TEMEL İSTATİSTİK

Değişkenlik Ölçüleri

Prof. Dr. Ezel Tavşancıl

Geçen haftadan hatırlatma…

- Önceki hafta, betimsel istatistiklerden bir olan dağılımın betimlenmesi ve grafik ya da tablo ile gösterilmesi anlatıldı.
- Geçen haftanın konu başlığı betimsel istatistiklerden biri, merkezî eğilim ölçüleri idi. Bu hafta bir diğeri olan **değişkenlik ölçüleri** işlenecek.
- Merkezi eğilim ölçüleri, frekans dağılımındaki bir noktanın ölçekteki değerini; değişim ölçüleri ise bir aralığın değerini verir (Büyüköztürk, Çokluk ve Köklü, 2018).

DEĞİŞKENLİK ÖLÇÜLERİ

- DİĞER İSİMLENDİRMELER:
  - Değişme ölçüleri (Baykul, 1999).
  - Değişkenlik ölçüleri, (Tan, 2016; Büyüköztürk, Çokluk ve Köklü, 2018)
  - Değişim ölçüleri (Arıcı, 1998).
- Bir dağılımda ölçümler arasında gözlenen farklılık ve değişikliğe değişim, veriler arasındaki değişimden kaynaklanan farklılıkların istatistiksel ölçülerine ise değişim ölçüleri denir (Büyüköztürk ve diğerleri, 2013).
- Değişim ölçüleri, bir merkezi eğilim ölçüsünden olan sapma miktarlarını gösterir ve araştırmalarda değişim ve onun kaynağı üzerinde durulur (Arıcı, 1998; Balcı, 1995; Howitt & Cramer, 1997).
- Kullanılan değişkenlik ölçüleri:
  - Ranj
  - Varyans
  - Standart Sapma
  - Çeyrek Sapma
  - Değişim katsayısı
  - Çarpıklık ve basıklık

Değişim Ölçüleri Neden Gerekli?

1. Merkezî eğilim ölçüleri, puanların dağılımında yetersiz kalmaktadır.

2. İki ya da daha çok grubun belli bir değişkene ilişkin ölçümlerini karşılaştırmak, salt ortalama ölçüleriyle mümkün değildir.

ÖRN: Bir grubun değişkenliğine ilişkin ortalama 50 olsun. Bu vasat bir öğrencinin alacağı değeri göstermektedir. Ancak , gözlenen en yüksek ve düşük puan ile ilgili bir şeyler söylemek için değişim ölçüleri gereklidir. YA DA ortalaması aynı iki gruptan hangisinin gerçek anlamda daha başarılı olduğunu anlamak için değişkenlik ölçülerine ihtiyaç var.

(Büyüköztürk ve diğerleri, 2018)

Ranj

- Bir ölçümün ranjı, ölçümlerin en büyüğü ile en küçüğü arasındaki farktır (Ranj= En Büyük Ölçüm- En Küçük Ölçüm).
- Grubun homojen ya da heterojen bir dağılım gösterdiği hakkında bilgi verir
- Puanların sıralanmış olması gerekmez.
- Örnek: 78, 89, 56, 36, 48, 92, 59, 60 **>** Ranj: 92-36=56
- Kaba ve basit bir değişme ölçüsüdür. Verilerin yayılması hakkında fazla bilgi verici değildir (Baykul, 1999).
- Ortalamaları eşit ve n'leri aynı olan iki grubun karşılaştırılmasında kullanılır (Büyüköztürk ve diğerleri, 2013).

Standart Sapma (Evren için σ, Örneklem için S)

- Bir veri grubunda verilerin aritmetik ortalamadan ne kadar uzaklaştığının ölçüsüdür.
- Puanların ortalamadan olan farklarının, kareleri toplamının ortalamasının, kareköküne eşittir.
- Bir dağılımdaki ölçümlerin tümünü işleme kattığı için ileri matematikse hesaplar için uygun, güvenilir bir değişim ölçüsüdür.

! Örneklem standart sapması, evren standart sapmasından küçük çıkma eğilimindedir. Bu nedenle örneklem için standart sapma bulunurken formülde payda değerinden (grup sayısından) 1 çıkarılarak düzeltme işlemi uygulanır.

