\alpha . x} & x >= 0 \\ 0 & d.d \end{cases}$$

Dağılışın beklenen değeri

$$E(x) = r . \frac{1}{\alpha} = \frac{r}{\alpha}$$

Dağılışın varyansı

$$Var(x) = r . \frac{1}{\alpha^2} = \frac{r}{\alpha^2}$$

- Gamma dağılışı parametreleri r ve x’ dir.
- r: olay sayıları
- poisson dağılımındaki birim zaman parametresi
- x: zaman(değişkendir)
- α kere poisson olasılığı olarak da tanımlanabilir.

Burada $\Gamma(r)$ tanımı şöyle açıklanmaktadır;

$$\Gamma(r) = \int_0^\infty x^{r-1} . e^{-x} . dx$$

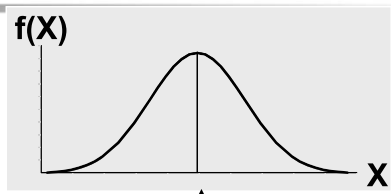
$$= (r-1).\Gamma(r-1)$$

$$= (r-1)!$$

$\Gamma\left(\frac{1}{2}\right) = \Gamma\pi$ özel bir durumdur.

Normal Dağılım

- Özellikleri:
- 1.‘Çan-Şekilli’ ve simetrik
- 2. Ortalaması, modu ve medyanı eşit
- 3.’Orta yayılımı’= 1.33 σ
- 4. Şans değişkeni sonsuz aralığa sahip



Ortalama
Mod
Medyan

Normal Dağılımın Önemi

❖ 1. Çoğu rassal süreçleri ve sürekli olayları tanımlar.

❖ 2. İstatistiksel yorumlamanın temelidir.

• Sürekli ve kesikli şans değişkenlerinin dağılımları birlikte ele alındığında istatistikte en önemli dağılım **Normal dağılımdır**.

• Normal dağılım ilk olarak **1733’te Moivre** tarafından p başarı olasılığı değişmemek koşulu ile binom dağılımının limit şekli olarak elde edilmiştir. **1774’te Laplace** hipergeometrik dağılımını limit şekli olarak elde ettikten sonra **19. yüzyılın ilk yıllarında Gauss** 'un katkılarıyla da normal dağılım istatistikte yerini almıştır.

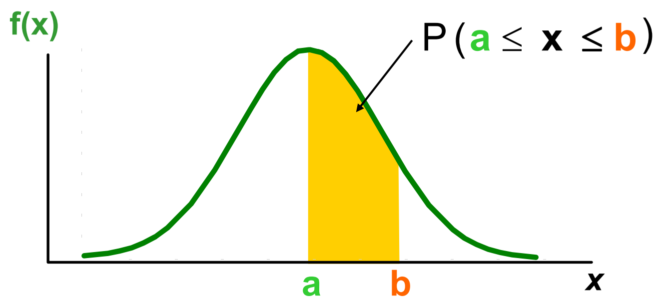
❖ Normal dağılımın ilk uygulamaları doğada gerçekleşen olaylara karşı başarılı bir biçimde uyum göstermiştir. Dağılımın göstermiş olduğu bu uygunluk adının Normal Dağılım olması sonucunu doğurmuştur.

❖ İstatistiksel yorumlamanın temelini oluşturan Normal Dağılım, bir çok rassal süreçlerin dağılımı olarak karşımıza çıkmaktadır.

❖ Normal dağılış kullanımının en önemli nedenlerinden biride bazı varsayımların gerçekleşmesi halinde kesikli ve sürekli bir çok şans değişkeninin dağılımının normal dağılışa yaklaşım göstermesidir.

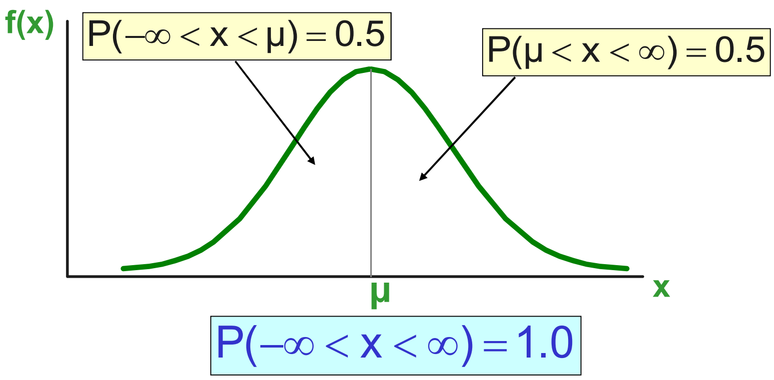
Olasılıkları Elde Etmek;

Olasılık eğrinin altındaki alan ile belirlenir.



