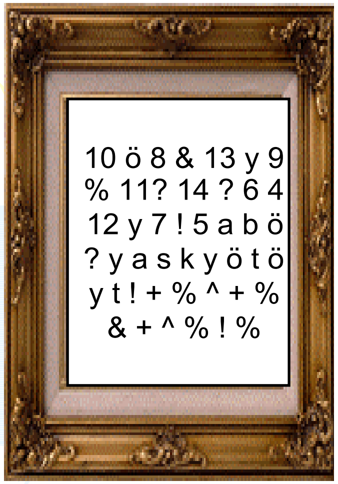
DENEYLERİN TASARIMI
DENEYSEL TASARIM

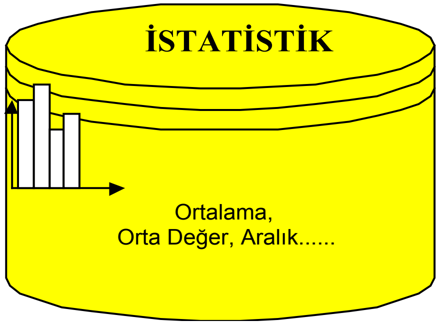
İSTATİSTİKSEL DENEY PLANLAMANIN TEMEL İLKELERİ

Design of Experiment

EXPERIMENTAL DESIGN

istatistiki temel ilkeler





İstasistik neden önemli?

Gerçekten istatististiksel sonuçlara ihtiyacımız var mı?

temel ve yorumlayıcı istatistik

- tanımlayıcı istatistik (descriptive statistics)
  - verilerin tanımlanmasında
  - verilerin ne işe yaradığını
  - verilerin dış görünümünü ortaya çıkarır
- yorumlayıcı istatistik (inferential statistics)
  - verilerin ne anlama geldiğini
  - verilerden ne gibi sonuçlar çıkarılacağını
  - verilerin iç dünyasını inceler

bir grup aritmetik ve grafik araçlar

populasyon ve örnek grup

- populasyon

  ve

- örnek grup

Örnek grupların populasyonu yansıtacak şekilde seçilmesi olduça önemlidir. Yanlı veya yanlış yapılacak seçimler dönüşü bazen olmayan pahalı sonuçlara ve yorumlara yol açabilirler.

İstatistiğin en çok kullanıldığı alanlardan birinin tıp olduğu gözönüne alınırsa tıp dünyasında yapılan yanlış istatistiklerin katostropik sonuçlar verebileceği unutulmamalıdır.

temel istatistik

- Aritmetik araçlar
  - ortalama (average)
  - orta değer (medyan)
  - mod (mode)
  - aralık (range)
  - değişim (variance)
  - standart sapma (standard deviation)
- Grafikler
  - dağınıklık/ilişki (scatter diagram)
  - ağaç-yaprak (stem-leaf)
  - çubuk (bar)
  - pay (pie)
  - olasılık (probability)
  - histogram (histogram)

Sayısal veriler yalan söyleyebilirler ama grafikler istatistiğin gerçekleri tüm çıplaklığı ile ortaya koyar

temel istatistik

- Ortalama $\overline{X}$, μ (Average, Mean)

verilerin aritmetik ortalamasıdır. Aynı kategorideki tüm verilerin toplamının verilerin sayısına bölünmesi ile elde edilir. Çerçeve içinde aynı kategoride olan sayıların ortalaması;

X : değişken
N : veri grubundaki değişken adedi
σ : populasyon ortalaması
$\overline{X}$: örnek ortalaması

$$\mu = \overline{X} = \frac{\sum_{i}^{N} X_i}{N}$$

temel istatistik

- Ortalama

32,32,35,36
38,37,39,40
39,39,40
42,45

$$\mu, \overline{X} = \frac{(32+32+35+36+37+38+38+39+39+39+40+40+42+45)}{14} = \frac{433}{14} = 30.9$$

temel istatistik

- Orta Değer (Median)

Büyükten küçüğe veya küçükten büyüğe sıralanmış aynı kategorideki verilerin ortasında yeralan veridir.

$$\textbf{Orta Değer} = \frac{X_1 + X_2}{2}$$

temel istatistik

- Orta Değer (Median)

32 32 35 36 36 37 38 38 39 39 39 40 40 42 45

- Veriler sıralandığında verilerin yarısı orta değerden büyük yarısıda küçüktür.
- Eğer verilerin toplam adedi çift sayı ise orta değer ortadaki iki değerin toplamının yarısına eşittir.

32 32 35 36 36 37

38 39

39 39 40 40 42 45

Orta Değer $= \dfrac{38+39}{2} = 38.5$

temel istatistik

- **Orta Değer (Median)**

32 32 35 36 36 37 38 38 39 39 39 40 40 42 45

– Veri grubunda bir ekstra değer daha olsaydı örneğin bu değer 46 olarak alınırsa bu durumda veri adedinin tekil sayı olduğu veri gruplarında verilerin ortasında sadece bir veri olacaktır. Bu veri orta değer olarak alınır.

32 32 35 36 36 37 38

39

39 39 40 40 42 45 46

temel istatistik

•Mod (Mode)

Verilerin içinde en çok gözleneni veya en popular olanıdır.

Bazı durumlarda birden fazla mod olma olasılığı olabilir. Bu durumda veri grubu çok modlu olarak anılır. Mod sayısına göre ikili, üçlü vb. Eğer hiçbir değer birden fazla tekrarlanmamışsa veri grununun belirgin bir modu yoktur yargısına varılabilir.

Mod: En popular veri

Örnekteki veri grubunda 39 sayısı diğer verilerden en fazla tekrarlananı, yani en popular olanıdır.

32 32 35 36 36 37 38 38 <u>39 39 39</u> 40 40 42 45

Mod=39

temel istatistik

- **Aralık (Range)**

Veri grubundaki en büyük veri ile en küçük veri arasındaki farktır.

Aralık= Enbüyük değer- Enküçük değer

Veri grubunda en büyük veri 45 ve en küçük veri 32 dir.

Aralık=45-32=13

temel istatistik

- Değişim (σ^2 , s^2) (Variance)
  - Değişim tüm verileri içine alır
  - Verilerin nasıl dağıldığını gösterir
  - Verilerin hepsi aynı olsalardı verilerin değişimi ve aralık değeri sıfır olurdu
  - Sayıların ortalaması orta değer ve mod birbirlerine eşit olur ve değerlerin gizemli bir tarafı kalmazdı
  - Sayılar ne kadar çok dağınık olurlarsa değişim de o kadar büyük olur

temel istatistik

- **Değişim (σ^2 , s^2) (Variance)**

  – **İki farklı değişim hesaplaması vardır.**
    - **Popuplasyon Değişimi**
    - **Örnek Değişimi**

  – **Populasyondaki değişimi bulmak için herbir verinin ortalama değerle arasındaki farkların karelerinin toplamının değerlerin <u>toplam adedine</u> bölümü ile elde edilir.**

- **Populasyonun değişimi;**

$$\sigma^2 = \frac{\sum_{i=1}^{N} S_i^2}{N}$$

temel istatistik

- **Populasyon Değişimi, σ^2**

  – **Populasyondaki değişimi bulmak için herbir verinin ortalama değerle arasındaki farkların karelerinin toplamının değerlerin toplam adedine bölümü ile elde edilir.**

- **Populasyonun değişimi;**

$$\sigma^2 = \frac{\sum_{i=1}^{N} S_i^2}{N}$$

temel istatistik

- Populasyon Değişimi, σ^2
  - Öncelikle herbir veri ile ortalama değer arasındaki farkı bulalım. Çıkan sonucun negatif olması sizleri ürkütmesin.

$$Fark = (X_i - \overline{X})$$

  - Şimdi farkların karelerini bulalım. Böylelikle tüm değerler pozitif hale gelecektir.

$$S^2 = Fark^2 = (X - \overline{X})^2$$

temel istatistik

- **Populasyon Değişimi, σ^2**
  - **32 değeri için değişim**

$$S^2 = (32 - 30.93)^2 = 1.08^2 = 1.17$$

  - **Karelerin toplamı**

$$SS = \sum S^2 = 1.17 + 1.17 + 16.65 + 25.81 + 36.97 + 50.13 + 50.23 + 65.29 + .. = 863.77$$

  - **Değişim**

$$\sigma^2 = \frac{\sum S^2}{N} = \frac{SS}{N} = \frac{863.77}{14} = 61.70$$

temel istatistik

- **Örnek Değişimi, s^2**
  - örnek alınan bir alt grup verinin değişimini
  - örnek grup değişiminde aynı ortalama değerden farklarının karelerinin toplamının verilerin toplam adedinin bir eksiğine bölünmesi ile elde edilir.
  - Veri setinde yer alan toplam veri adedi yerine adedin bir eksiğine bölmekle örnek değişimin "bias" yani yönlenmesi giderilmiştir.

$$s^2 = \frac{\sum_{i=1}^{N} S_i^2}{N-1}$$

$$SS = \sum S^2 = 1.17 + 1.17 + 16.65 + 25.81 + 36.97 + 50.13 + 50.23 + \ldots = 863.77$$

$$s^2 = \frac{863.77}{14-1} = 66.44$$

temel istatistik

- Standart Sapma (σ, s)
  - Verilerin ortalama değere ne kadar yakın olduğunu gösteren değerdir ve değişimin karekökü alınarak hesaplanabilir.