Standart Sapmanın Genel Formülü

$$S = \sqrt{\frac{\sum (X - \overline{X})^2}{n-1}}$$

S : Standart sapma

X_i : i'nci ölçüm değeri

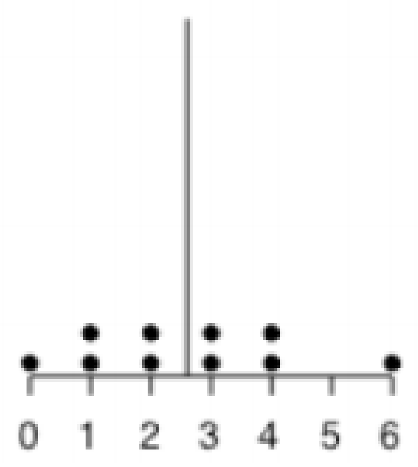
$\bar{X}$: n sayıda ölçümün ortalaması

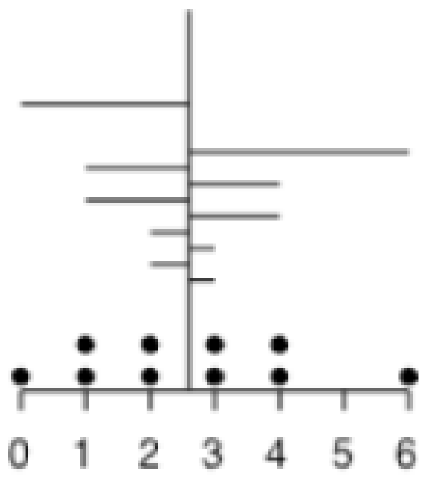
n : Ölçüm sayısı

$$S = \sqrt{\frac{n\sum X^2 - (\sum X)^2}{n(n-1)}}$$

!Matematiksel olarak: Her bir gözlemin ortalamadan uzaklıkları kareleri ortalamasının karekökü.