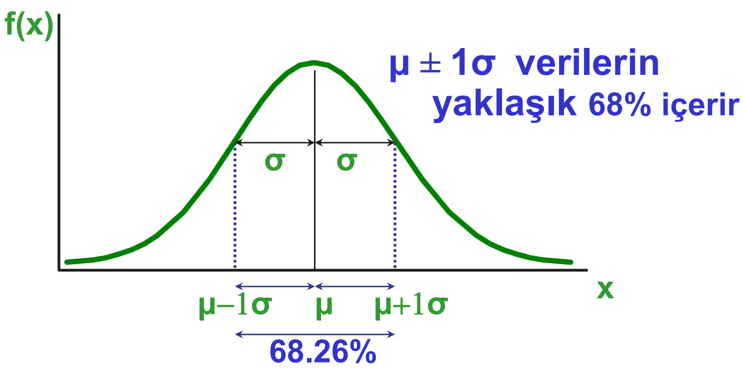
Probability as Area Under the Curve

The total area under the curve is 1.0, and the curve is symmetric, so half is above the mean, half is below



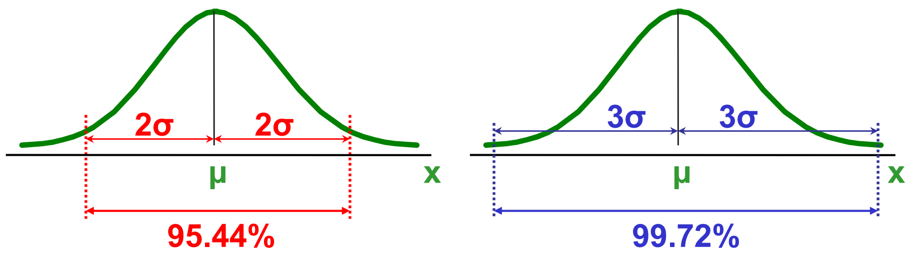
Empirical Rules

What can we say about the distribution of values around the mean? There are some general rules:



The Empirical Rule

- μ ± 2σ verilerin yaklaşık 95% içerir
- μ ± 3σ verilerin yaklaşık 99.7% içerir

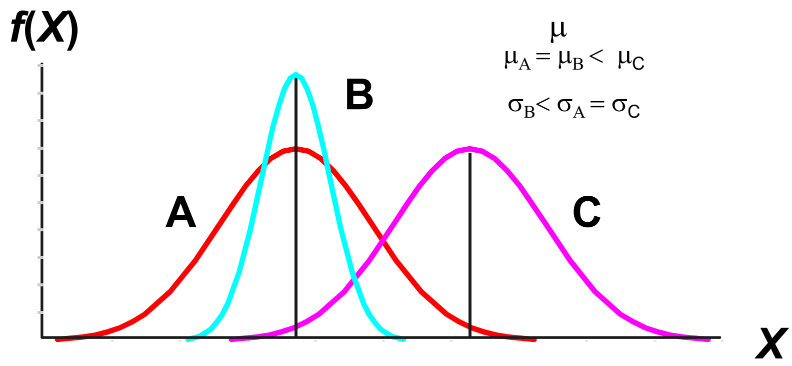


Normal Dağılım
Olasılık Yoğunluk fonksiyonu

$$f(X)=\frac{1}{\sqrt{2\pi}\cdot\sigma}e^{-\left(\frac{1}{2}\right)\cdot\left(\frac{(X-\mu)}{\sigma}\right)^2}$$