- **populasyonun standart sapma**
- **örnek grubun standart sapması**

temel istatistik

- Populasyonun standart sapması, σ

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N} S_i^2}{N}}$$

$$\sigma = \sqrt{s^2}$$

$$\sigma = \sqrt{\frac{863.77}{14}} = \sqrt{61.70} = 7.85$$

temel istatistik

- Örnek grup standart sapması, **s**

$$s = \sqrt{\frac{\sum_{i=1}^{N} S_i^2}{N-1}}$$

$$s = \sqrt{\frac{863.77}{13}} = \sqrt{66.44} = 8.15$$

Veri gösterim araçları

- Dağınıklık /ilişki diyagramları

- Dağılım ve dağılım diyagramları
  - Frekans diyagramları
  - Histogramlar
  - Çubuk diyagramları
  - Frekans Poligonları
    - Mutlak Frekans Poliganları
    - Rölatif frekans Poliganları
    - Kümülatif Mutlak Frekans Poliganları
    - Kümülatif Rölatif Frekans Poligonları
  - Gövde yaprak diyagramları

veri gösterim araçları

- Dağınıklık ilişki diyagramları

Dağınıklık diyagramları genelde değişkenler arasındaki ilişkileri incelerler ve bu ilişkilerin büyüklüğünü dile getirirler.

Bu ilişkilerde sebep-sonuç bağlantısı olabileceği gibi bazı durumlarda iki değişken birbirleri ile sebep-sonuç ilişkisi olmadan da ilişkili olabilir.

Doğrusal olan ilişkilerin doğrultusu negatif veya pozitif olabilir. Bu negatif veya pozitif ilişkiler zayıf veya güçlü olabilirler veya iki değişken arasında hiçbir ilişki olmayabilir.

veri gösterim araçları

Dağınıklık diyagramlarının oluşumunda adımlar :

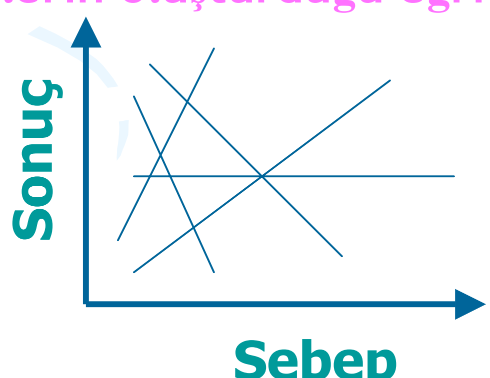
- **Öncelikle iki değişkene ihtiyaç vardır. Bu değişkenler biri bağımsız (x, sebep) diğeri ise bağımlı değişkendir (y, sonuç).**

- **Bu değişkenlerden bağımsız (sebep) değişken olanı genellikle yatay eksende ve bağımlı (sonuç) değişken dikey eksende olacak şekilde diyagram çizilir. Her bağımsız değişkene karşılık diyagramda bir bağımlı değişken değeri yer alır.**

veri gösterim araçları

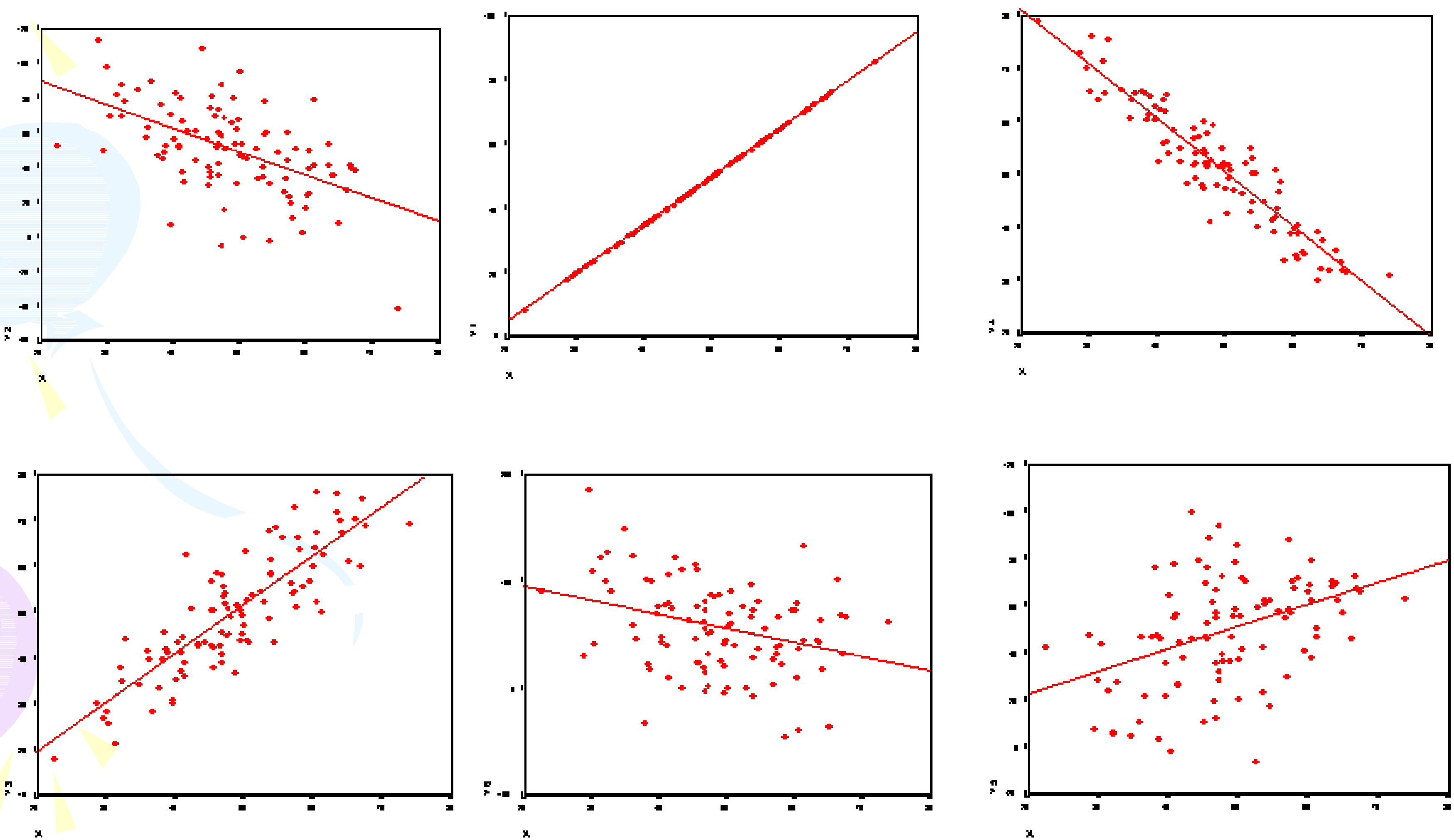
Dağınıklık diyagramları nasıl yorumlanır.

- pozitif ilişki : artan bağımsız değişkene karşılık bağımlı değişkende artıyorsa
- negatif ilişki : sebep değişken arttıkça sonuç değişken azalıyorsa
- ilişki yok: sebep-sonuç değişken arasında belirgin bir ilişki tanımlanamıyorsa başka bir deyişle veriler dağınık veya birbiri içine çok girmiş ise
- veriler doğru oluşturacak şekilde biraraya geldikçe aralarındaki ilişki o kadar güçlü hale gelir
  - eğrinin eğimide ilişkinin güçlü olup olmadığı hakkında bilgi verir
  - Verilerin oluşturduğu eğri dikleştikçe ilişkide güçlenir.



veri gösterim araçları

Değişik dağınıklık diyagramları



veri gösterim araçları

İstatistik ve dağınıklık diyagramları

•Dağınıklık diagramları verilerin istatistik analizleri ile ilgili bilgi vermezler ama verilerin doğasını çıplak olarak yansıtırlar.

•İstatistik analizler bazen yanıltıcıdırlar, istatistik ve dağınıklık diyagramları arasındaki işbirliği:

veri gösterim araçları

İstatistik ve dağınıklık diyagramları

X1	Y1	X2	Y2	X3	Y3
10	8.04	10	9.14	10	7.46
8	6.95	8	8.14	8	6.77
13	7.58	13	8.74	13	12.74
9	8.81	9	8.77	9	7.11
11	8.33	11	9.26	11	7.81
14	9.96	14	8.10	14	8.84
6	7.24	6	6.13	6	6.08
4	4.26	4	3.10	4	5.39
12	10.84	12	9.13	12	8.15
7	4.82	7	7.26	7	6.42
5	5.68	5	4.74	5	5.73

- Bu veri gruplarının herbirinin istatistiki analizi aynıdır.

veri gösterim araçları

İstatistik ve dağınıklık diyagramları

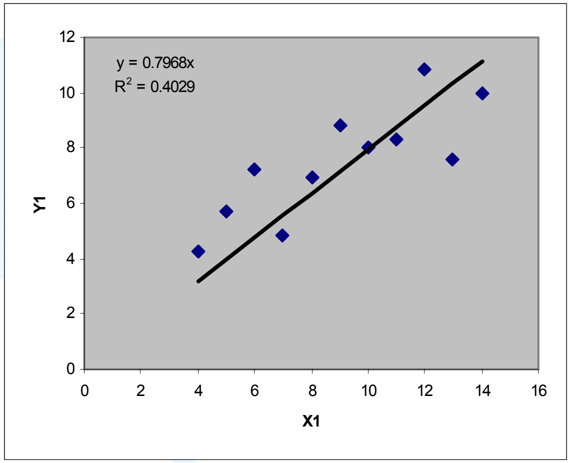
- Bu veri gruplarının herbirinin istatistiki analizi aynıdır.