Standart Sapmanın Görsel Anlatımı

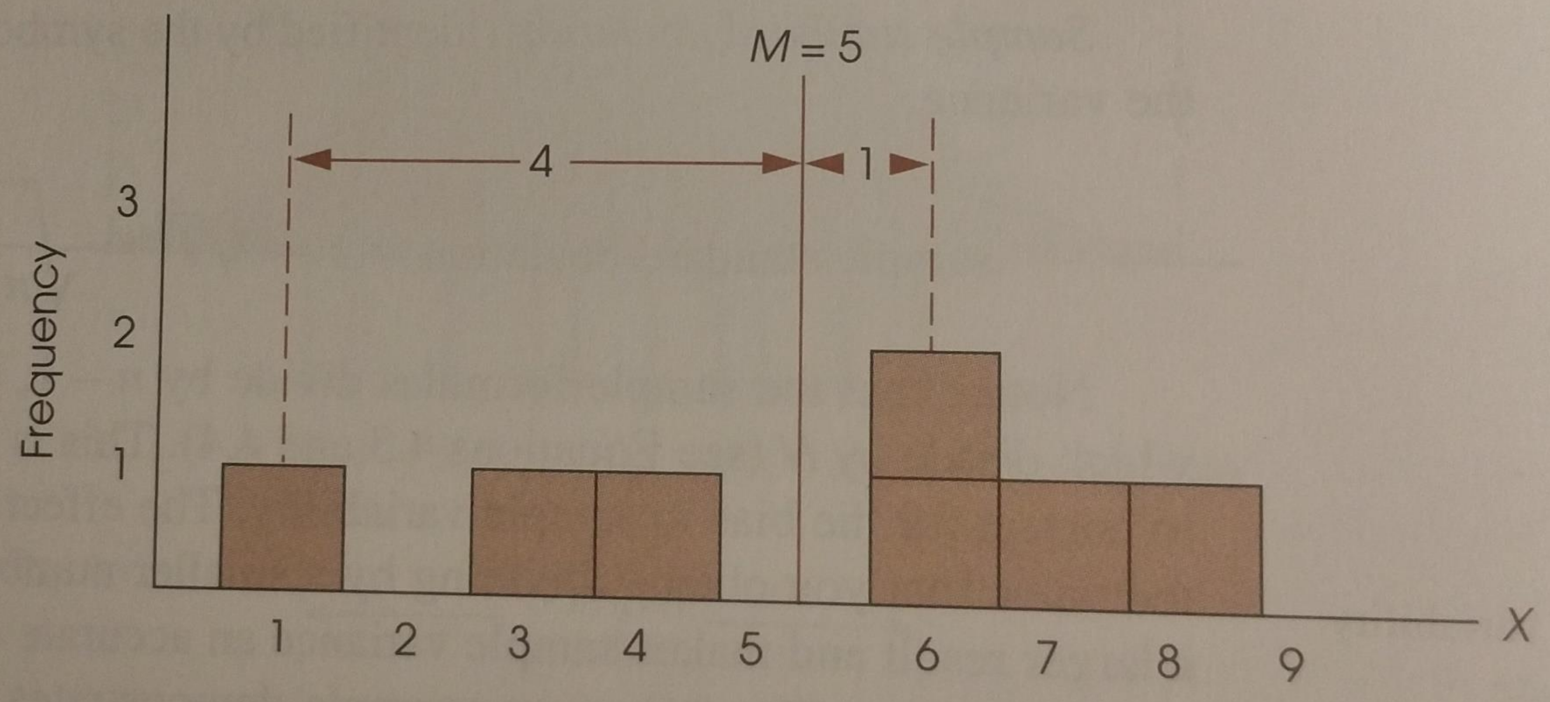




Değişim, uzaklık olarak betimlenmektedir. Puanlar, birlikte kümeleniyor mu yoksa geniş bir mesafeye dağılmış mı? Bir puan ortalamadan ne kadar uzakta? (Grawetter & Wallnau, 2013)

FIGURE 4.5

The frequency distribution histogram for a sample of $n = 7$ scores. The sample mean is $M = 5$. The smallest distance from the mean is 1 point, and the largest distance from the mean is 4 points. The standard distance (standard deviation) should be between 1 and 4 points, or about 2.5.



Hesaplanma Basamakları

1. Aritmetik ortalamanın hesaplanması
2. Her bir puanın aritmetik ortalamadan farkının hesaplanması
3. Bu farkların karelerinin hesaplanması ve toplanması
4. Farkların karelerinin toplamının öğrenci sayısının bir eksiğine bölünmesi

Hesap Makinesi ile Birlikte Hesaplayalım

Örnek Soru: Aşağıdaki 9 öğrencinin bir testten aldıkları puanlar verilmiştir. Bu puanlara ait standart sapmayı hesaplayınız.

Öğrenci	Test Puanı, X	$X_i - \bar{X}$	$(X_i - \bar{X})^2$
1	6	-4	16
2	7	-3	9
3	8	-2	4
4	9	-1	1
5	10	0	0
6	11	1	1
7	12	2	4
8	13	3	9
9	14	4	16
Toplam, Σ	90	000	60

Kareler Toplamı

$$S = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n-1}}$$

$$= \sqrt{\frac{60}{9-1}} = \sqrt{\frac{60}{8}} = \sqrt{7.5}$$