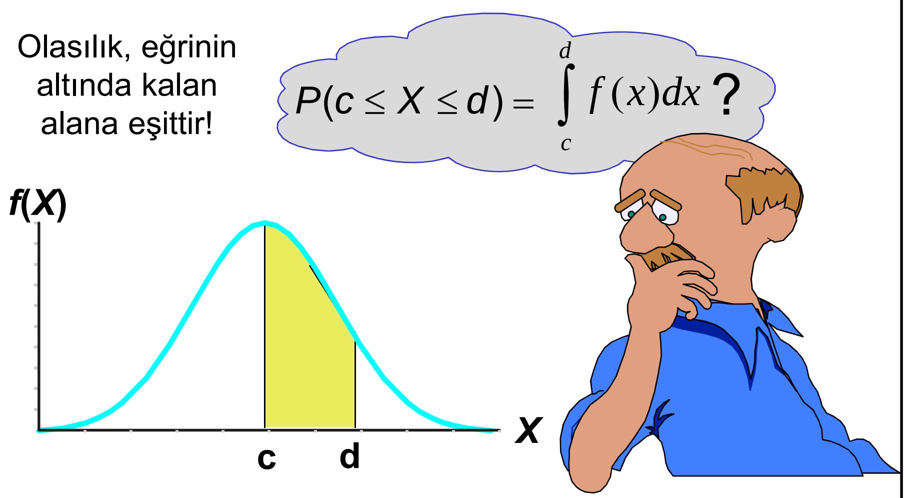
$f(X)$	=	X şans değişkeninin frekansı
π	=	3.14159; e = 2.71828
σ	=	populasyonun standart sapması
X	=	Şans değişkeninin değeri ($-\infty < X < \infty$)
μ	=	populasyon ortalaması

Parametrelerin Değişikliğinin Etkileri (μ & σ)



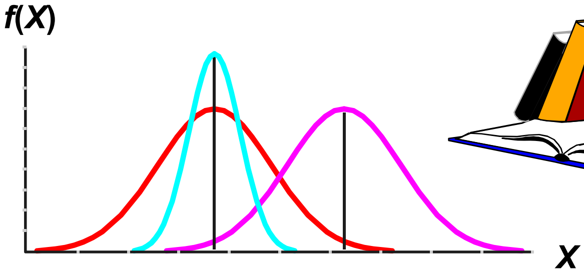
Normal Olasılık Dağılımı

Olasılık, eğrinin altında kalan alana eşittir!



Normal Dağılım Tablolarının Sonsuz Sayısı

Normal dağılımlar, ortalama ve standart sapma açısından farklılık gösterirler.



Her dağılım için bir tablo gerekir.

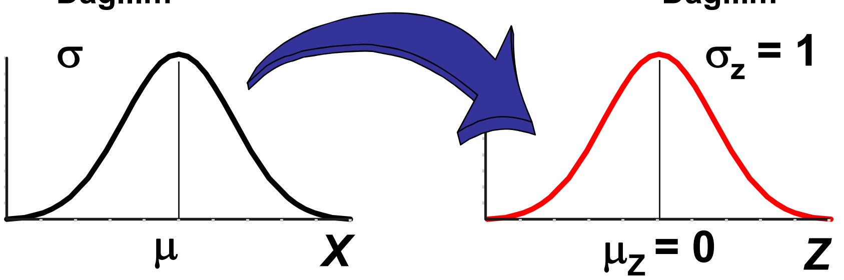
Bu da sonsuz sayıda tablo anlamına gelir!

Standart Normal Dağılım

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

Normal Dağılım

Standart Normal Dağılım

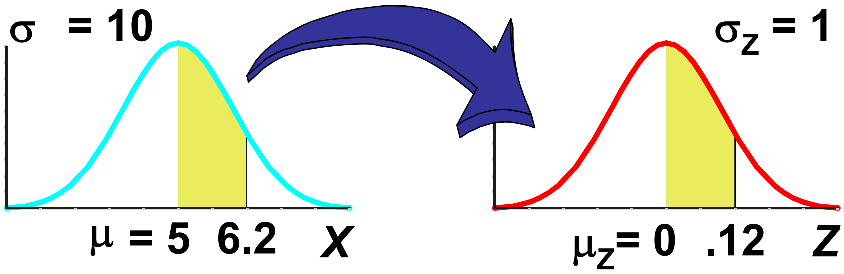


Artık tek tablo yeterli!!!

