ÖLÇME SONUÇLARI ÜZERİNE İSTATİSTİKSEL İŞLEMLER

Sıklık (Frekans) Dağılımları

Örnek: 10 puanlık bir istatistik sınavında N=20 öğrencinin aldıkları puanlar aşağıdaki gibidir.

Puanlar: 8, 9, 8, 7, 10, 9, 6, 4, 10, 8
7, 8, 10, 9, 8, 6, 9, 7, 8, 7

Tablo 3. Frekans Tablosu Oluşturma

X	f
10	3
9	4
8	6
7	4
6	2
4	1
	$\sum f = N = 20$

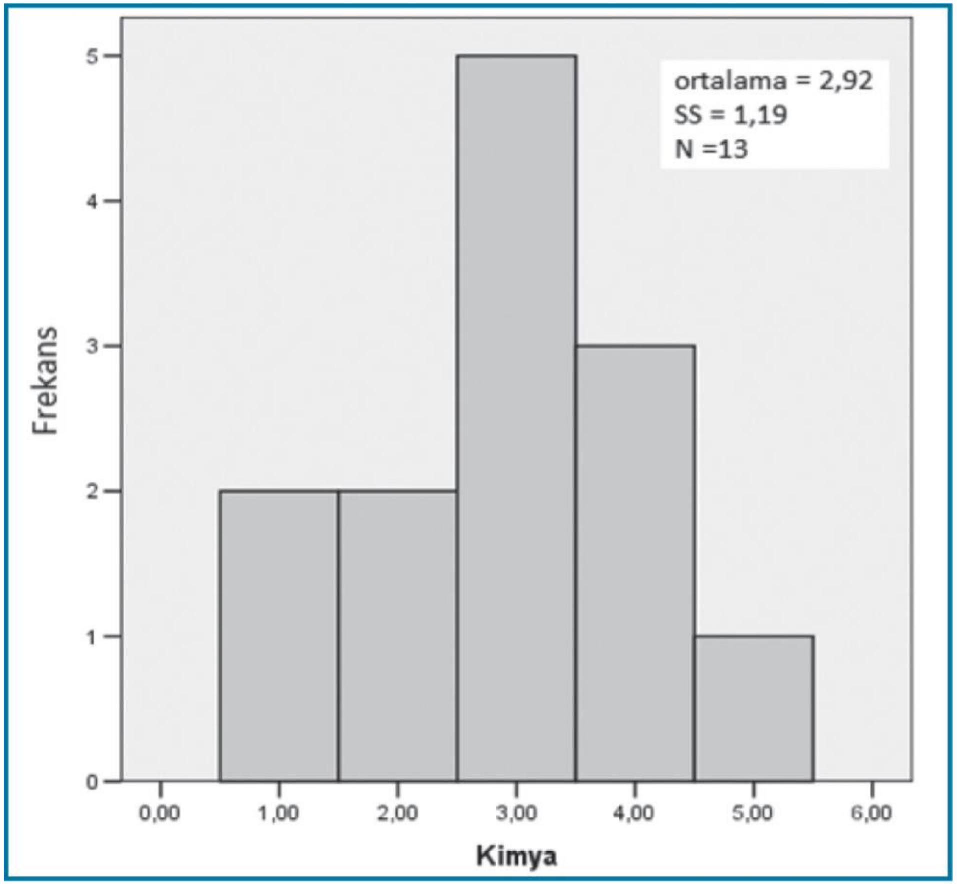
$\sum fX$'in frekans dağılım tablosundan elde edilmesi: her bir puan tekrar eden sayısı ile çarpılarak üçüncü sütunda gösterilir ve bu sütundaki sayıların toplamı $\sum fX$ e yani $\sum X$ e eşittir.

Toplam Frekans ve Yüzdeler

Puan (X)	Frekans (f)	Toplamlı Frekans (tf)	Frekans Yüzdesi (%)	Toplamlı Frekans Yüzdesi (% tf)
26	2	2	6,7	6,7
45	3	5	10	16,7
46	3	8	10	26,7
50	2	10	6,7	33,4
57	2	12	6,7	40,1
58	4	16	13,3	53,4
61	3	19	10	63,4
64	2	21	6,7	70,1
66	3	24	10	80,1
70	3	27	10	90,1
82	1	28	3,3	93,3
90	2	30*	6,7	100
Toplam	30			

* Ölçüm Sayısı

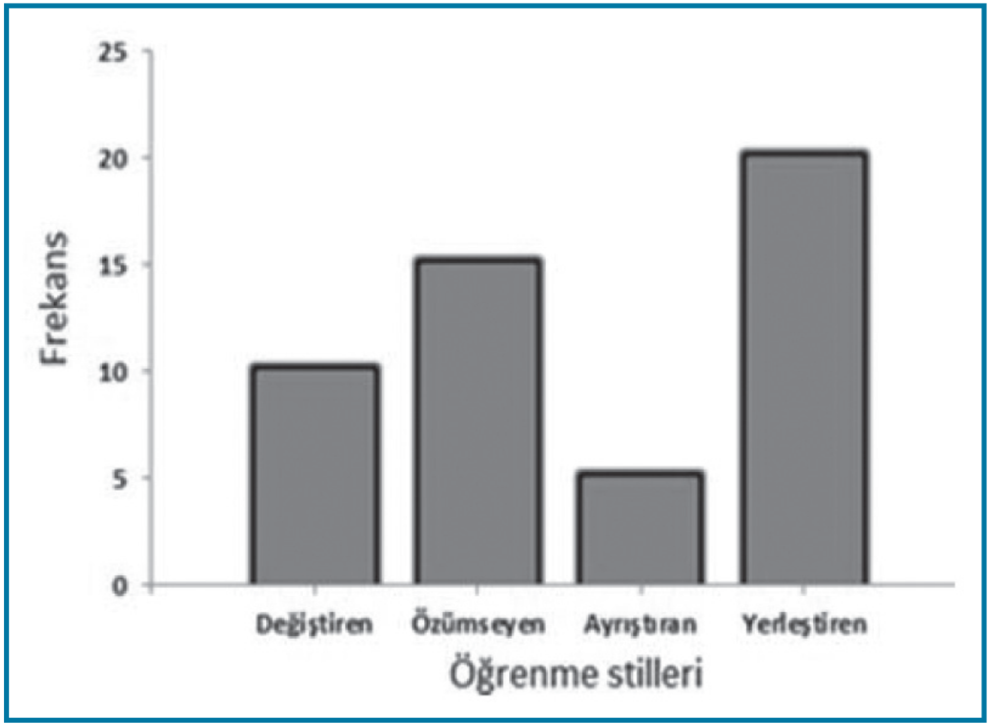
Histogram



Grafik 1. Histogram Grafiği Örneği

Tablo 9. Grafik 3'de Kullanılan Veriler

X	f
Değiştiren	10
Özümseyen	15
Ayrıştıran	5
Yerleştiren	20



Grafik 3. Bar Grafiği Örneği

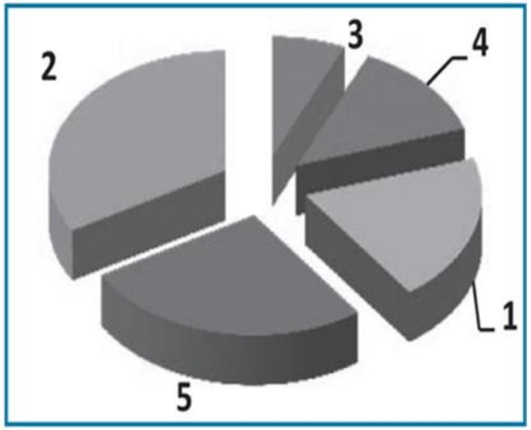
- **Pasta Grafikleri**: Çoğunlukla kategorik verilerin gösterilmesinde ve karşılaştırılmasında kullanılan, dairenin pasta dilimi şeklinde ayrılmasıyla oluşturulan gösterim şeklidir.

Örnek: Ortaokul 6. sınıf öğrencilerinin babalarının meslek durumunu gösteren veri-lerin aşağıdaki tablodakiler olduğunu düşünelim.