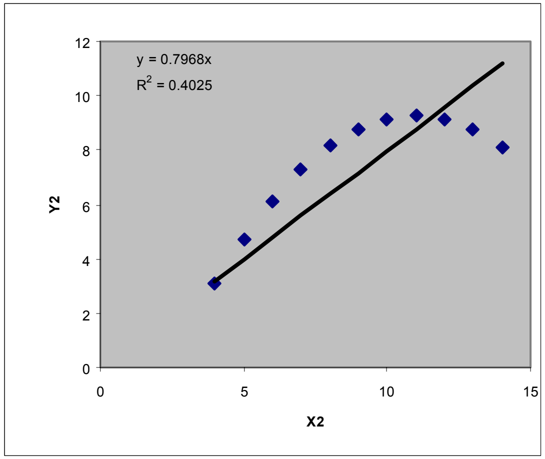
– N	:11
– X değerlerinin ortalaması	:9
– Y değerlerinin ortalaması	: 7,5
– Eğim	: 0,5
– Kesişim	: 3
– Standart sapma	:1,126
– Korelasyon	:% 81,7

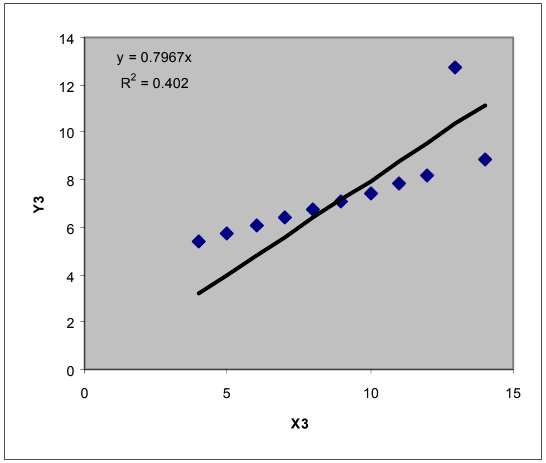
veri gösterim araçları

İstatistik ve dağınıklık diyagramları

- Bu veri gruplarının herbirinin istatistiki analizi aynıdır.







veri gösterim araçları

Dağınıklık diyagramları

- Değişkenler arasındaki ilişkilerin herzaman doğrusal olmayabilir

- Dışardakiler (outliers) olarak bilinen ve diyagramlarda rahatlıkla ayıredilebilen veri çiftleri vardır.
  - dışardakiler diyagramlarda (xn,yn) gibi veri çiftlerinden oluşurlar ve bu veri grupları diyagramda değişkenler arasındaki ilişkiyi yansıttığına inanılan patternin dışında yeralırlar.

veri gösterim araçları

Dağılım diyagramları

Dağılım diyagramları grafiksel anlamda bir grup olarak elde edilmiş verilerin grup içinde varolma ve/veya gözlenme sıklıklarını ele alarak dağılımlarını görüntüleyen diyagramlar ailesidir.

Temelde aynı amaca hizmet etmekle beraber belli başlı dağılım diyagramları:

- Frekans diyagramlar
- Histogramlar
- Çubuk diyagramları
- Frekans Poligonları
- Gövde yaprak diyagramları

veri gösterim araçları

Frekans diyagramları

Frekans diyagramlarını oluşturmadan önce verilerin organize edilmesine yönelik olarak frekans tablolarının hazırlanması gerekmektedir.

- **Frekans tabloları genellikle iki sütundan oluşur.**
  - **Sütunlardan biri olası verileri veya veri aralıklarını**
  - **Diğeri ise verilerin oluşum frekansını veya o verinin söz konusu veri setinde kaç kez oluştuğunu belirler.**
  - **Frekans dağılımlarının doğru olarak çizilmesi için varolan veri seti içinde bulunmayan verilerin frekansları sıfır olarak alınmalıdır.**

veri gösterim araçları

Histogram

- **Histogramlar oluşturulan frekans tablolarında yer alan verilerin aralıklarının orta noktalarının x eksenine ve frekanslarının y eksenine verilmesi ile elde edilir.**

- **Aralıkta yeralan her değer için bir sütun çizilir bu sütunun genişliği veri aralığı kadar olup x eksenindeki orta noktası gösterir. Sütunun yüksekliği ise sözkonusu verinin frekansına eşit olmalıdır.**

veri gösterim araçları

Histogramların oluşturulmasında,

- Değişken içinde yeralan verilerin belirgin aralıklarda gruplandır ve gözlem frekanslarının belirle.
- Veri setinin aralık genişliğini tespit et.
- Veri aralık değerini istenilen aralık sayısına böl (Önerilen aralık adedi 10)

  elde edilen değer tam sayı değilse bulunan sayı en yakın tek sayıya yuvarla.
- İlk aralığın sınırı için ilk aralık değerine bulunan aralık değerinin ekle ve bu değerden bir çıkar
- Aynı işlemlere devam et ve diğer aralıkların sınırlarıda belirle.
- Aralıkların orta noktaları için herbir aralığın başlangıç ve son sınır değerlerinin ortalamasını bul.
- Son adımda aralıkların orta noktaları x-ekseninde, y-ekseninde ise belirlenen aralıklara düşen verilerin frekansları olacak şekilde histogram oluşturulur.
- Histogramlarda orta noktalar arasındaki sınır değerlerine gerçek sınır değerler olarak tanımlanır. Gerçek sınır değerler arasındaki farklarda aynı zamanda bulunan aralık genişliği değerine eşittir.

veri gösterim araçları

Histogram

Aşağıda bir sınıfın matematik sınav sonuçları yeralmaktadır.

53,35,67,48,63,42,48,55,33,50,46,45,59,40,47,51,66,53

Veri	Frekans
33	1
35	1
40	1
42	1
45	1
46	1
47	1
48	2
50	1
51	1
53	2
55	1
59	1
63	1
66	1
67	1
Toplam Frekans	18

veri gösterim araçları

Histogram

_Verilerin aralığı=67-33=34
_Aralık adedi 10 olarak alındığında
_Aralık boyutu=34/10=3.4
_Bu değere en yakın tek sayı 3 dür.
_Dolayısıyla diyagramdaki aralık boyutu 3 dür.

veri gösterim araçları

Histogram

Bu değere göre söz konusu 10 aralığın üst ve alt limitleri aşağıdaki şekilde olacaktır.

	Açık Sınır		Gerçek Sınır		
Aralık	Alt Limit	Üst Limit	Alt Limit	Üst Limit	Limit Orta Noktası
33-35	33	35	32.5	35.5	34
36-38	36	38	35.5	38.5	37
39-41	39	41	38.5	41.5	40
42-44	42	44	41.5	44.5	43
45-47	45	47	44.5	47.5	46
48-50	48	50	47.5	50.5	49
51-53	51	53	50.5	53.5	52
54-56	54	56	53.5	56.5	55
57-59	57	59	56.5	59.5	58
60-62	60	62	59.5	62.5	61
63-65	63	65	62.5	65.5	64
66-68	66	68	65.5	68.5	67