Standartlaştırma Örneği

$$Z = \frac{X-\mu}{\sigma} = \frac{6.2-5}{10} = 0.12$$

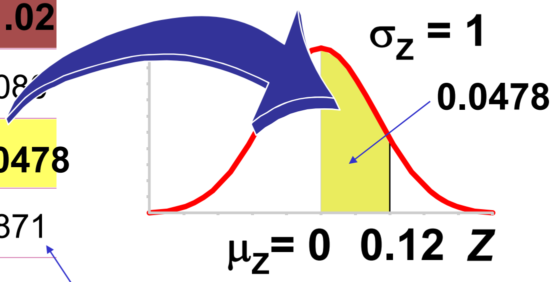
Normal Dağılım | **Standart Normal Dağılım**



Olasılığın Elde Edilmesi

Standart Normal Olasılık Tablosu (Kısmen)

Z	.00	.01	.02
0.0	.0000	.0040	.008
0.1	.0398	.0438	**.0478**
0.2	.0793	.0832	.0871
0.3	.1179	.1217	.1255

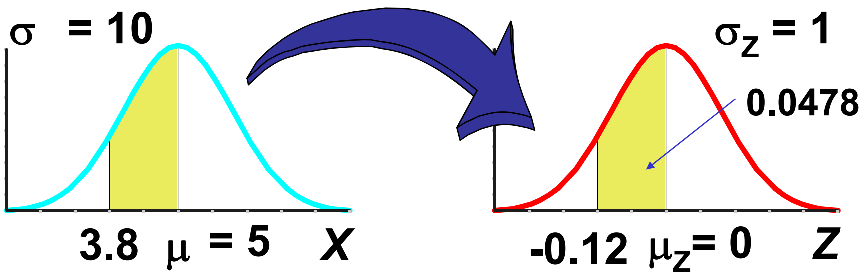


Olasılıklar

Örnek; $P(3.8 \leq X \leq 5) = ?$

$$Z = \frac{X-\mu}{\sigma} = \frac{3.8-5}{10} = -0.12$$

Normal Dağılım **Standart Normal Dağılım**



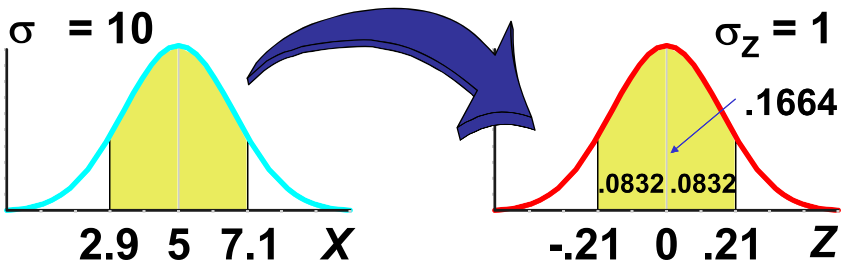
Örnek

$P(2.9 \leq X \leq 7.1) = ?$

$$Z = \frac{X-\mu}{\sigma} = \frac{2.9-5}{10} = -.21$$

$$Z = \frac{X-\mu}{\sigma} = \frac{7.1-5}{10} = .21$$

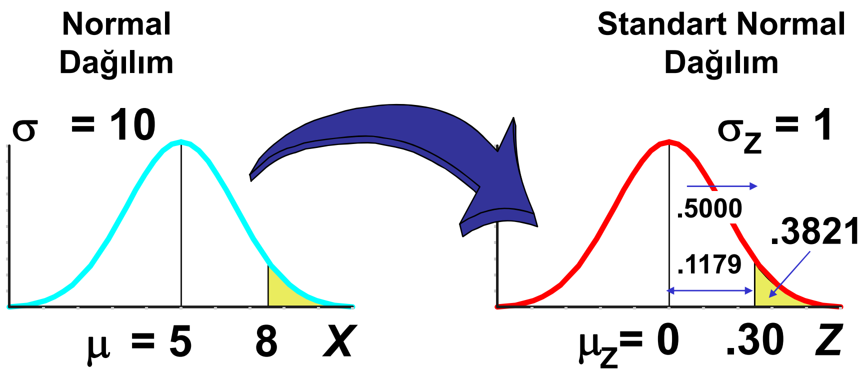
Normal Dağılım **Standart Normal Dağılım**