$$= 2.73861$$

$$= 2.74$$

Tan, 2016

ÖRNEK I
(Gruplandırılmamış Veriler İçin)

Tablo 3.4: Günlük Ders Çalışma Saatleri

X	$X-\bar{X}$	$(X-\bar{X})^2$	$(X)^2$
8	2	4	64
1	-5	25	1
5	-1	1	25
10	4	16	100
3	-3	9	9
9	3	9	81
$\sum = 36$	**0**	**64**	**280**

$$S= \sqrt{\frac{n\sum X^2-(\sum X)^2}{n(n-1)}} = \sqrt{\frac{6(280)-(36)^2}{6(6-1)}} = \sqrt{\frac{1680-1296}{30}}$$

$$S= \sqrt{12.8} =3.58$$

Aynı veriler için dağılımın varyansı, standart sapmanın karesi olup:

$S^2=(3.58)^2=12.8$'dir.

Gruplandırılmış Veriler için Standart Sapma

$$S = \sqrt{\frac{\sum f(X_0 - \overline{X})^2}{n-1}}$$

$$S = \sqrt{\frac{n\sum fX_0^2 - (\sum fX_0)^2}{n(n-1)}}$$

S : Standart sapma

X_0 : İlgili aralığın orta noktası

$\bar{X}$: İlgili aralığın ortalaması

n : Ölçüm sayısı

Gruplandırılmamış Ancak Tekrarlı Veriler için Standart Sapma

$$S = \sqrt{\frac{n\sum fX^2 - (\sum fX)^2}{n(n-1)}}$$

ÖRNEK II
(Gruplandırılmamış Ancak Tekrarlı Veriler için Standart Sapma)

$$S = \sqrt{\frac{\Sigma f(X-\bar{X})^2}{n-1}} = \sqrt{\frac{\Sigma fx^2}{n-1}}$$

Tablo 5-3. Matematik Testi Puanları

Puan (X)	(f)	$(X-\bar{X})$ (x)	$(X-\bar{X})^2$ (x^2)	$f(X-\bar{X})^2$ (fx^2)
9	1	+4	16	16
8	1	+3	9	9
7	3	+2	4	12
5	4	0	0	0
4	2	−1	1	2
3	2	−2	4	8
2	1	−3	9	9
1	1	−4	16	16
	$n = 15$			$\Sigma fx^2 = 72$