veri gösterim araçları

Histogram

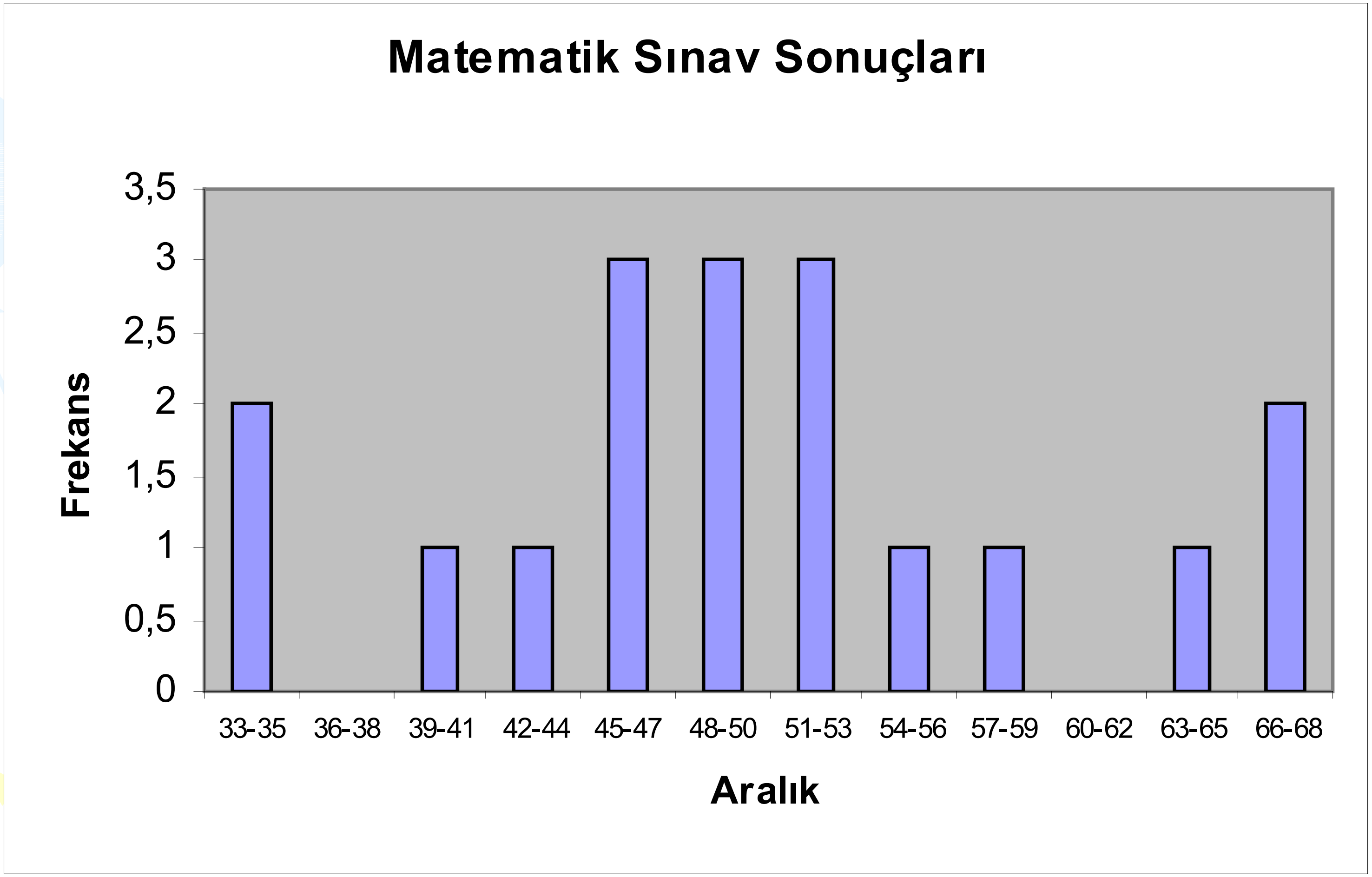
10 aralık olduğunda aralık-frekans tablosu aşağıdaki gibidir

Aralık	Frekans
33-35	2
36-38	0
39-41	1
42-44	1
45-47	3
48-50	3
51-53	3
54-56	1
57-59	1
60-62	0
63-65	1
66-68	2

veri gösterim araçları

Histogram

Aralık boyutu 3 olduğunda



veri gösterim araçları

Histogram

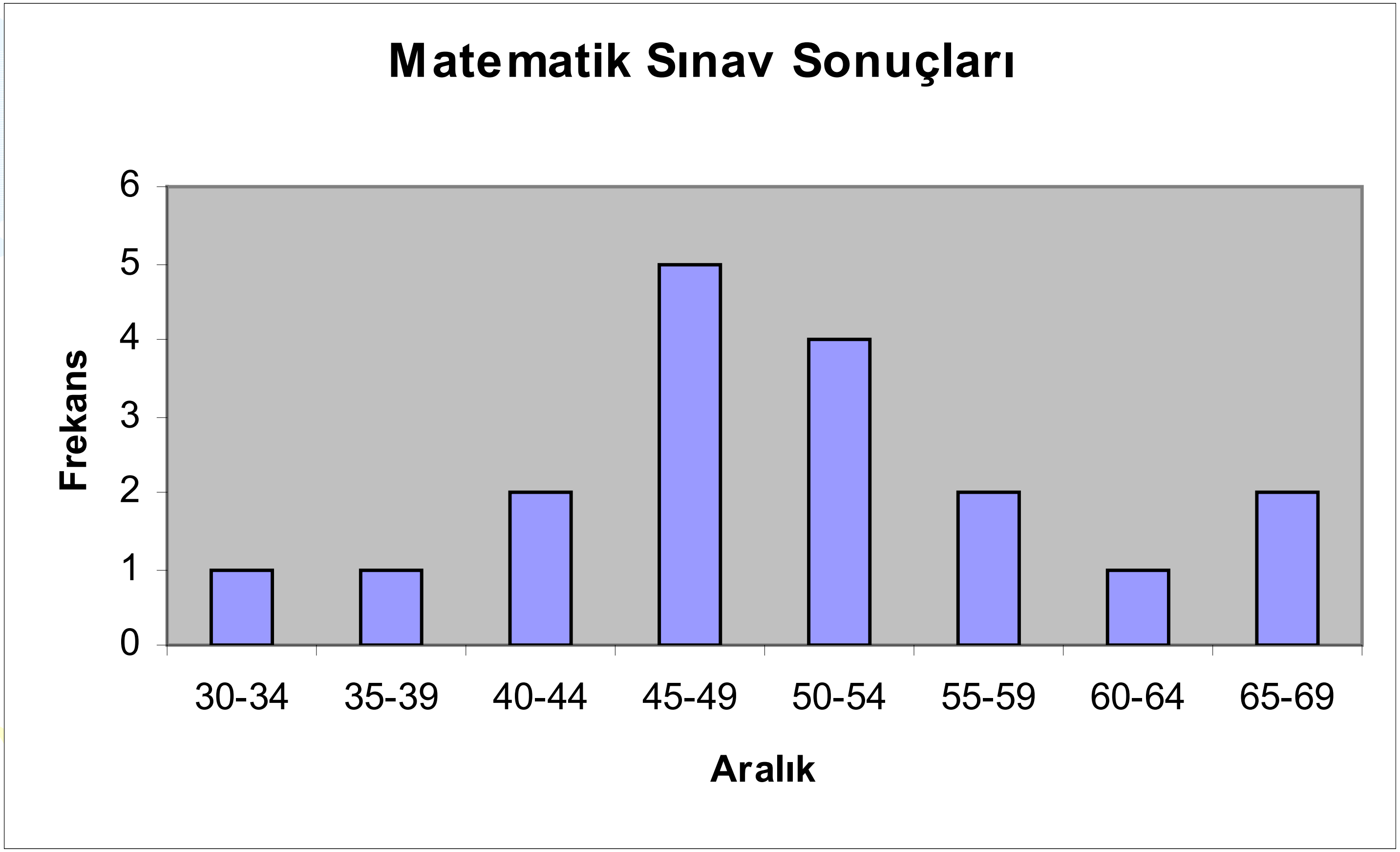
- Aralık boyutu 5 olduğunda

	Açık		Gerçek			
Aralık	Alt Limit	Üst Limit	Alt Limit	Üst Limit	Limit Orta Noktası	Mutlak Frekans
30-34	30	34	29.5	34.5	32	1
35-39	35	39	34.5	39.5	37	1
40-44	40	44	39.5	44.5	42	2
45-49	45	49	44.5	49.5	47	5
50-54	50	54	49.5	54.5	52	4
55-59	55	59	54.5	59.5	57	2
60-64	60	64	59.5	64.5	62	1
65-69	65	69	64.5	69.5	67	2
						18