Örnek

P(X ≥ 8) = ?

$$Z = \frac{X - \mu}{\sigma} = \frac{8 - 5}{10} = .30$$

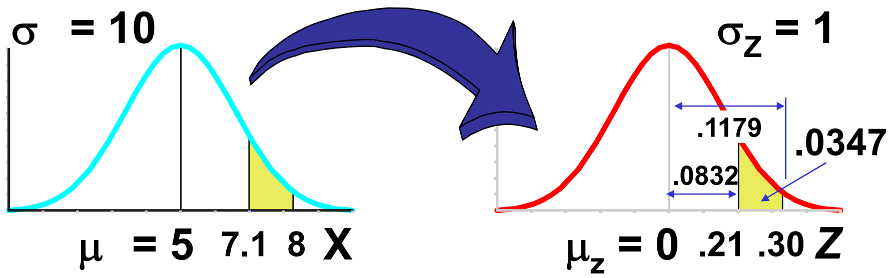


Örnek

P(7.1 ≤ X ≤ 8) = ?

$$Z = \frac{X - \mu}{\sigma} = \frac{7.1 - 5}{10} = .21$$

$$Z = \frac{X - \mu}{\sigma} = \frac{8 - 5}{10} = .30$$



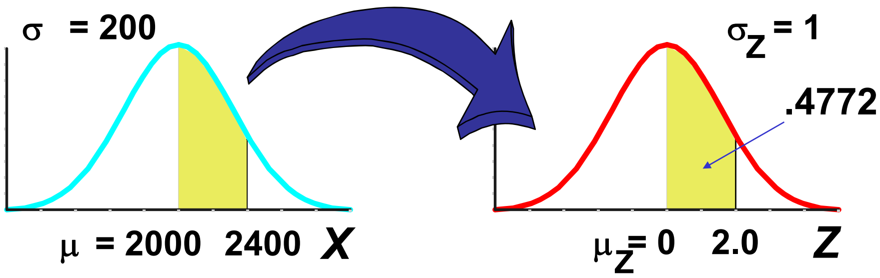
Normal Dağılım Alıştırması

- General Electric için Kalite Kontrol uzmanı olarak çalışıyorsunuz. Bir ampulün ömrü μ **= 2000** saat, σ **= 200** saat olan Normal dağılım göstermektedir. Bir ampulün
- A. **2000** & **2400** saat arası dayanma
- B. **1470** saatten az dayanma olasılığı nedir?

Çözüm
A) P(2000 ≤ X ≤ 2400) = ?

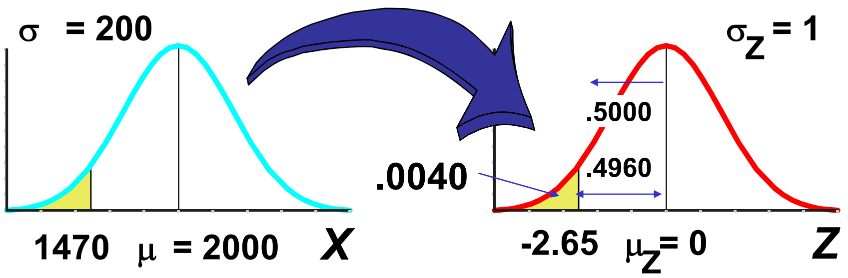
$$Z = \frac{X - \mu}{\sigma} = \frac{2400 - 2000}{200} = 2.0$$

Normal Dağılım | **Standart Normal Dağılım**



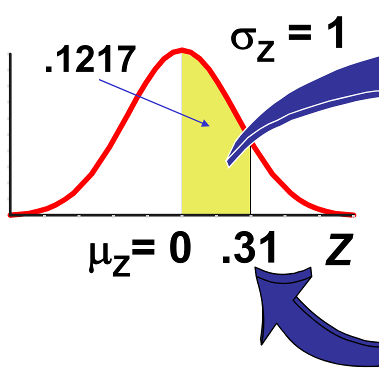
Çözüm
B) P(X ≤ 1470) = ?