Tablo 10. 6. Sınıf Öğrencilerin Baba Meslek Durumları

X	f
1 (öğretmen)	4
2 (hukukçu)	6
3 (doktor)	1
4 (mühendis)	3
5 (işçi)	5

Bu verilere ait pasta grafiği Grafik 4'de verilmiştir.

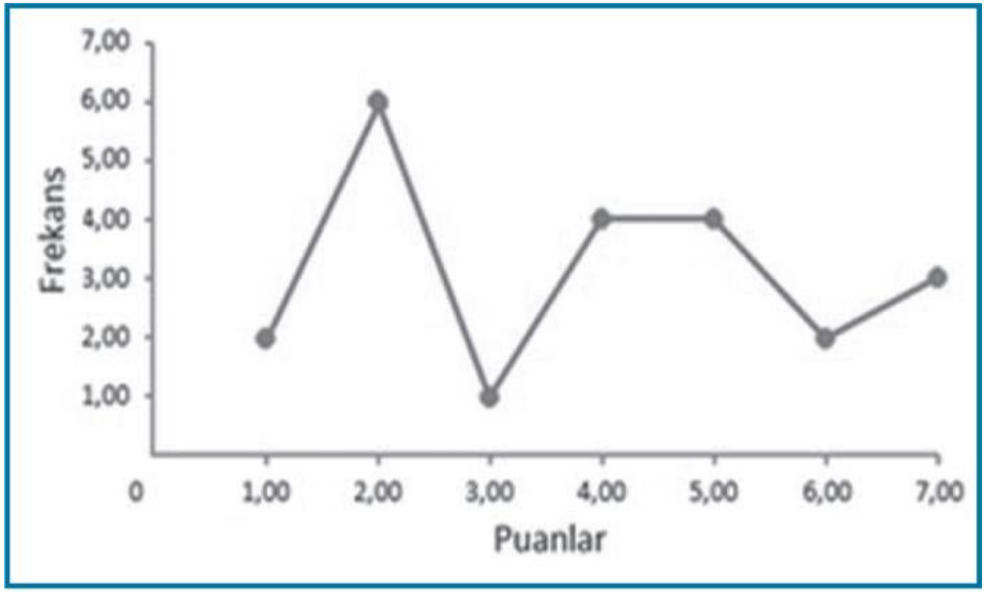


Grafik 4. Pasta Grafiği Örneği

Frekans Poligonları

Tablo 12. Grafik 6'da Kullanılan Veriler

X	f
7	3
6	2
5	4
4	4
3	1
2	6
1	2



Grafik 6. Frekans Poligonu Grafiği Örneği

MERKEZİ EĞİLİM ÖLÇÜLERİ

- **Mod**
- **Medyan**
- **Aritmetik ortalama**

Merkezi Eğilim Ölçüleri

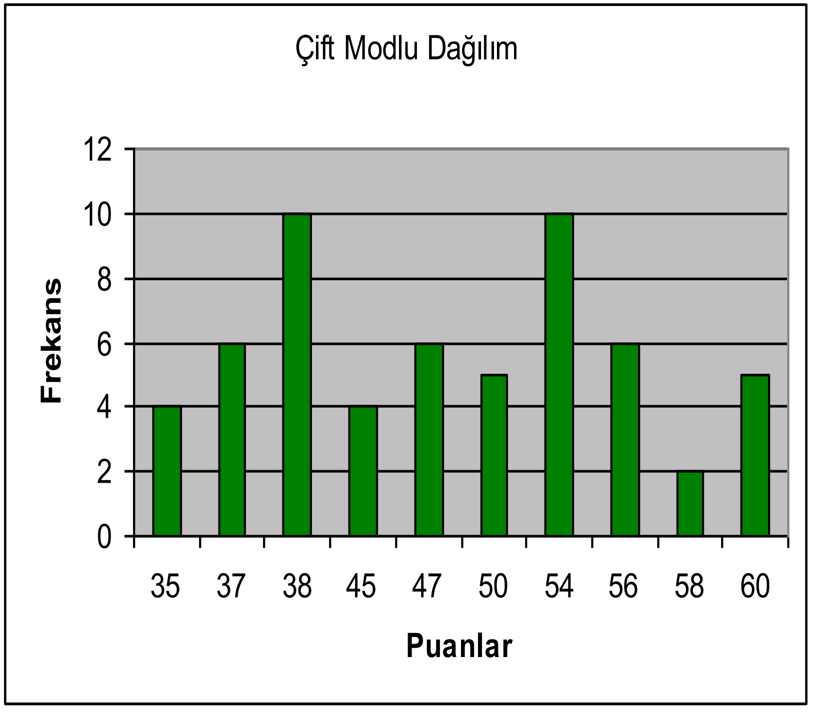
- Mod (Tepe Değeri): Bir dağılımda en sık gözlenen, frekansı en yüksek olan puandır.
- Modu tanımlamak için herhangi bir sembol yoktur. Ayrıca örneklem modu ve evren modu arasında bir ayrım yoktur.
- Mod herhangi bir ölçekteki tipik veya ortalama değeri belirlemek için kullanıldığından kullanışlı bir merkezi eğilim ölçüsüdür.

Restaurant	f
College Grill	5
George & Harry's	16
Luigi's	42
Oasis Diner	18
Roxbury Inn	7
Sutter's Mill	12

n = 100 öğrencilik bir örneklemin favori restaurantı listelenmiştir.
Mod bir kategori veya bir puandır, bir frekans değildir. Bu durumda verilen örnek için mod Luigi'stir, *f* = 42 değildir. Böylece favori restaurant olarak en sık listelenen yer Luigi'stir.

- Bir dağılımın birden fazla moda sahip olması mümkündür. İki moda sahip bir dağılım iki modlu, ikiden fazla moda sahip olan bir dağılımsa çok modlu dağılım olarak adlandırılır.

Puan (X)	Frekans (f)
35	4
37	6
38	10
45	4
47	6
50	5
54	10
56	6
58	2
60	5



Verilen puan dağılımında 2 mod değeri vardır: 38 ve 54 dağılımın modlarıdır.

Merkezi Eğilim Ölçüleri

- <u>Medyan (Ortanca)</u>: Bir dağılımda puanlar küçükten büyüğe veya büyükten küçüğe sıralandığında dağılımın tam ortasındaki puandır. Dağılımdaki bireylerin %50'si medyanda veya medyanın altında puana sahiptir.

- Medyanı tanımlamak için herhangi bir sembol yoktur. Ayrıca bir örneklem ve bir evren için medyanın tanımı ve hesaplamaları aynıdır.

- Medyan değeri dağılımda n birey sayısını ifade etmek üzere, $[(n + 1)/ 2]$. bireye karşılık gelen puandır.

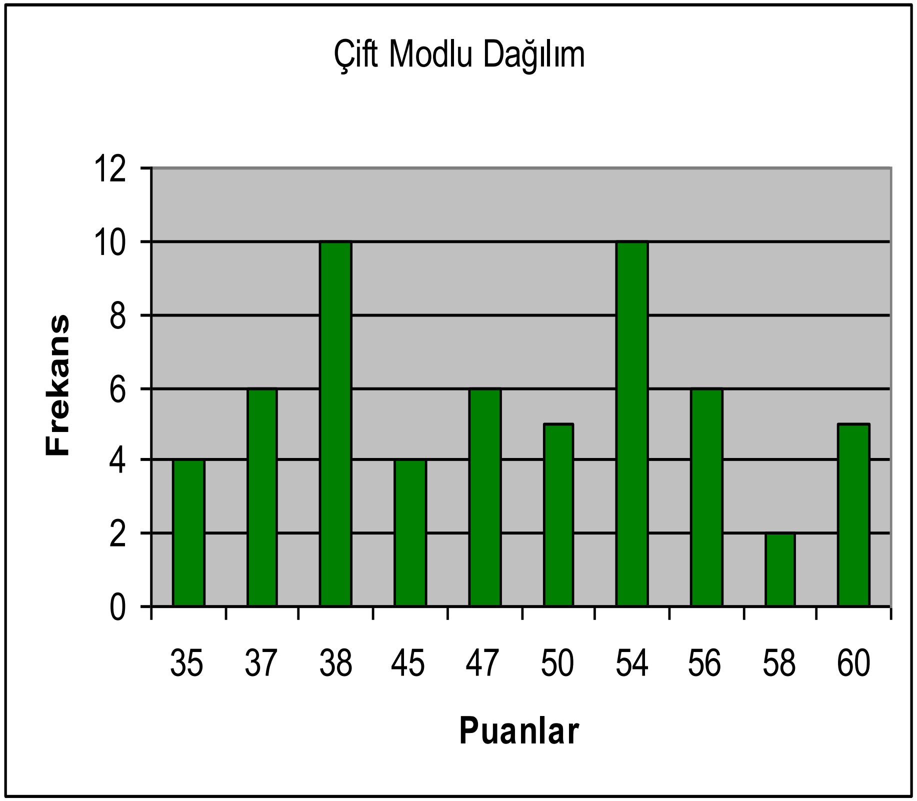
- Örnek:
  - Aşağıda 5 öğrencinin aldığı puanlar verilmiştir:

    8, 5, 11, 3, 10
  - Bu dağılımın medyanının hesaplayabilmek için ilk olarak puanların sıraya dizilmesi gerekmektedir.