veri gösterim araçları

Histogram

- **Aralık boyutu 5 olduğunda**



veri gösterim araçları

Histogram

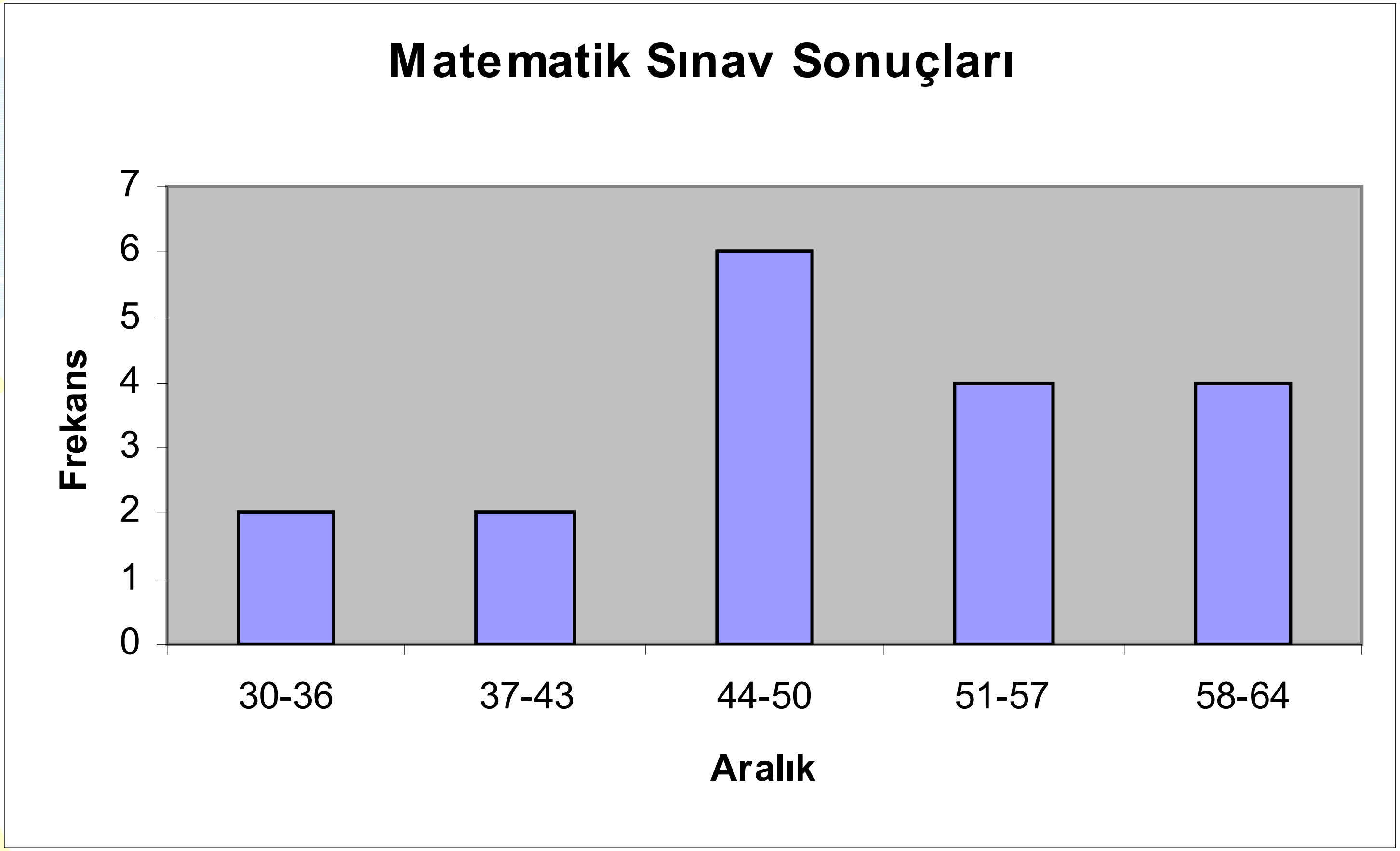
- **Aralık boyutu 7 olduğunda**

Aralık	Frekans
30-36	2
37-43	2
44-50	6
51-57	4
58-64	4

veri gösterim araçları

Histogram

- **Aralık boyutu 7 olduğunda**



veri gösterim araçları

Histogram

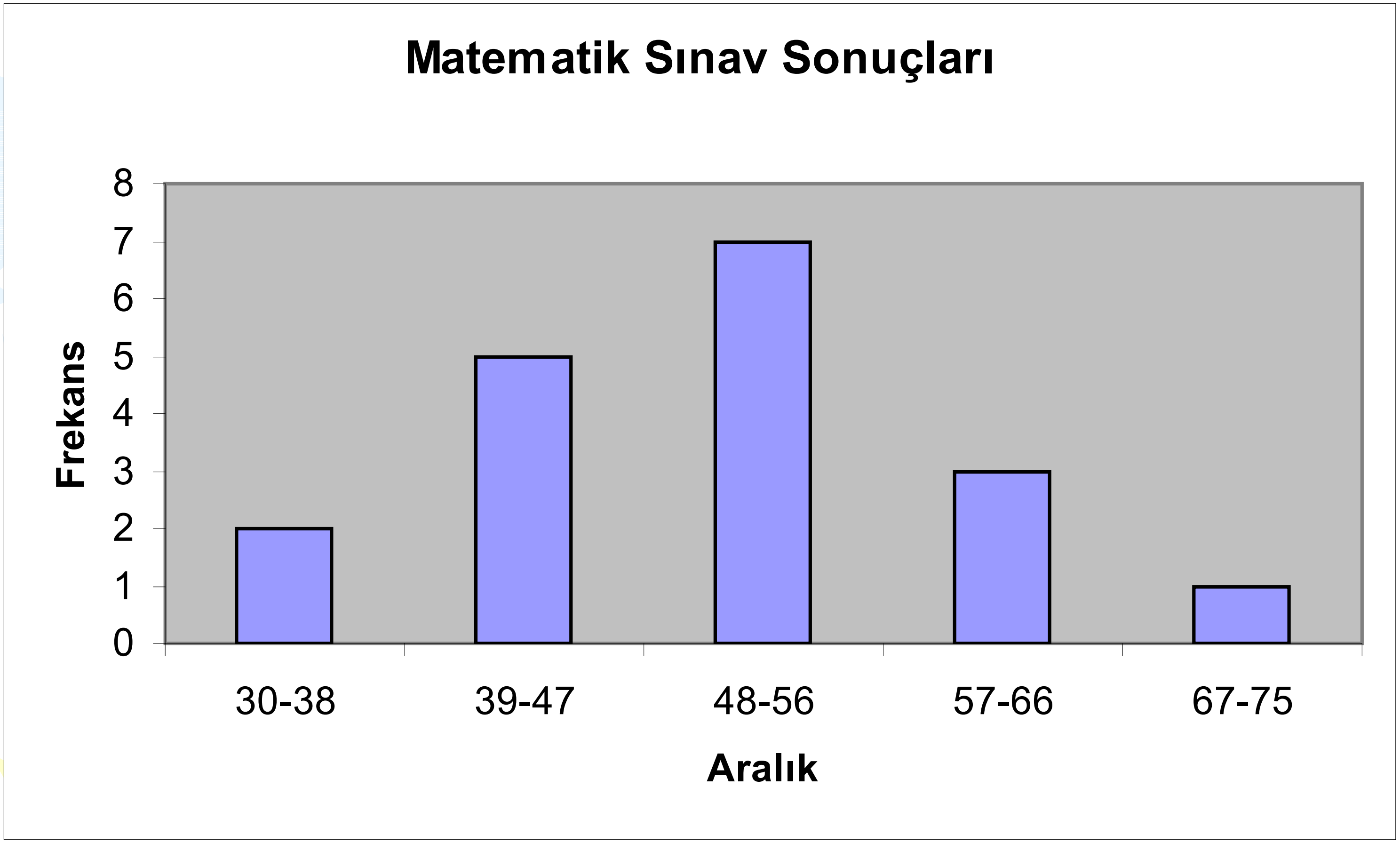
- **Aralık boyutu 9 olduğunda**

Aralık	Frekans
30-38	2
39-47	5
48-56	7
57-66	3
67-75	1

veri gösterim araçları

Histogram

- **Aralık boyutu 9 olduğunda**



veri gösterim araçları

Histogram

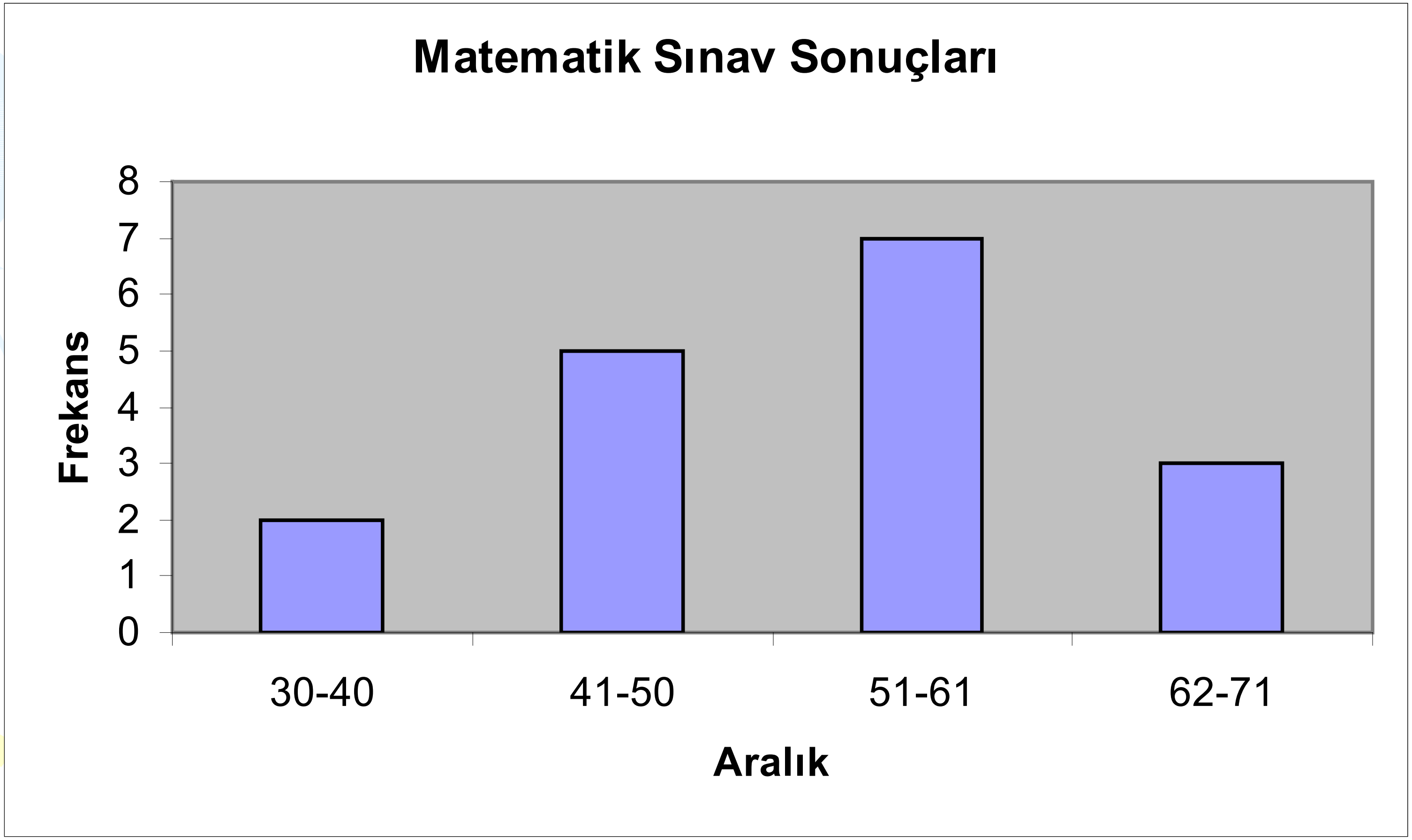
- **Aralık boyutu 11 olduğunda**

Aralık	Frekans
30-40	2
41-50	5
51-61	7
62-71	3

veri gösterim araçları

Histogram

- **Aralık boyutu 11 olduğunda**



veri gösterim araçları

Çubuk Diyagramları

- **Çubuk diyagramı da histogramla aynı kategorididir.**

- **Çubuk diyagramlarında çoğu zaman farklı değişkenler aynı anda yer alabileceği gibi diyagramda yeralan çubuklar histogramın aksine bitişik değil aralıklıdır.**