$$Z = \frac{X - \mu}{\sigma} = \frac{1470 - 2000}{200} = -2.65$$



Bilinen Olasılıklar İçin Z Değerlerinin Bulunması

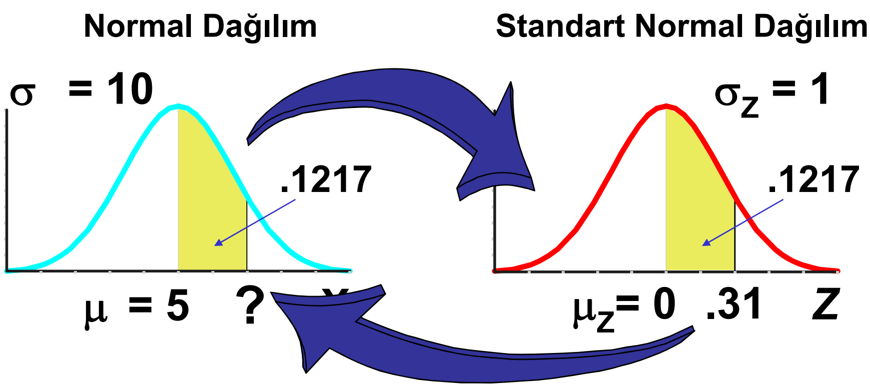
***P*(*Z*) = 0.1217 ise Z nedir?**



Standart Normal olasılık Tablosu (Kısmen)

Z	.00	**.01**	0.2
	000	.0040	.0080
0.	.0398	.0438	.0478
0.2	.0793	.0832	.0871
0.3	.1179	**.1217**	.1255

Bilinen Olasılıklar İçin X Değerlerinin Bulunması

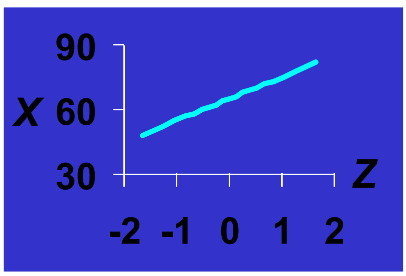


$$X = \mu + Z \cdot \sigma = 5 + (0.31) \cdot 10 = 8.1$$

Normallik Varsayımı

- 1. Verilerin karakteristiklerini Normal dağılımın özellikleriyle karşılaştırın
- 2. Normal Olasılık Plot'unu değerlendirin
  - Bilgisayarla çizin yada
  - Verileri standartlaştırılmış kantil değerlerine karşı işaretleyin.

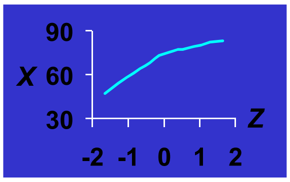
Normal Dağılım İçin NormalOlasılık Plot'u



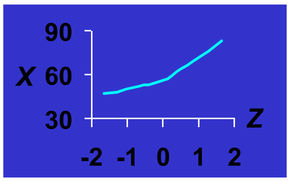
Düz bir çizgi Olmalı!!!

Normal Olasılık Plot'ları

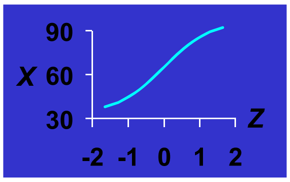
Sola çarpık



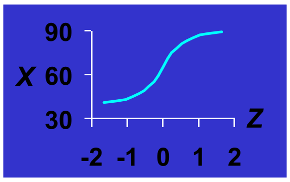
Sağa çarpık



Dikdörtgensel



U-Şekilli



Bir ana kütlenin aritmetik ortalamasının sınırlarını %99,34 güvenle tahmin edebilmek için kullanılacak standart hata katsayısı kaçtır?

a)**2,66**

b)**2,68**

c)**2,70**

d)**2,72**

e)**2,74**

Normal dağılıma sahip bir ana kütle ortalamasının 135 olup olmadığının sınanması için rasgele çekilen bir örneğin ortalaması 141, standart hatanın tahmini değeri 12'dir. Ortalamanın dönüştüğü z değerinin sağında kalan bölgenin oranlanmış alanı kaçtır?

a)0,1915 b) 0,3085 c) 0,5000

d) 0,6915 e) 0,8035

Ana kütle ortalamasının 70 olup olmadığının sınanmasında çekilen 50 birimlik bir örneğin ortalaması 76, standart sapması 21'dir. Örnek ortalamasına karşılık gelen z değeri kaçtır?