    3, 5, 8, 10, 11
  - Dizilen puanlardan tam ortadaki puan medyandır: $[(5 + 1)/ 2] = 3$, soldan veya sağdan 3. öğrenciye karşılık gelen puan $X = 8$ olup medyan 8'e eşittir.
- Örnek:
  - Aşağıda 6 öğrencinin aldığı puanlar verilmiştir:

    3, 3, 4, 5, 7, 8
  - Dizilen puanlardan tam ortadaki puan medyandır: $[(6 + 1)/ 2] = 3,5$, 3. ve 4. öğrencilere karşılık gelen puanlar sırasıyla 4 ve 5 olup medyan $4 + 5/2 = 4.5$'e eşittir.

Puan (*X*)	Frekans (*f*)	Yığılmalı Frekans (*Yf*)
35	4	4
37	6	10
38	10	20
45	4	24
47	6	30
50	5	35
54	10	45
56	6	51
58	2	53
60	5	58



Verilen puan dağılımında 58 birey olup [(*58* + 1)/ 2] = 29,5, 29. ve 30. öğrencilere karşılık gelen puanlar sırasıyla 47 ve 47 olup medyan 47 + 47/2 = 47'ye eşittir.

Merkezi Eğilim Ölçüleri

- Ortalama: Bir dağılımdaki bütün puanların toplanıp toplamın puan sayısına bölünmesiyle hesaplanır.
- Bir evren için ortalama μ ile bir örneklem içinse $\overline{X}$ ile gösterilir.
  - Evren ortalamasının formülü aşağıdaki gibidir:

$$\mu = \frac{\sum_{i=1}^{N} X_i}{N}$$

  - Burada
    - X_i: i bireyinin testten elde ettiği puan
    - N: Evrendeki birey sayısı

- Örneklem ortalamasının formülü aşağıdaki gibidir:

$$\overline{X} = \frac{\sum_{i=1}^{N} X_i}{n}$$

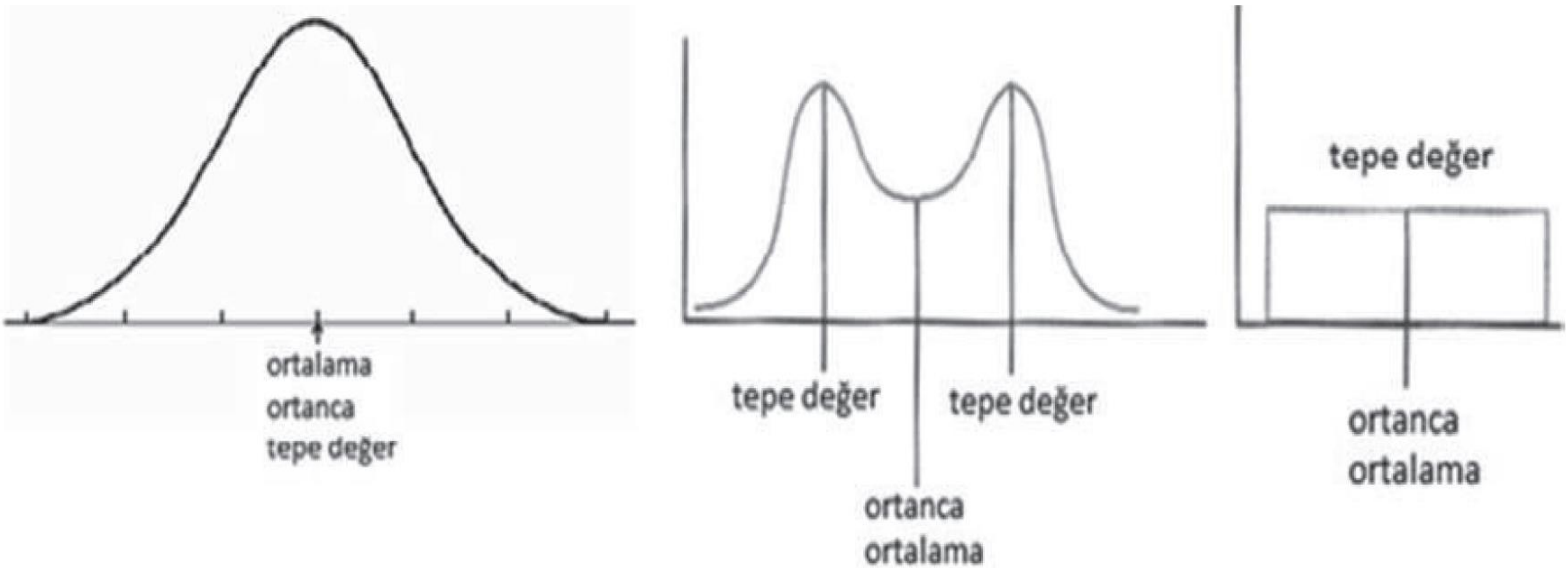
- Burada
  - *Xi*: *i* bireyinin testten elde ettiği puan
  - *n*: Örneklemdeki birey sayısı
- Örnek:

  Asağıda bir örneklemde bulunan 20 öğrencinin puanları verilmiştir:

  12 15 18 10 14 14 15 16 16 20 12 14 15 18 14 15 16 16

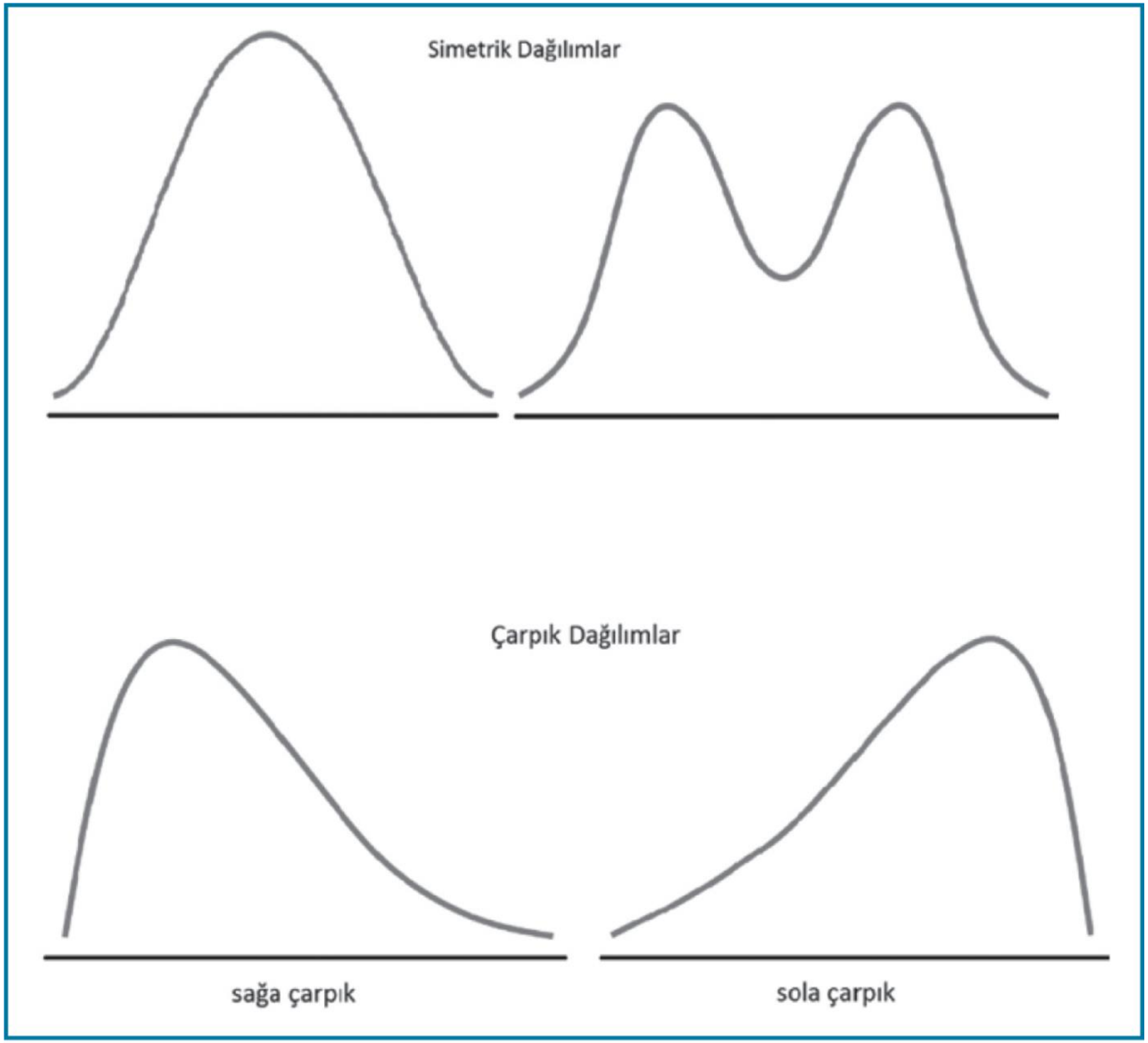
$$\overline{X} = \frac{12 + 15 + 18 + \ldots + 16 + 15 + 15}{20} = 15$$