$\Sigma fx^2 = 72$ olduğundan,

$$S = \sqrt{\frac{72}{14}} = \sqrt{5.14} = 2.27$$

! Ortalamadan farkların gözlenme sayısı- kadar ağırlıklandırılması

(Arıcı, 1998)

ÖRNEK II (Gruplandırılmış Veriler için Standart Sapma)

Tablo 3.6: *Boy Uzunluklarına Ait Frekans Dağılımı*

Puan Aralığı	f	X_0	fX_0	$f(X_0)^2$	x'	fx'	$(x')^2$	$(x')^2$
150-154	1	152	152	23104	0	0	0	0
155-159	9	157	1413	221841	1	9	1	9
160-164	8	162	1296	209952	2	16	4	32
165-169	13	167	2171	362557	3	39	9	117
170-174	16	172	2752	473344	4	64	16	256
175-179	3	177	531	93987	5	15	25	75
$\sum =$	50		8315	1384785		143		489

$$S=a\sqrt{\frac{n\sum fx'^2-(\sum fx')^2}{n(n-1)}}=5\sqrt{\frac{50(489)-(143)^2}{50(50-1)}}$$

$$S=5\sqrt{\frac{4001}{2450}}=5\sqrt{1.63}=5\ (1.28)=6.39\text{'dir.}$$

$$S=\sqrt{\frac{n\sum fX_0{}^2-(\sum fX_0)^2}{n(n-1)}}=\sqrt{\frac{50(1384785)-(8315)^2}{50(50-1)}}$$

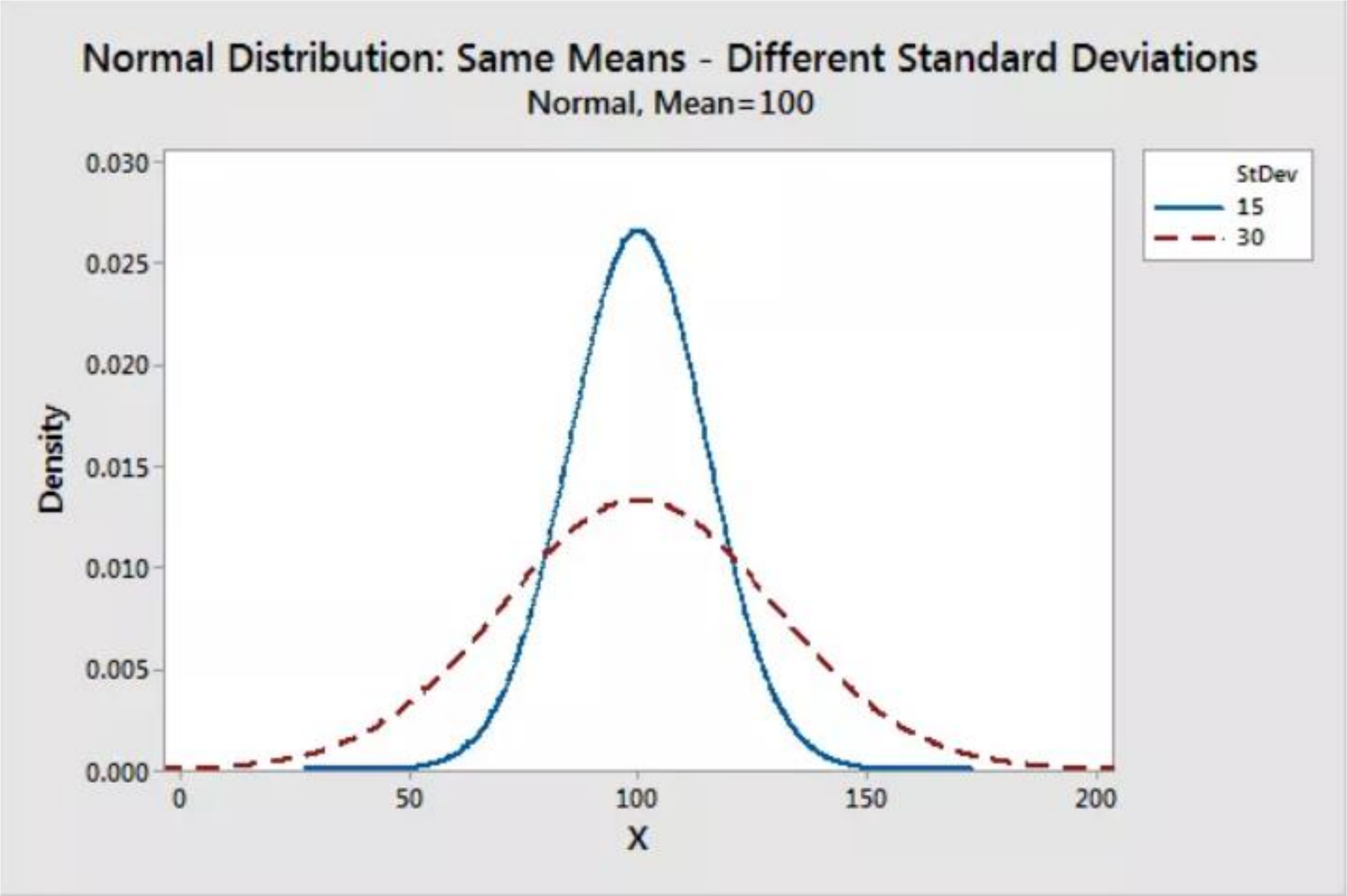
$$S=\sqrt{\frac{100025}{2450}}=\sqrt{40.83}=6.39\text{'dır.}$$