veri gösterim araçları

Frekans Poligonları

- mutlak,
- rölatif,
- kümulatif mutlak
- ve kümulatif rölatif frekans poligonları

veri gösterim araçları

Mutlak Frekans Poligonları

– Mutlak frekans poligonları histogramların tamamen aynısı olup tek fark sütunlar yerine poligon diyagramlarında sütunların orta noktaları ve frekanslardan oluşan veri çiftlerinin noktaları alınır. Bu noktalar belirlendikten sonra noktalar çizgilerle birleştirilerek diyagram oluşturulur.

Rölatif Frekans Poligonları

– Rölatif diyagramlarını çizebilmek için rölatif frekansların belirlenmesi gerekmektedir. Bir verinin rölatif fekansı o verinin frekansının toplam frekansa bölünmesi ile elde edilir. Rölatif frekans poligonlarıda ve mutlak frekans poligonlarında aynı anlayışta çizilir. Tek fark y-ekseninde frekans değerleri yerine rölatif frekans değerleri yeralır.

veri gösterim araçları

Kümulatif Mutlak Frekans Poligonları

– Kümulatif mutlak frekans belirli bir skorun altında olan skorların toplamıdır. Kümulatif frekans belirlenmiş değerlerin altında olan frekans değerlerinin toplanması ile elde edilir. Bu tür diyagramlarda y ekseninde kümulatif frekans değerleri yer alır. Mutlak kümulatif frekans diyagramları herzaman artış gösteren diyagramlardır bu diyagramlar hiç bir zaman azalan özellik sergilemezler ama belirli bir seviyede sabit olarak kalabilirler.

Kümulatif Rölatif Frekans Poligonları

– Kümulatif rölatif poligonlar da rölatif poligonlarla aynı mantıkla hazırlanırlar. Bu diyagramların hazırlanmasında ilk adım kümulatif rölatif frekanslarının belirlenmesidir. Kümulatif rölatif fekanslar, mutlak kümulatif frekansların toplam kümulatif frekansa bölünmesi ile elde edilir. Bu diyagramlarda kümulatif mutlak diyagramlar gibi herzaman artan bir eğilim gösterirler. Kümulatif rölatif diyagramlarda erişilebilecek en yüksek frekans değeri 1 (bir) dir.

veri gösterim araçları

Gövde yaprak diyagramları

- Verilerin gövde ve yaprak olarak ayır
- Gövde ve yapraklar küçükten büyüğe sırala
- Bu sıralama sonucunda elde edilen tablo verilerin tanımlayıcı istatistik araçlarını belirlemekte oldukça yararlıdır.

0.6, 2.6, 0.1, 1.1, 0.4, 2.0, 0.8, 1.3, 1.2, 1.5, 3.2, 1.7, 1.9, 1.9, 2.0

Gövde-yaprak diyagramı sayılar sıralanmadan hazırlandığında;

Gövde	Yaprak						
0	6	1	4	8			
1	1	3	2	5	7	7	9
2	6	0	2				
3	2						

veri gösterim araçları

Gövde yaprak diyagramları

Gövde-yaprak diyagramı sayılar sıralandıktan sonra hazırlandığında;

Gövde	Yaprak						
0	1	4	6	8			
1	1	2	3	5	7	9	9
2	0	2	6				
3	2						

Düzenlenmiş gövde-yaprak diyagramından orta değer ve modu bulmak oldukça kolaydır.

Gövde-yaprak diyagramında 15 adet veri olduğuna göre sekizinci değer orta değeri verirki 1.5’e eşittir.

Mod ise en popular veri olarak tanımlandığına göre 1.9 en çok tekrarlanan veri olarak bu veri grubunun modu olarak alınır.

olasılık ve şans

Olasılık modelleri

Olasılık modelleri gerçekte frekans dağılımları veya poligonlarına alternatif modellerdir.

Frekans diyagramları x ekseni üzerindeki herbir veriye karşılık o verinin örnek içerisindeki frekansını y ekseninde dağılımını gösterir.

Olasılık diyagramlarında istatistiği biraz daha yoğun olarak kullanarak örnek içindeki dağılımı y eksenine olasılık ve x eksenine ise örneğin temel istatistik hesaplamalarını yerleştirerek yansıtır.

olasılık ve şans

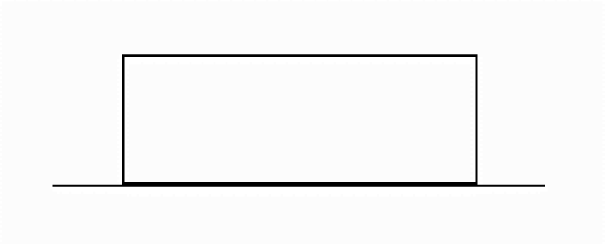
Dağılım modelleri

Verilerin büyük bir titizlikle ve doğrulukla ölçülmesi durumunda dünyanın gerçek yapısı ortaya çıkarılır. Dünya gerçeklerinin herzaman teorik dağılım modellerine uymaları zordur fakat ampirik dağılım diyagramları yararlı kararların alınmasında dünya gerçeklerini yeteri düzeyde ortaya koyarlar. Gerçek dünyanın tanımlanmasında değişik dağılım diyagramları kulanılabilir.

- **Dikdörtgen Dağılım veya Tek biçimli (Uniform) dağılım**
- **Negatif Eksponensiyel Dağılım**
- **Üçgen Dağılım**
- **Normal Dağılım veya Normal Eğri (Simetrik dağılım Diyagramı)**

olasılık ve şans

Dağılım modelleri

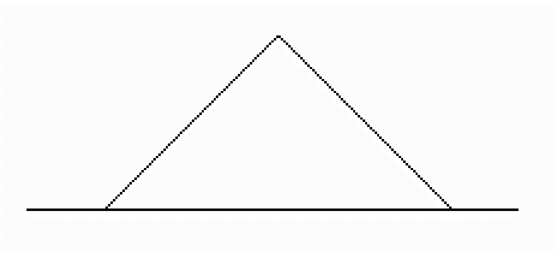


- Dikdörtgen Dağılım veya Tek biçimli (Uniform) dağılım

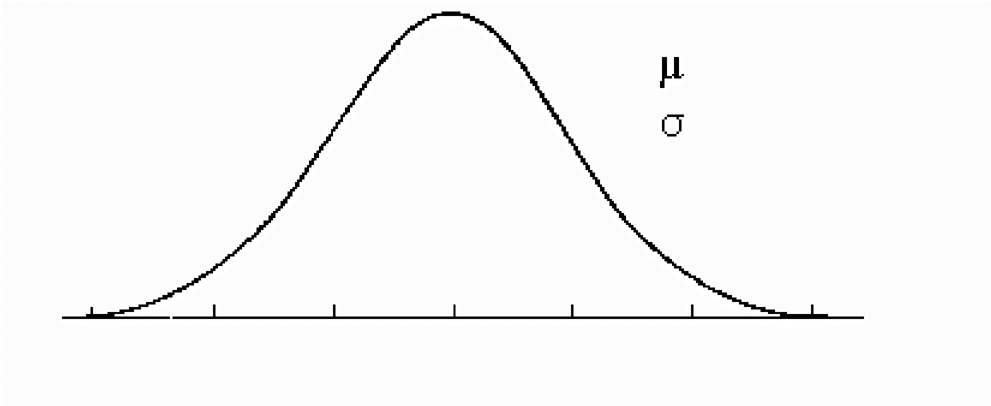
- Negatif Eksponensiyel Dağılım

olasılık ve şans

Dağılım modelleri



- Üçgen Dağılım



- Normal Dağılım veya Normal Eğri (Simetrik dağılım Diyagramı)

olasılık ve şans

Normal Dağılım

- 1773 yılında DeMoivre tarafından geliştirilmiş
- 1924 de kadar bu yaklaşımın kayıtları 1924 bulunamamıştır
- 1783 de Laplace dağılımın hatalarını tanımlamak için normal dağılım yaklaşımını kullanmıştır
- 1809 da Gaus astronomik verileri yorumlamak için aynı yaklaşımı kullanmıştır

Bu yüzden bu dağılım çoğu zaman Gaussian dağılımı veya şeklinden dolayı çan eğrisi diyagramı adı verilir.

olasılık ve şans

Dağılım eğrileri

Simetrik dağılım eğrilerinde:

μ populasyonun ortalaması

σ ise standart sapma değeridir

Bu tür dağılım eğrilerinde

populasyon ortalaması, mod ve orta değer birbirlerine eşittir.