Simetrik Dağılım



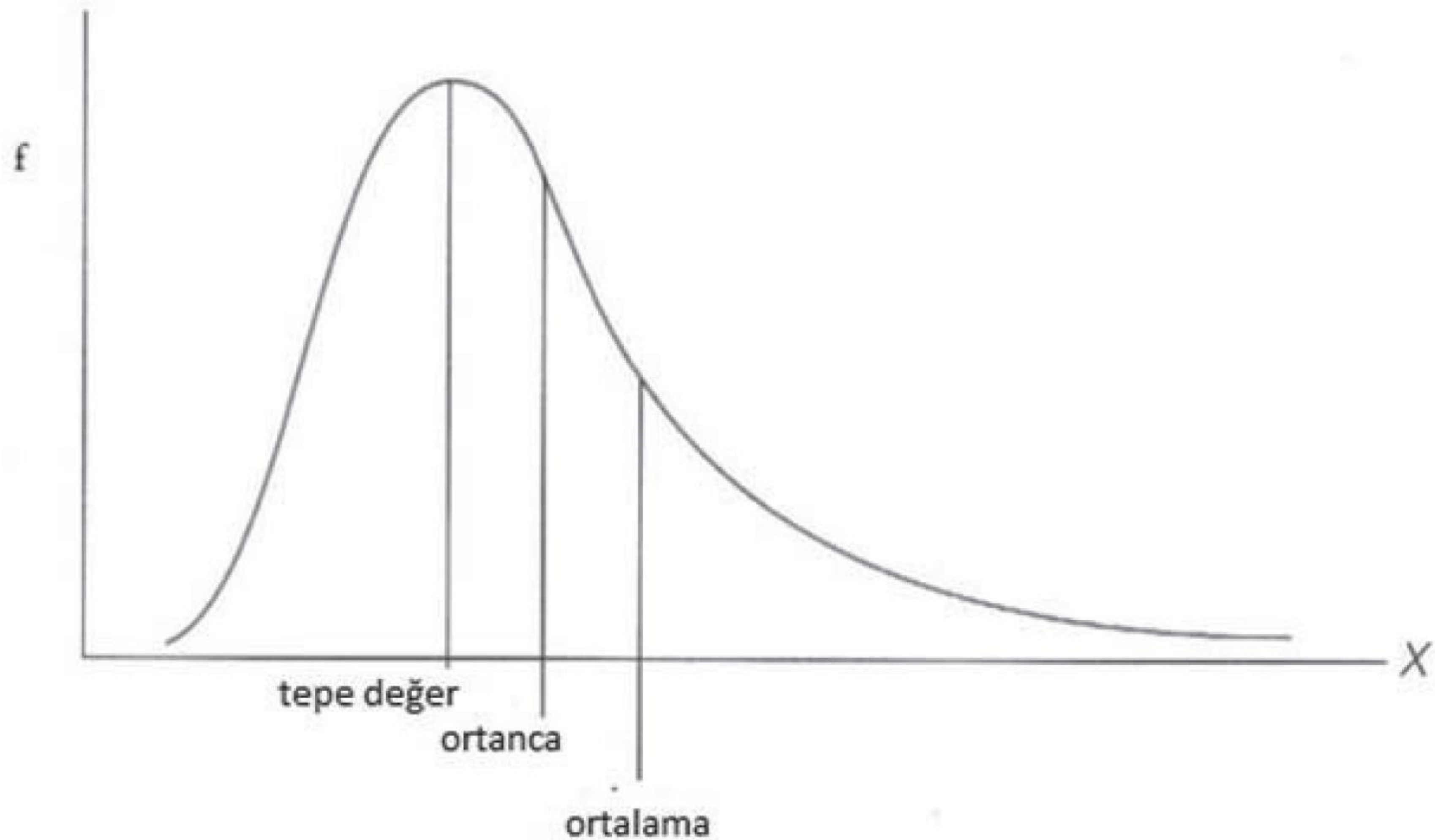
Grafik 11. Simetrik Dağılım Özellikleri

• Simetrik ve Simetrik Olmayan Frekans Dağılım Şekilleri

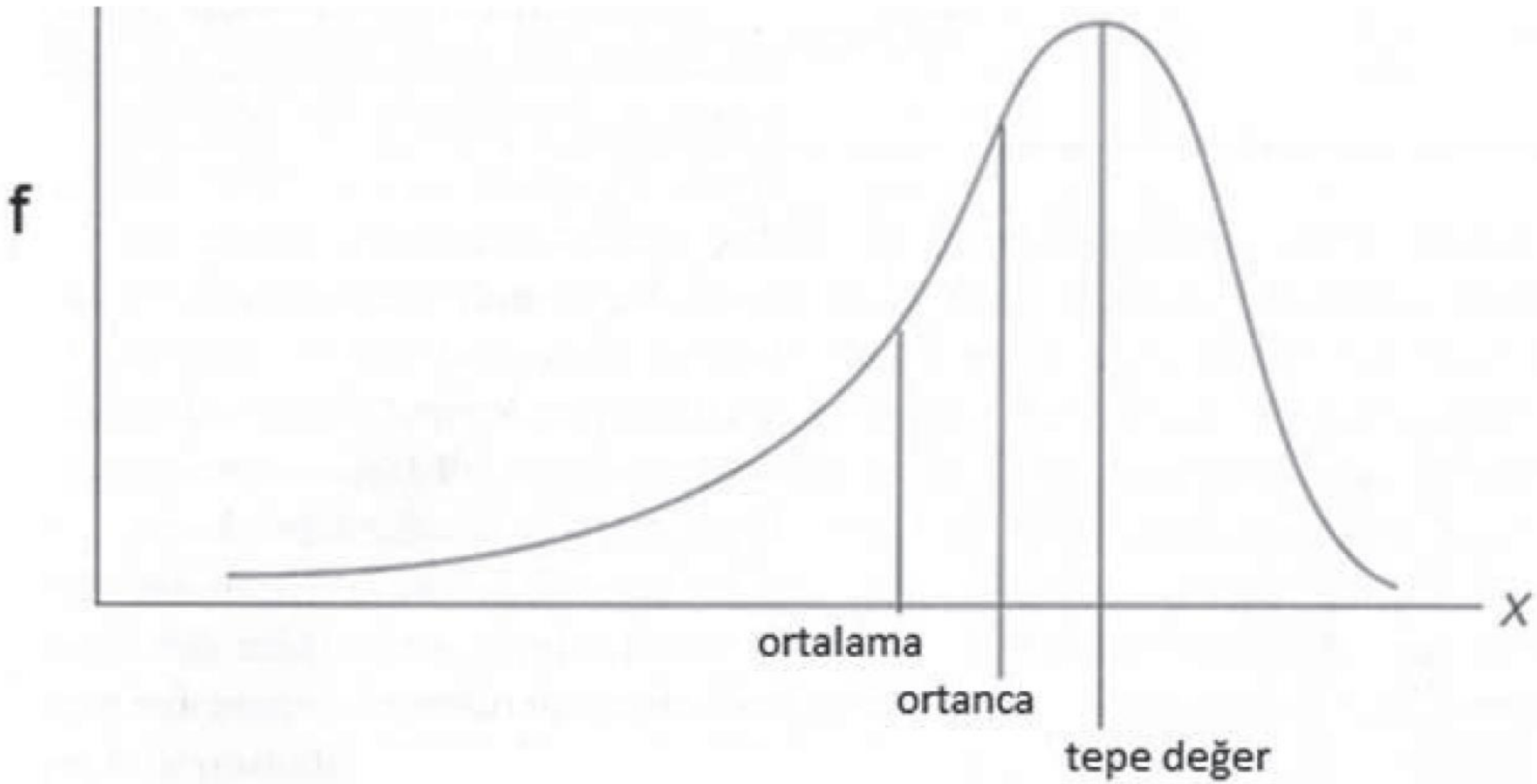


Şekil 1. Dağılım Şekillerine Örnekler

- ***Çarpık Dağılımlar***



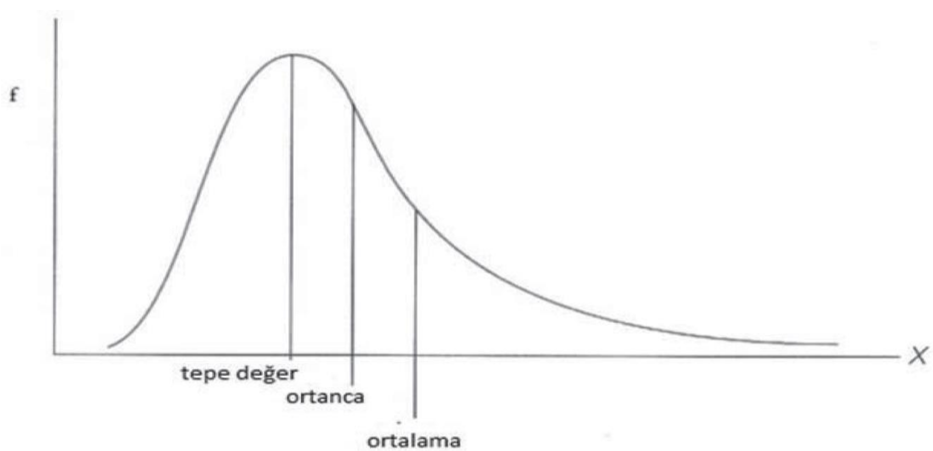
Grafik 12. Sağa Çarpık Dağılım



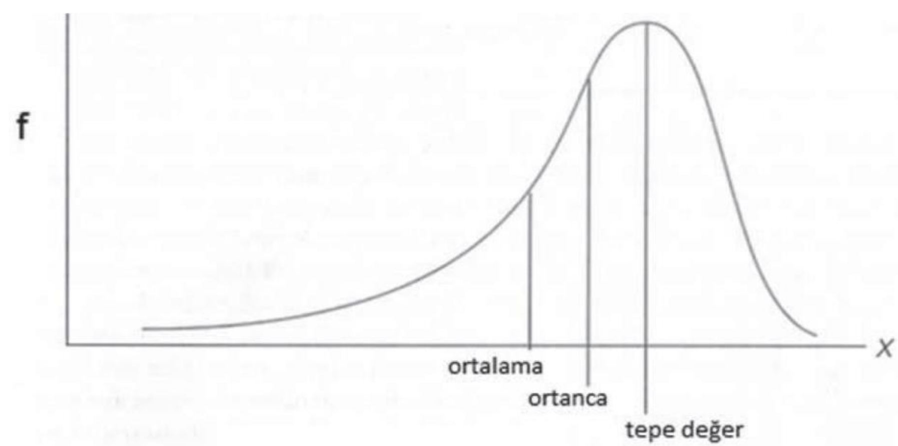
Grafik 13. Sola Çarpık Dağılım

Çarpık dağılımlar: Pozitif çarpık (sağa çarpık) dağılımlarda mod solda, tepe noktasındaki değerdir. Dağılımı eşit bir şekilde iki parçaya ayıran değer, yani ortanca, modun sağ tarafında olacaktır. Aritmetik ortalama az sayıdaki uç değerden etkileneceği için ortancanın daha sağında konumlanacaktır.

Negatif çarpık (sola çarpık) dağılımlarda ise mod sağda tepe noktasındaki değerdir. Bu sefer dağılımı eşit bir şekilde iki parçaya ayıran değer, yani ortanca, modun sol tarafında yer alır. Aritmetik ortalama az sayıdaki uç değerlerden etkileneceği için ortancanın daha solunda konumlanacaktır.