Değişkenlik Ölçüleri ve Alan İlişkisi...

- Veri setindeki ölçümlerin dağıldığı alan ile standart sapma arasında doğru bir orantı vardır:
  - Standart sapma büyüdükçe veri saçılır (ya da veri saçıldığı için standart sapma büyüktür), örneklem **heterojenleşir**, ayrıklaşır.
  - Standart sapma küçüldükçe veri daha az saçılır (ya da veri daha az saçıldığı için standart sapma küçüktür), örneklem **homojenleşir**, benzeşir.
- Bir veri setinin benzeşik olup olmadığına ya da ikiden çok veri setinden hangilerinin daha benzeşik-ayrık olduğunu belirlemede değişim katsayıları kullanılabilir.

(Büyüköztürk ve diğerleri, 2018)

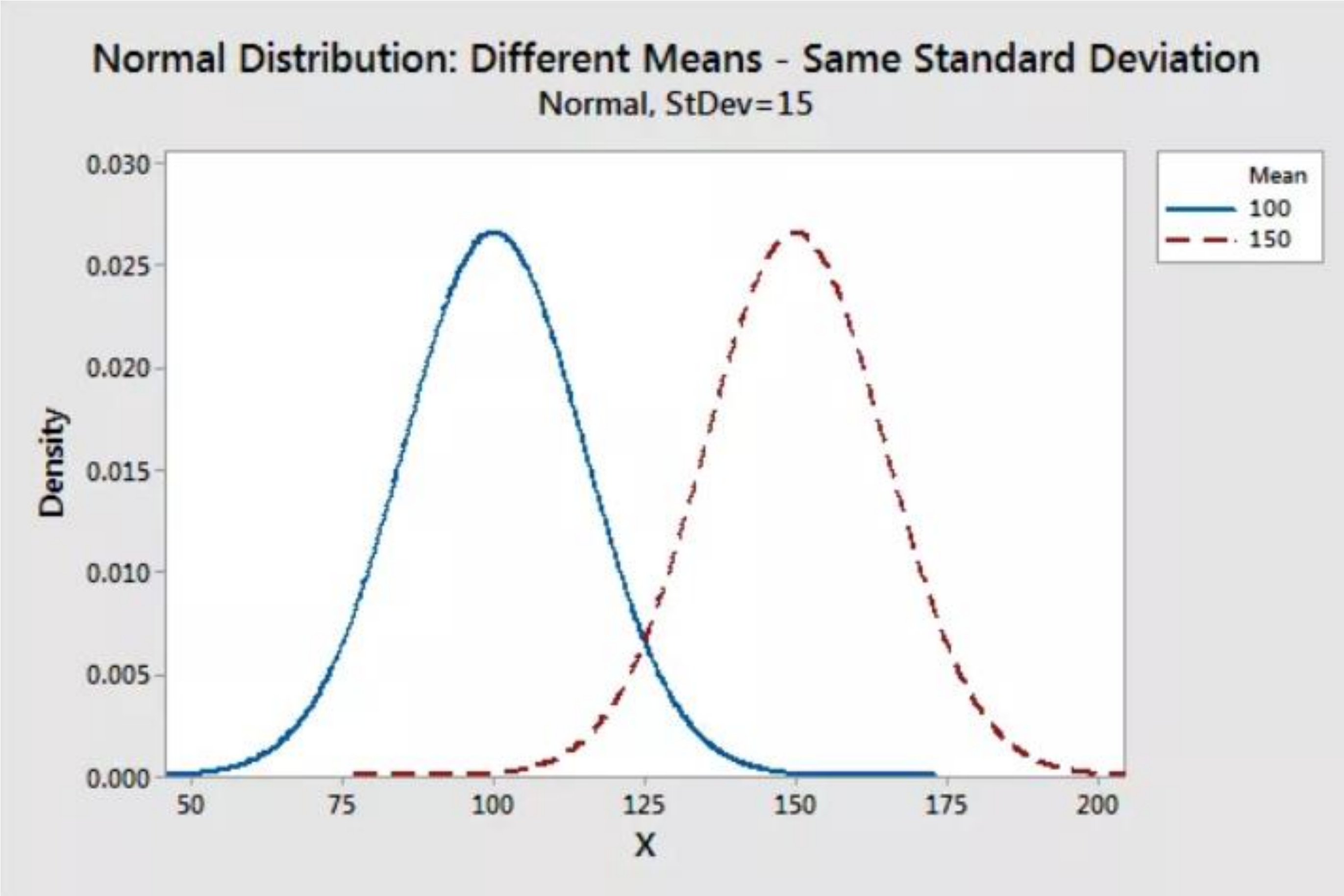
…Değişkenlik Ölçüleri ve Alan İlişkisi…



Eğriler, iki sınıftaki 14 yaş grubu kız çocukların boy uzunluklarına ait.

Ortalamaları 100 cm
standart sapmaları 15 cm ve 30 cm olan iki veri setine ait dağılım eğrileri.

…Değişkenlik Ölçüleri ve Alan İlişkisi



Eğriler, iki sınıftaki 14 yaş grubu kız çocukların boy uzunluklarına ait.

Ortalamaları 100 cm ve 150 cm; standart sapmaları 15 cm olan iki veri setine ait dağılım eğrileri.

Çeyrek Sapma

- Merkezi eğilim ölçüsü olarak ortalama yerine ortancanın kullanıldığı durumlarda değişkenlik ölçüsü olarak kullanılır.
- Ortancadan sapmaya ilişkin bilgi verir.
- Standart sapam gibi aşırı uç değerlerden etkilenmez.
- Çeyrek sapma üçüncü çeyrek ile birinci çeyrek arasındaki farkın yarısına eşittir.

$$\varphi = \frac{Y_{75} - Y_{25}}{2}$$

Tablo 3.2: *Öğrencilerin Türkçe Başarı Testi Puanları*

Puan Aralığı	f	X_0	fX_0	X_0^2	fX_0^2	x^1	fx^1	$(x^1)^2$	$f(x^1)^2$	t_f
90-98	2	94	188	8836	17672	9	18	81	162	100
81-89	6	85	510	7225	43350	8	48	64	384	98
72-80	9	76	684	5776	51984	7	63	49	441	92
63-71	12	67	804	4489	53868	6	72	36	432	83
54-62	17	58	986	3364	57188	5	85	25	425	71
45-53	13	49	637	2401	31213	4	52	16	208	54
36-44	15	40	600	1600	24000	3	45	9	135	41
27-35	14	31	434	961	13454	2	28	4	56	26
18-26	8	22	176	484	3872	1	8	1	8	12
09-17	4	13	52	169	676	0	0	0	0	4
$\sum =$			**5071**		**297277**		**419**		**2251**	

Örnek 3.10 Tablo 3.2'de verilen öğrencilerin Türkçe başarı testi puanları için çeyrek sapmayı bulunuz.

Çözüm. Çeyrek sapmayı bulmak için öncelikle Formül 3.6 kullanılarak 25. ve 75.yüzdelikleri hesaplayalım.

$$Y_{25}=26.5 + \frac{25-12}{14}\,9=34.86 \text{ ve } Y_{75}=62.5 + \frac{75-71}{12}\,9=65.5$$

Bulunan bu değerleri Formül 3.14'te yerine koyalım.

$$\varphi=\frac{Y_{75}-Y_{25}}{2}=\frac{(65.5)-(34.86)}{2}=15.32\text{'dir.}$$

VARYASYON KATSAYISI (DEĞİŞİM KATSAYISI)

Standart sapma dağılımın yaygınlığını gösteren bir ölçüdür.