Şekil de farklı ortalama ve standart sapma değerlerine sahip değişik normal dağılım eğrileri verilmiştir. Şekilllerden de görüldüğü üzere populasyon ortalamasındaki, μ, değişim normal dağılım eğrilerinin yüksekliklerini etkilerken, standart sapma, σ, ise normal dağılım eğrisinin yayılımını etkiler.

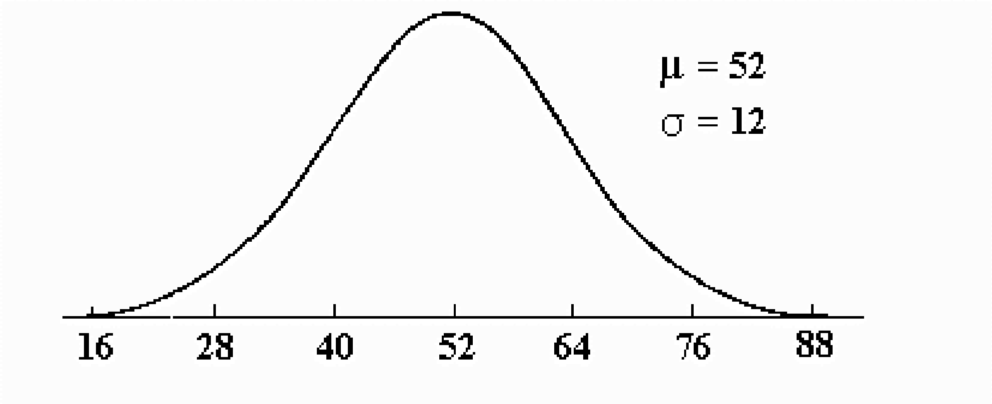
Normal dağılım eğirilerinin karakteristik özelliklerinden bir diğeri de normal dağılımdaki gözlemlerin veya verilerin yaklaşık

- %68 i $\pm$ 1σ
- %95 i $\pm$2σ
- ve %99 u $\pm$ 3σ standart sapma aralığında yer almasıdır.

olasılık ve şans

Normal Dağılım

Simetrik dağılım eğrilerinde:

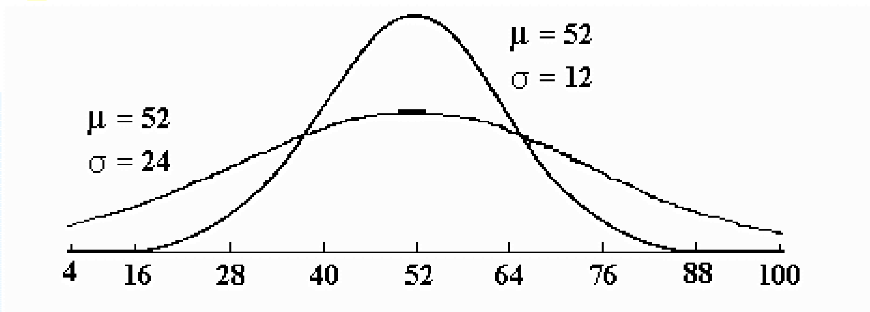


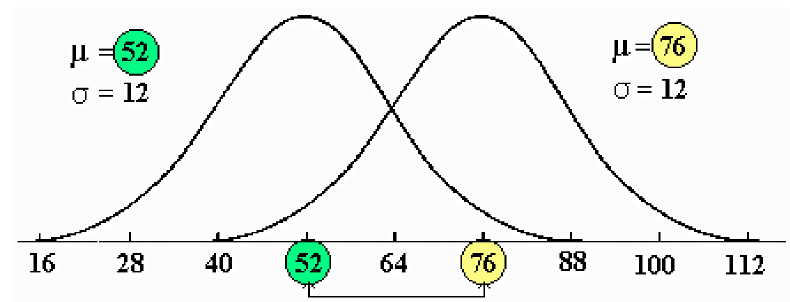
populasyon ortalaması= mod=orta değer

μ populasyonun ortalaması
σ ise standart sapma değeridir

olasılık ve şans

Normal Dağılım



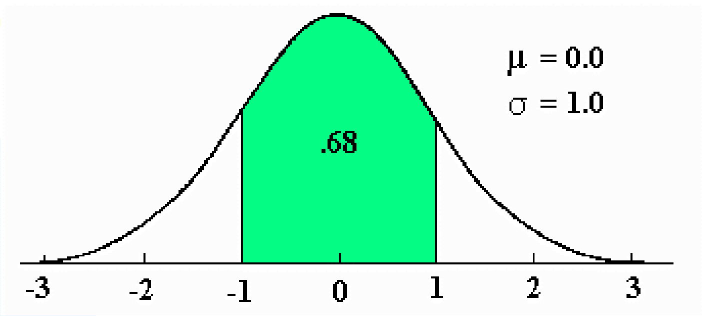


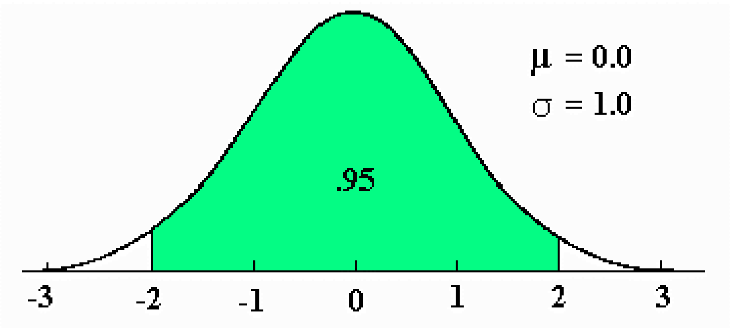
μ daki değişim normal dağılım eğrilerinin yüksekliklerini etkilerken

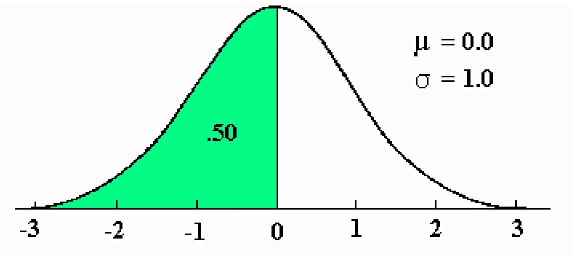
σ daki değişim ise normal dağılım eğrisinin yayılımını etkiler.

olasılık ve şans

Normal Dağılım





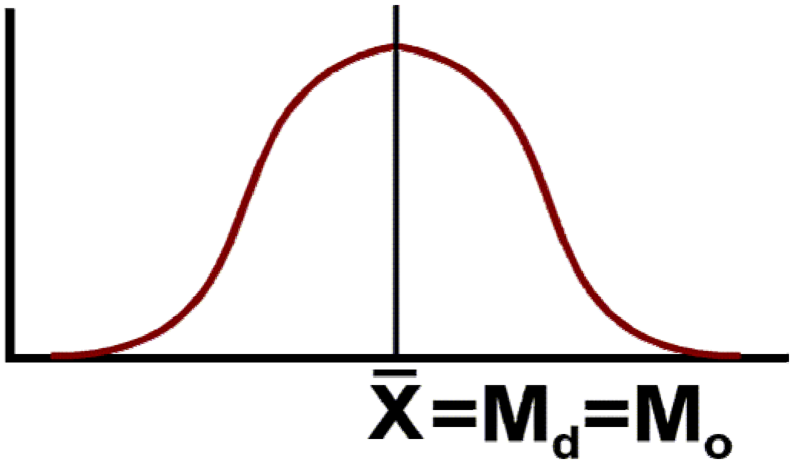


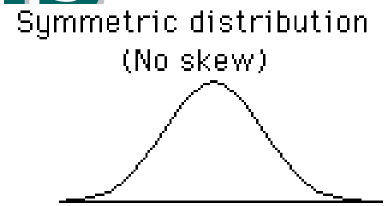
Normal dağılım eğirilerinin karakteristik özelliklerinden bir diğeri de normal dağılımdaki gözlemlerin veya verilerin yaklaşık

- %68 i $\pm 1\sigma$
- %95 i $\pm 2\sigma$
- ve %99 u $\pm 3\sigma$ standart sapma aralığında yer almasıdır.

olasılık ve şans

Dağılım eğrileri





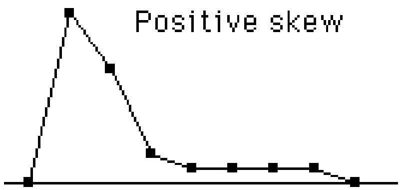
Simetrik dağılım diyagramlarında verilerin ortalaması, orta değer ve mod birbirlerine eşittir.

Verilerin ortalama değeri, modu ve orta değerini grafiklediğinden verilerin merkezi eğilimlerini (cental tendency) gösterir.

İkinci olarak verilerin dağılımı veya yayılımı ile ilgili olarak bilgi verir ve bunu veri aralığını ve standart sapma değerini içererek yapar.