Grafik 12. Sağa Çarpık Dağılım



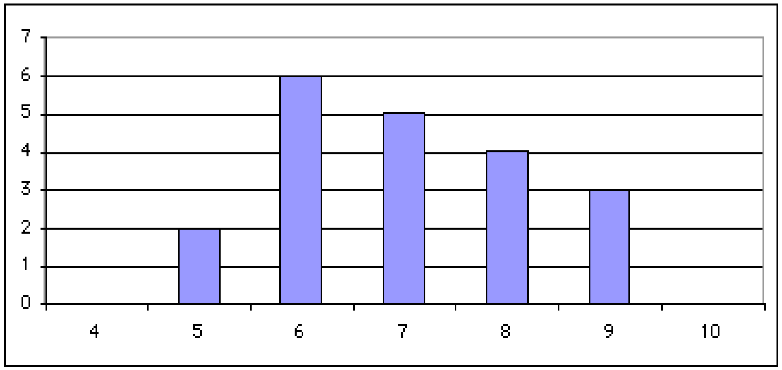
Grafik 13. Sola Çarpık Dağılım

DAĞILIM (YAYILMA, DEĞİŞİM) ÖLÇÜLERİ

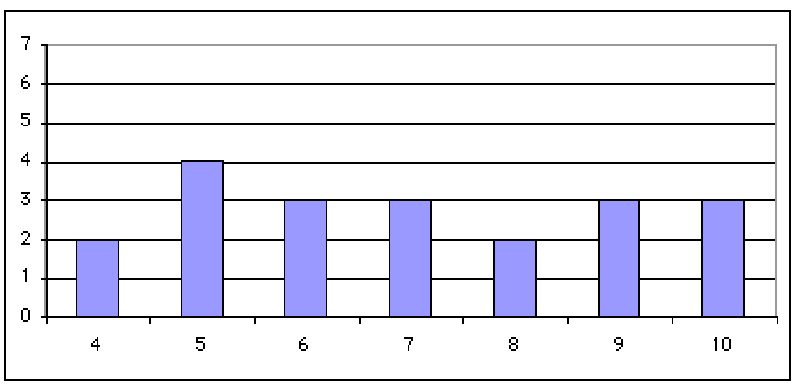
- **Ranj**
- **Çeyrek Sapma**
- **Varyans**
- **Standart Sapma**

Merkezi Değişim Ölçüleri

Quiz 1’den Elde Edilen Puanlara İlişkin Sütun Grafiği



Quiz 2’den Elde Edilen Puanlara İlişkin Sütun Grafiği



- Yandaki grafiklerden yukarıdaki grafik bir İngilizce dersinde bir grup öğrenciye uygulanan quiz 1´den, alttaki grafik ise aynı gruba uygulanan quiz 2´den alınan puanları göstermektedir.
- Quiz 1´den alınan puanların aritmetik ortalaması quiz 2´den alınan puanların aritmetik ortalamasına eşit olup 7´dir.
- İki quizden alınan puanların ortalaması birbirine eşit olmasına rağmen puanların dağılımı birbirinden oldukça farklıdır.