a)0,5 b) 1 c) 1,5

d) 2 e) 2,5

X lisesi mezunlarının ÖSS sınavında başarı oranı %50'dir. Bu yıl mezun olacak 324 öğrenciden 153 veya daha fazlasının ÖSS'de başarılı olma olasılığı nedir?

a) 0,1587
b) 0,3413
c) 0,5000
d) 0,6826
e) 0,8413

%99,60 güvenle tahmin yapmak için hangi standart hata katsayısı kullanılır.

a)2,80
b)2,88
c)2,90
d)2,92

Dağılımı normal, belli bir ana kütleden seçilen 101 birimlik bir örneğin ortalaması 20, standart sapması 5 olarak bulunmuştur. Ana kütle ortalaması %95,44 güvenle hangi aralıkta yer alır?

a)19-21

b)15-25

c)17-23

d)18-22

Dağılımı normal, belli bir ana kütleden seçilen 101 birimlik bir örneğin ortalaması 20, standart sapması 5 olarak bulunmuştur. Ana kütle ortalaması %68,26 güvenle hangi aralıkta yer alır?

a)18,5-21,5

b)19-21

c)19,5-20,5

d)19,55-20,45

4. Normal dağılıma sahip bir ana kütleden rasgele seçilen 226 birimlik örneğin ortalaması 16, standart sapması 4'tür. Buna göre ana kütle ortalaması %99,30 güvenle hangi aralıkta yer alır?

a) 14,20-16,72

b) 14,28-17,72

c) 15,26-16,74

d) 15,28-16,72

e) 15,28-17,72

5. Bir milletvekili kendi seçim bölgesinde partisinin oy oranını tahminlemek istiyor. Bu amaçla rassal olarak 750 seçmen seçiyor ve bunların 495'inin kendi partisine oy vereceğini tespit ediyor. Bu bilgilere göre istenen tahminin %95 güven sınırları nedir?

a) $0,627 < \pi < 0,693$

b) $0,600 < \pi < 0,702$

c) $0,598 < \pi < 0,637$

d) $0,640 < \pi < 0,712$

e) $0,602 < \pi < 0,701$

6. Ortalaması 50 kg. ve varyansı 23 kg. olan 50 birimlik bir örneğin, ana kütle ortalaması %95 güvenle hangi aralıkta yer alır?

a) 40,6-59,4

b) 45,08-59,34

c) 45,2-54,8

d) 48,66-51,34

e) 49,74-50,26

1. Bir ana kütlenin aritmetik ortalamasının sınırlarını %99,34 güvenle tahmin edebilmek için kullanılacak standart hata katsayısı kaçtır?

a) 2,66

b) 2,68

c) 2,70

d) 2,72

e) 2,74

2. Normal dağılıma sahip bir ana kütle ortalamasının 135 olup olmadığının sınanması için rasgele çekilen bir örneğin ortalaması 141, standart hatanın tahmini değeri 12'dir. Ortalamanın dönüştüğü z değerinin sağında kalan bölgenin oranlanmış alanı kaçtır?

a) 0,1915 b) 0,3085 c) 0,5000

d) 0,6915 e) 0,8035

3. Ana kütle ortalamasının 70 olup olmadığının sınanmasında çekilen 50 birimlik bir örneğin ortalaması 76, standart sapması 21'dir. Örnek ortalamasına karşılık gelen z değeri kaçtır?

a) 0,5 b) 1 c) 1,5

d) 2 e) 2,5

4. X lisesi mezunlarının ÖSS sınavında başarı oranı %50'dir. Bu yıl mezun olacak 324 öğrenciden 153 veya daha fazlasının ÖSS'de başarılı olma olasılığı nedir?

a) 0,1587

b) 0,3413

c) 0,5000

d) 0,6826

e) 0,8413

6. 400'ü kız ve 600'ü erkek olmak üzere toplam 1000 öğrencinin okuduğu bir liseden, 150 öğrencilik bir örneklem rassal olarak seçilecektir. Bu örneklem içinde 50'den az kız öğrenci bulunması olasılığını hesaplayınız.