Bununla beraber dağılım diyagramının şekline bağlı olarakta verilerin simetrikliği, kurtosis ve skewness olup olmadığı ile ilgili bilgi alınır.

olasılık ve şans

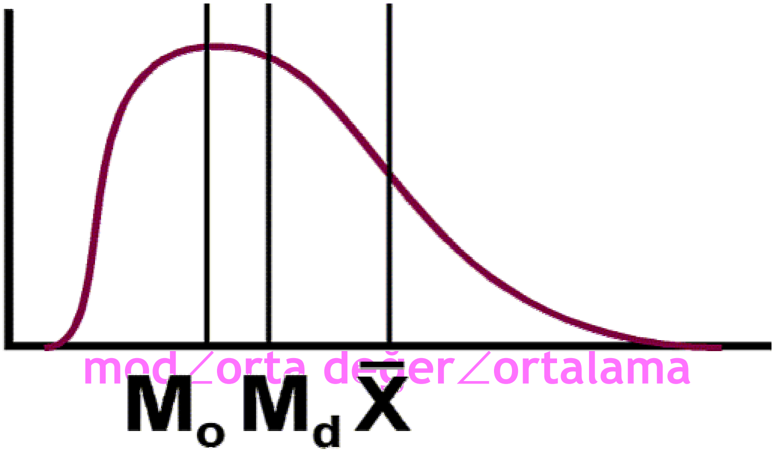


Dağılım eğrileri

Simetrik olmayan dağılım eğrileri

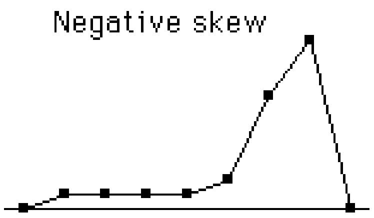
Diyagramlarında görülen asimetriklik çarpılma olarak da anılır ve bu çarpılma pozitif veya negatif yöne doğru olabilir.

- Pozitif olarak çarpılmış dağılım diyagramlarında mod orta değerden ve orta değer de ortalama değerden daha küçüktür.



mod∠orta değer∠ortalama

olasılık ve şans

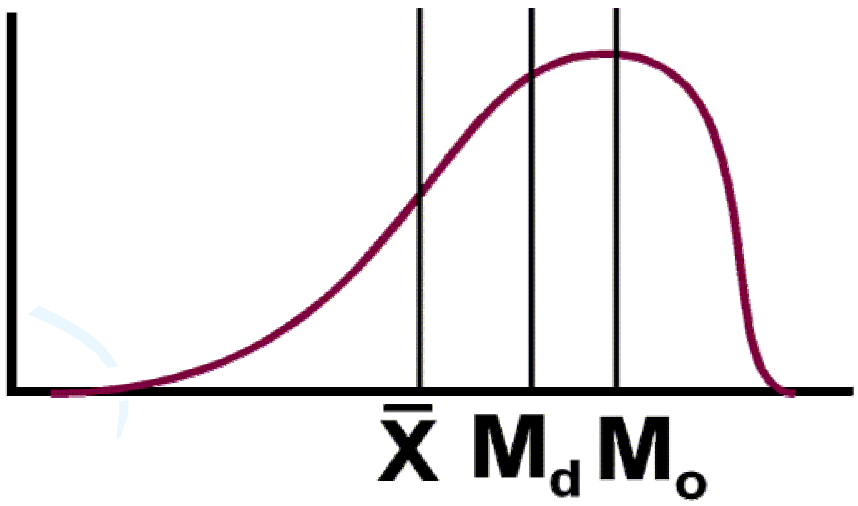


Dağılım eğrileri

Simetrik olmayan dağılım eğrileri

Skewness (yamulma)

Bu durumda dağılımın ortalaması dağılımın orta değerinden ve orta değerde dağılımın modundan küçüktür.

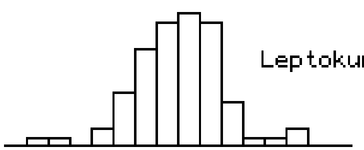


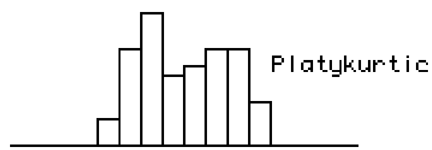
mod>orta değer > ortalama

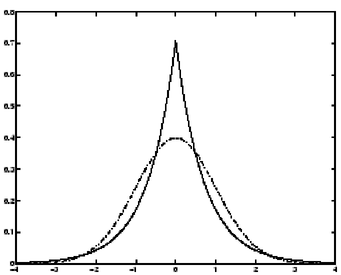
olasılık ve şans

Dağılım eğrileri

Kurtosis : Yunanca da dışbükey veya eğrilik

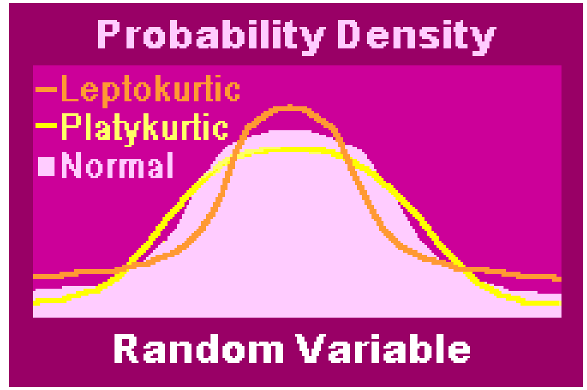






Leptokurtosis and Platykurtosis

Exhibit 1



Kurtosis dağılımın kısalığı ve düzlüğü veya uzunluğu ve sivriliği ile ilgilenir

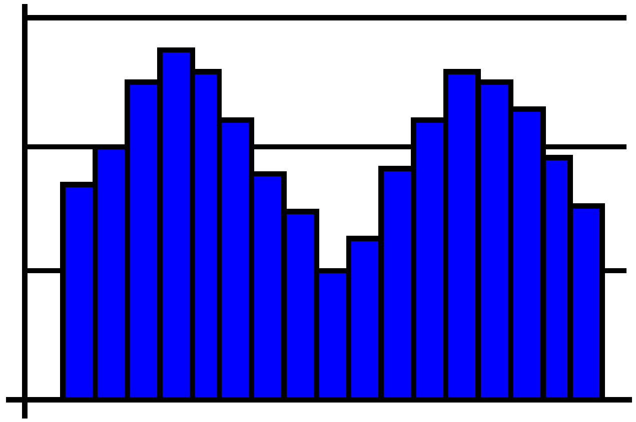
Kısa ve düzse platykurtic yani yassı olarak eğrilmiş
Sivri ise leptokurtic sivri olarak eğrilmiş
Simetrik dağılımlara ise meso kurtic veya orta eğrilmiş

olasılık ve şans

Dağılım eğrileri

iki modlu simetrik dağılımlar

Bu eğrilerde mod ve orta değer eşit olmasına rağmen dağılımda iki adet aynı freakansa sahip değer bulunur

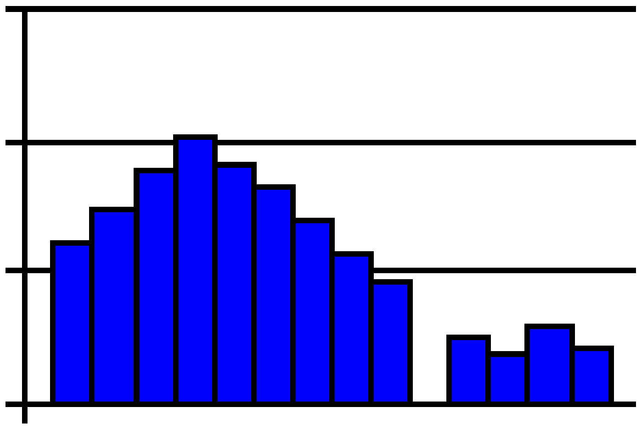


olasılık ve şans

Dağılım eğrileri

Simetrik olmayan dağılım eğrileri

Dışardakiler

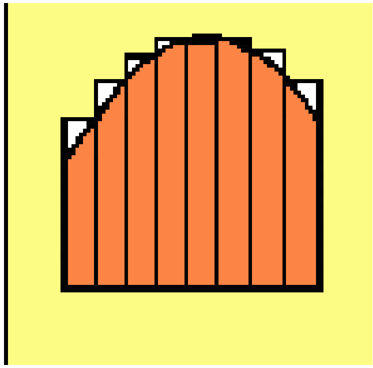


olasılık ve şans

Olasılık Kavramı ve Normal Dağılım Eğrisi

Normal eğri veya normal dağılım eğrisi bir tek eğriden değil gerçekte aynı alcebrik eşitlik tarafından tanımlanan sonsuz sayıda olasılık eğrisinden oluşur.

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



X : sebep değişken

π = 3.14

e =2.81

σ : standart sapma

μ :ortalama

olasılık ve şans

Olasılık Dağılımlarının Ortak Özellikleri

Değişkenler dağılım diyagramlarının vazgeçilmez parçasıdırlar.

- Dağılım diyagramlarının rölatif frekans toplamları 1'e eşittir.

- Dağılım diyagramlarında iki değer arasındaki alan olasılık değerini verir. Bu ana kadar yapılan tanımlardan anlaşılacağı üzere normal dağılımlarda tekil ölçümler yerine devamlı ölçümlerden bahsedilmektedir. Bu nedenle dağılım eğrileri tekil verilerin rölatif olasılıklarından değil, belirli bir aralıkda yeralan değerlerin olasılığından bahsetmektedir. Aksi takdirde tekil değerlerin söz konusu olduğu olasılıklarda normal dağılım eğrileri altındaki toplam alanların bire eşit olmaları olanaksızdır.