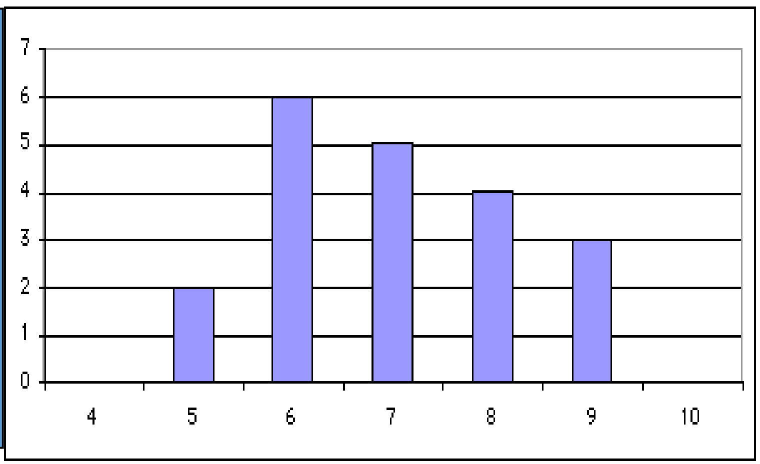
- Öğrencilerin puanları arasındaki fark quiz 2´de quiz 1´e göre daha fazladır.

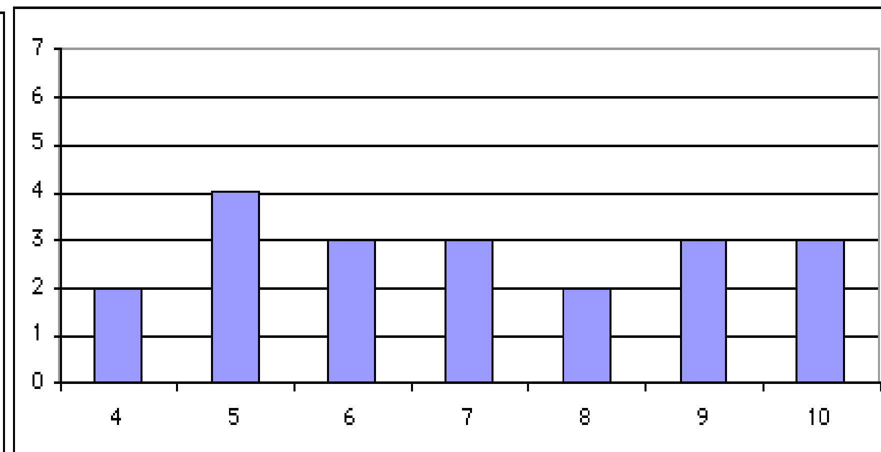
Merkezi Değişim Ölçüleri

- <u>Değişim Ölçüleri</u>
  - <u>Ranj</u>: Bir dağılımda elde edilen en yüksek puan (X_{max}) ve en düşük puan (X_{min}) arasındaki uzaklıktır.

<u>Örnek</u>: $$Ranj = X_{\max} - X_{\min}$$

Quiz 1



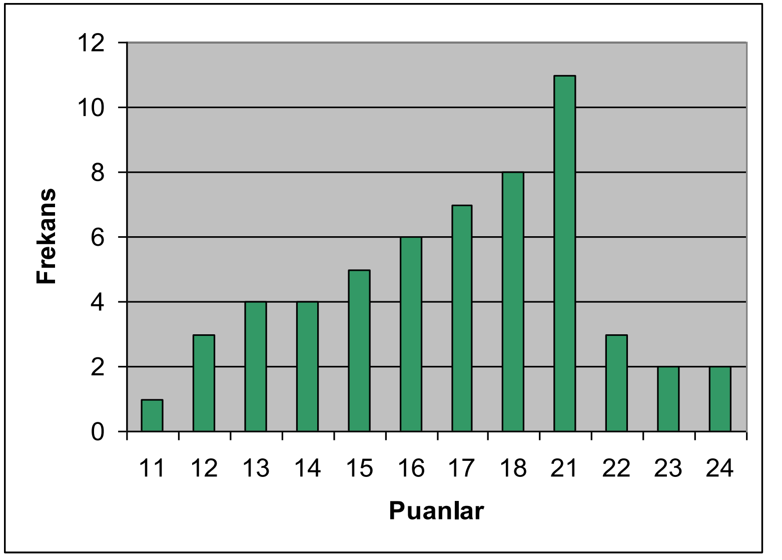


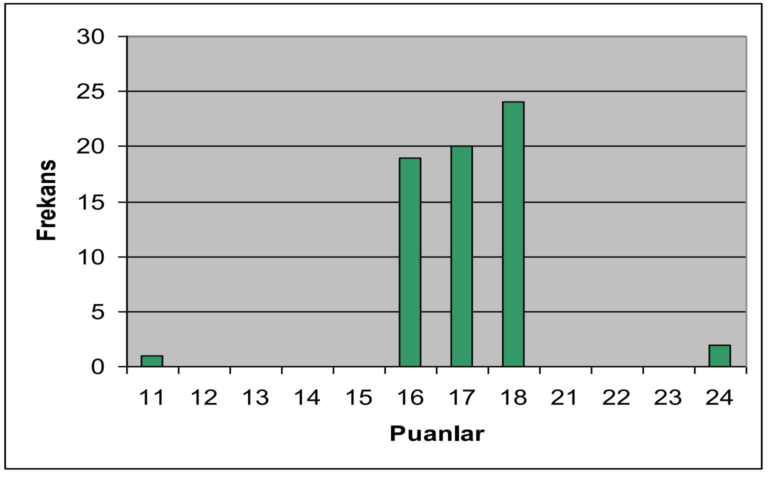
Quiz 2

Quiz 1´de, en düşük puan 5, en yüksek puan 9 olup ranj 9-5= 4´tür.
Quiz 2´de ranj daha büyüktür: en düşük puan 4 en yüksek puan, 10 olup ranj 10-4=6´dır.

Merkezi Değişim Ölçüleri

- Ranj tamamen iki uç değer kullanılarak ve dağılımdaki diğer puanlar göz ardı edilerek hesaplanan bir değişim ölçüsü olduğundan yeterince hassas ve kullanışlı değildir.
- Bir dağılımdaki uç değerlerin diğer puanlardan kopuk olması, puanların yayılımı hakkında yanıltıcı bilgi verir.
  - Yandaki iki grafikteki puan dağılımlarının ranjı 24-11=13 olup puanların 13 puanlık bir aralıkta dağıldığını ifade eder. Ancak birinci grubun daha heterojen, ikinci grubunsa daha homojen bir dağılım gösterdiği gözlenmektedir
- Ranj, gruptaki puanların birbirine ne kadar yakın veya uzak olduğu ile ilgili bir bilgi vermemektedir.





Çeyrek Sapma

Ortalama yerine ortanca kullanıldığında ya da veri setinde aşırı uç değerler bulunduğunda değişim genişliği yerine çeyrek sapma kullanılır. Çeyrek sapma Q ile gösterilir.

$$Q = \frac{Q_3 - Q_1}{2}$$

Eşitlikte,

Q : Çeyrek sapma

Q_1: Birinci çeyreklik

Q_3: Üçüncü çeyrekliktir.

Dağılımdaki bütün değerler kullanılmadığı için Q yeterli bir dağılım ölçüsü değildir.

Örnek 2.19: İlaçla tedavi edilen 8 hastanın iyileşme süreleri gün olarak aşağıda verilmiştir. Çeyrek sapma değerini hesaplayınız.

30, 20, 24, 40, 65, 70, 10, 62

Önce iyileşme sürelerini küçükten büyüğe sıraya dizelim:

$x_1,\ x_2,\ x_3, x_4, x_5, x_6,\ x_7,\ x_8$

10, 20, 24, 30, 40, 62, 65, 70

$n = 8, \quad çift$

$$Q_1 = \frac{X_{(2)} + X_{(3)}}{2} = \frac{20 + 24}{2} = 22$$

$$Q_3 = \frac{X_{(6)} + X_{(7)}}{2} = \frac{62 + 65}{2} = 63.5$$

O halde çeyrek sapma $Q = \frac{63.5-22}{2} = 20.75$ dir.