- Ancak standart sapma ile dağılım hakkında çok fazla bir şey söylemek olanaksızdır.
- Örneğin; «B*ir dağılımın standart sapması 6 ise bu değer büyük müdür, yoksa küçük müdür?»* buna karar verebilmek için VARYASYON KATSAYISINI hesaplamak gerekir.
- Varyasyon katsayısı; standart sapmanın ortalamaya göre yüzde kaçlık bir değişim gösterdiğini belirtir.

$$V = \frac{S}{X} 100$$

Örnek: Kuruyemiş satan bir dükkanda bir haftalık sürede satılan leblebi, fıstık ve bademlerin ortalamaları ve standart sapmaları aşağıda verilmiştir. Buna göre kuruyemişleri değişkenlikleri açısından karşılaştırınız ve kuruyemişin değişkenliğinin daha fazla olduğunu belirtiniz.

	$\bar{x}$	s
Leblebi	30 kg.	5 kg.
Fıstık	40 kg.	4 kg.
Badem	10 kg.	3 kg.

$$C_{V_{leblebi}} = \frac{s}{\bar{X}} * 100 = \frac{5}{30} * 100 = 16{,}67 = \%16{,}67$$

$$C_{V_{fıstık}} = \frac{s}{\bar{X}} * 100 = \frac{4}{40} * 100 = 10 = \%10$$

$$C_{V_{BADEM}} = \frac{s}{\bar{X}} * 100 = \frac{3}{10} * 100 = 30 = \%30$$

Üç kuruyemişin değişkenlikleri karşılaştırıldığında en küçük standart sapma değeri bademde olmasına rağmen en büyük varyasyon katsayısına sahip olduğundan en fazla değişkenliğin bademde olduğu görülür. Aritmetik ortalamalar içerisinde standart sapma yüzdelerine bakıldığında en büyük yüzde bademdedir.

ÖRNEK

- Ortalaması 31.7 ve standart sapması 8.37 olan bir dağılımın varyasyon katsayısı,

$$V = (8.37 / 31.7) \times 100$$

$$= \%\ 26.4$$

- Bu dağılımdaki değerler ortalamaya göre %26.4'lük bir değişim göstermektedir.

Varyans (Evren için σ^2, Örneklem için S^2)

- Puanların dağılımı ile ilgili olarak bilgi vermektedir.
- Puanların ortalamadan sapmalarının karelerinin evren için N'e örneklem için (n-1)'e bölümüdür.
- Uygulamada, varyans yerine standart sapma kullanılmaktadır.
- Değer olarak varyans, standart sapmanın karesidir.

(Büyüköztürk ve diğerleri, 2018)

- Standart sapma, ortalamadan standart uzaklığı ölçerken varyans ortalamadan uzaklıklar karesinin ortalamasıdır.

(Grawetter & Wallnau, 2013)

Çarpıklık Katsayısı (Kayış Ölçüleri)

- Bir dağılımda ortalama ve ortanca ayrı noktalarda ise dağılım çarpıktır. Ancak bu çarpıklığı yorumlamaya yetmez, daha güvenilir değerlere ihtiyaç var.
- Çarpıklık katsayısının sıfırdan küçük olması dağılımın sola (negatif) çarpık olduğunu gösterir ve ortanca ortalamadan büyüktür. Çarpıklık katsayısının sıfırdan büyük olması ise dağılımın sağa (pozitif) çarpık olduğunu gösterir ve ortalama ortancadan büyüktür.