Merkezi Değişim Ölçüleri

- **Varyans ve Standart Sapma:** Standart sapma en sık kullanılan ve en önemli değişim ölçüsüdür.
- Standart sapma dağılımın ortalamasını referans noktası olarak kullanır ve her bir puan ve ortalama arasındaki uzaklığı ele alarak değişkenliği ölçer.

- Böylece puanların genel olarak ortalamaya ne kadar yakın veya ortalamadan ne kadar uzak olup olmadığını belirler. Diğer bir ifadeyle puanların bir arada toplanıp toplanmadığını veya yayılıp yayılmadığını belirtir.

- Özetle standart sapma ortalamadan ortalama uzaklığı tahmin eder.

$$Standart\ Sapma = \sqrt{Varyans} = \sqrt{\frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}} \qquad orneklem\ standart\ sapmasi = \sqrt{orneklem\ varyansı} = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n-1}}$$

- **Merkezi Değişim Ölçüleri**

- Standart sapma, sapma puanı kavramına dayanır. Bir grupta *X* ham puanına sahip her bir birey için sapma puanı, *x*, aşağıdaki şekilde tanımlanabilir:

$$x_i = X_i - \mu$$

  - Burada
    X_i: *i* bireyinin puanı
    μ: Ortalama puan
  - Böylece bir sapma puanı, her bir bireyin puanı ve grup ortalaması arasındaki uzaklıktır.

- Örnek:
- Puanların ortalamasının 50 (μ = 50) olduğu bir puan dağılımında 53 puan alan (X_1 = 53) bir bireyin sapma puanı

$$x_1 = X_1 - \mu = 53 - 50 = 3$$

Bireyin puanı ortalamanın 3 puan üstündedir.

- Puanların ortalamasının 50 (μ = 50) olduğu bir puan dağılımında 45 puan alan (X_2 = 45) bir bireyin sapma puanı

$$x_2 = X_2 - \mu = 45 - 50 = -5$$

Bireyin puanı ortalamanın 5 puan altındadır.

- Bir sapma puanının iki kısmı vardır: işaret (+ veya -) ve sayı.
  - İşaret ortalamadan olan uzaklığın yönünü belirtir. (+) işareti bireyin puanının ortalamanın üstünde, (-) işaretiyse altında olduğunu belirtir.
  - Sayı bireyin puanının ortalamadan olan gerçek uzaklığını verir.

- Standart sapma hesaplanırken amaç ortalamadan olan standart uzaklığın ölçümünü hesaplamak olduğundan, her bir birey için sapma puanı elde edildikten sonraki basamak sapma puanların ortalamasını hesaplamak olmalıdır.
  - Bu ortalamayı hesaplamak için önce sapma puanlar toplanır, sonra elde edilen toplam birey sayısına bölünür.
  - Örnek:

Verilen 4 puanın toplamı 12 olup ortalaması 3'tür.

X	X - μ
8	+5
1	-2
3	0
0	-3
$12 = \sum_{i=1}^{4} X_i$	$0 = \sum_{i=1}^{4} (X_i - \mu)$

Sapma puanları hesaplamak için her bir puandan 3 çıkarılır. Verilen 4 sapma puanın toplamı 0 olup ortalaması 0'dır.

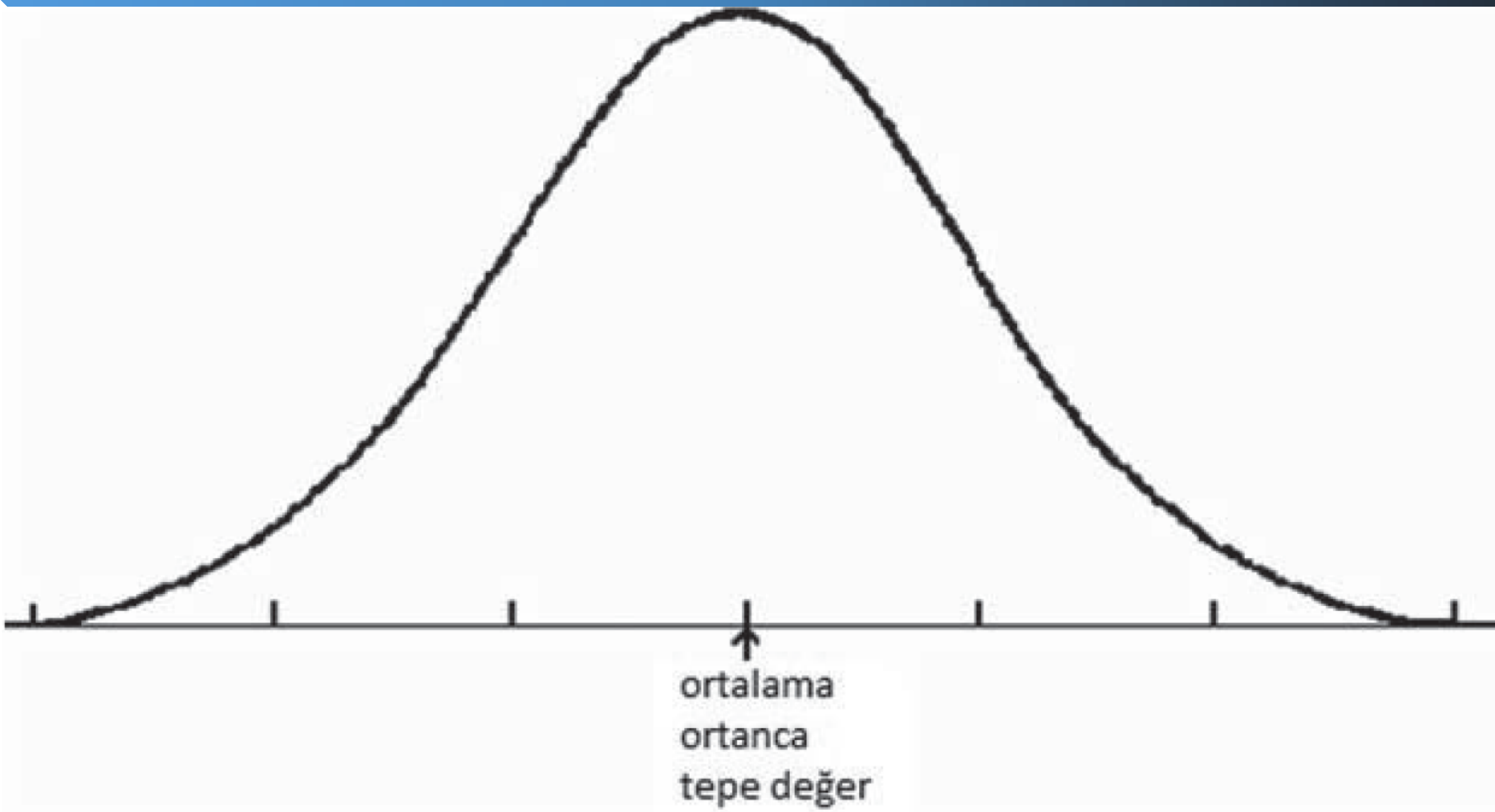
- Ortalama dağılımın denge noktası olduğundan, ortalamanın üstündeki uzaklıkların toplamı, ortalamanın altındaki uzaklıkların toplamına eşittir. Bu nedenle sapma puanların toplamı HER ZAMAN sıfıra eşittir.