$$Çarp = \frac{3(\overline{X} - Ortanca)}{S}$$

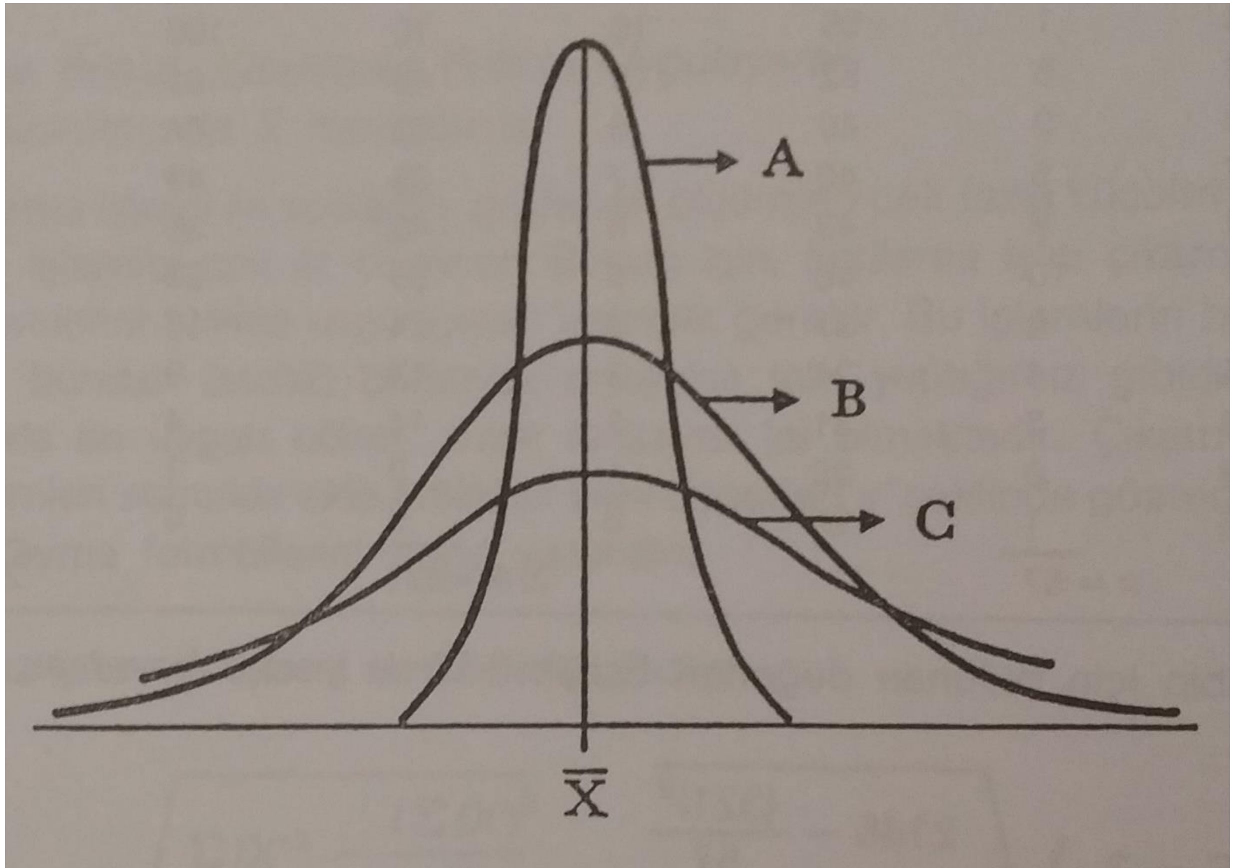
$$Çarp = \frac{\sum (X - \overline{X}) / n}{S^3}$$

Basıklık Katsayısı (Sivrilme Ölçüleri)

- Dağılımın genişliğini yorumlamada kullanılır.
- Basıklık katsayısı negatifse basık; pozitifse sivrilmiş dağılım gösterir.

$$Bas = \frac{\sum (X - \overline{X})^4 / n}{S^4} - 3$$

ŞEKİLDEKİ DAĞILIMLAR İÇİN NE SÖYLERSİNİZ?



Şekil 5-2. Ortalamaları Aynı, Standart Kaymaları Farklı Üç Dağılım

Üç dağılımın ortalamaları birbirine eşittir; ancak standart sapmaları çok farklıdır. Standart sapmaları yönünden yorumlarsak;

- A'nın ölçümleri dar alana dağılmış, standart sapma küçük.
- B'nin ölçüleri daha geniş alana dağılmış, standart sapma daha büyük
- C'nin ölçümleri en geniş alana dağılmış, standart sapması en büyük

Dağılımların basıklıkları yönünden yorumlarsak;

- A sivrilmş, B normal, C basık ya da yassı dağılım göstermektedir.

(Arıcı, 1998)

Bilgisayar Uygulaması için Bazı Terimler

- **Standard Deviation**: Standart Sapma
- **Variance**: Varyans
- **Range**: Ranj
- **Skewness**: Çarpıklık
- **Kurtosis**: Basıklık
- **Mean**: Ortalama
- **Median:** Ortanca
- **Frequency**: Frekans

KAYNAKLAR

Arıcı, H. (1998). *İstatistik: Yöntemler ve uygulama*. Kendi Yayını.

Baykul, Y. (1999). *İstatistik: Metodlar ve uygulamalar*. Ankara: Anı Yayıncılık.

Büyüköztürk, Ş., Çokluk, Ö. ve Köklü N. (2013). *Sosyal bilimler İçin istatistik*. Ankara: Pegem Akademi.

Gravetter, F. J., & Wallnau, L. B. (2013). *Statistics for the behavioral sciences*. Pacific Grove, CA: *Cole Publishing Company.*

Tan, Ş. (2016). *SPSS ve Excel Uygulamalı Temel İstatistik I.* Ankara: Pegem Akademi.