- Sapma puanların ortalaması her zaman sıfır olduğundan, sapma puanların ortalaması değişkenlik ölçüsü olarak kullanılamayacaktır.
  - Sapma puanların ortalamasının sıfır olması problemi artı ve eksi değerlerin birbirini iptal etmesinden kaynaklanır. Çözüm işaretlerden (+ ve -) kurtulmaktır. Bunu gerçekleştirmek için kullanılan standart yöntem her bir sapma puanın karesini almaktır.
  - Sapma puanların karelerinin ortalaması varyans olarak adlandırılır ve evren için σ^2 ile örneklem için s^2 ile gösterilir. Varyans ortalamadan olan ortalama kare uzaklığı ifade eder.
  - Örnek:

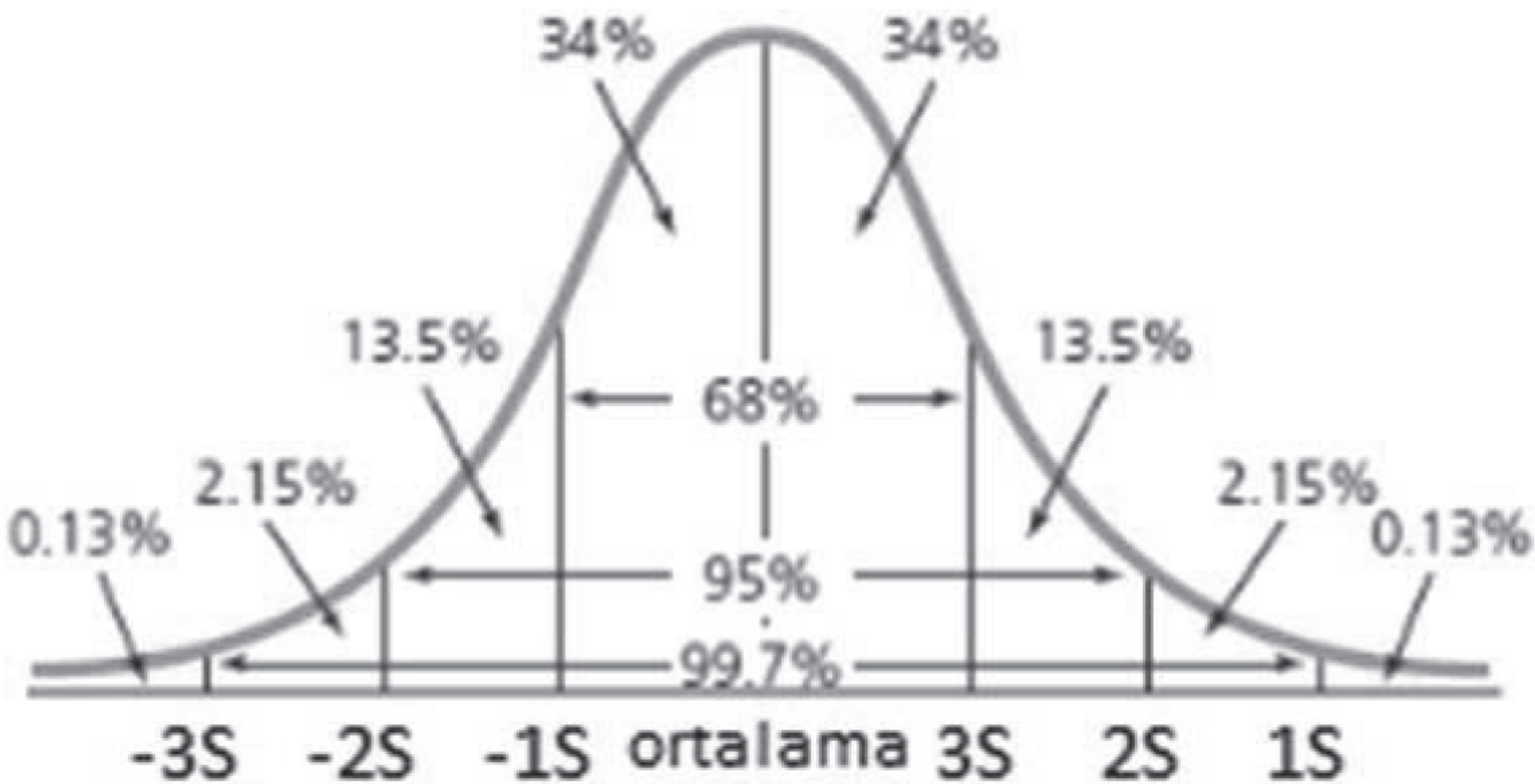
X	X-μ	$(X-\mu)^2$
8	+5	25
1	-2	4
3	0	0
0	-3	9
$12=\sum_{i=1}^{4} X_i$	$0=\sum_{i=1}^{4}(X_i-\mu)$	$38=\sum_{i=1}^{4}(X_i-\mu)^2$

Sapma puanların kareleri toplamı 38 olup ortalaması 38/4 = 9.5'tir. Elde edilen 9.5 değeri evren için varyanstır.

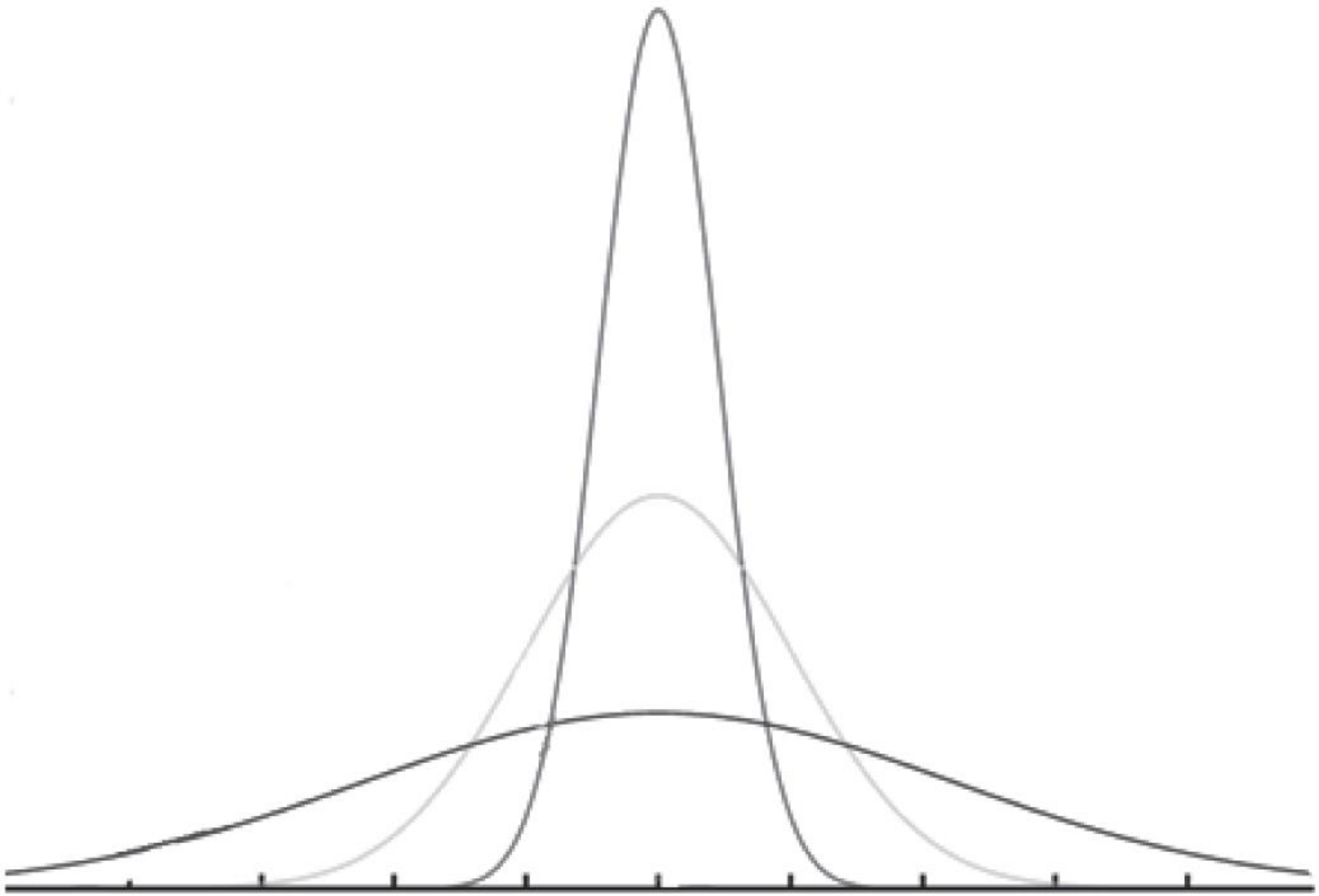
NORMAL DAĞILIM



Şekil 2. Normal Dağılım Eğrisi



Şekil 3. Normal Dağılım Eğrisi Aralıkları ve Yüzdeleri



Şekil 5. Normal Dağılım Örnekleri