HACETTEPE ÜNİVERSİTESİ
İSTATİSTİK BÖLÜMÜ

İST156 İSTATİSTİĞE GİRİŞ II

UYGULAMA 3

Ar. Gör. Dr. Derya Turfan – Ar. Gör. Leyla Bakacak Karabenli

1) 12 hastaya ilişkin kolesterol değerleri ölçülmüştür. Kitle ortalaması için %95 güven aralığını hesaplayınız.

Kolesterol	178	254	185	219	205	182	310	191	245	229	245	240

Veri girişi için SPSS’te *Variable View* penceresinde *kolesterol* ismi ile değişken tanımlanmış ve *Data View*’de gözlem değerlerinin girişi yapılmıştır.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	kolesterol	Numeric	8	0		None	None	8	Right	Scale	Input

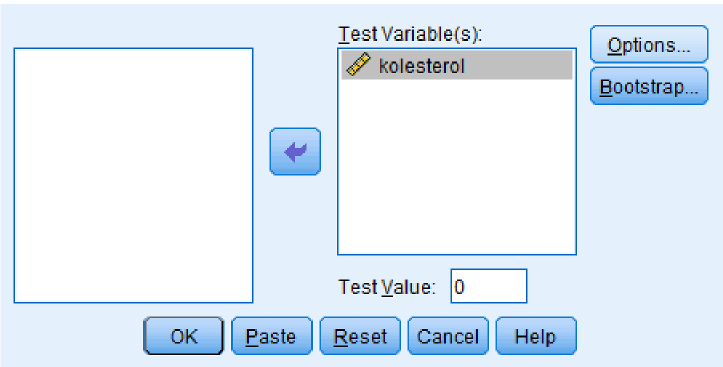
	kolesterol
1	178
2	254
3	185
4	219
5	205
6	182

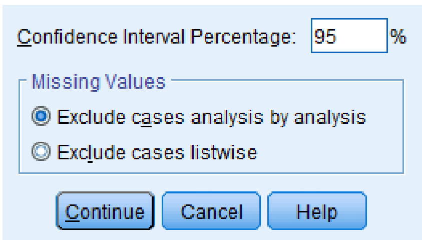
. . .

	kolesterol
7	310
8	191
9	245
10	229
11	245
12	240

Kitle varyansı (σ^2) bilinmediği ve n<30 olduğu için t dağılımından yararlanılır. SPSS’te kitle ortalamasına ait güven aralığının hesaplanması için aşağıdaki adımlar izlenir.

Analyze → Compare Means → One Sample T-Test





One-Sample Test

	Test Value = 0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
kolesterol	20.100	11	.000	223.583	**199.10**	**248.07**

Yorum: Hesaplanan güven aralığının kitle ortalamasını kapsama olasılığı %95’tir. Hastaların ortalama kolesterol değerleri [199.10, 248.07] aralığındadır.

Hazırlayanlar: Ar. Gör. Ceren Ünal, Ar. Gör. Leyla Bakacak Karabenli

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{178 + \cdots + 240}{12} = 223.58$$

$$S^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{(178-223.58)^2 + \cdots + (240-223.58)^2}{12-1} = 1484.811 \rightarrow S = \sqrt{1484.811} = 38.533$$

$$1 - \alpha = 0.95 \rightarrow \alpha = 0.05 \rightarrow t_{0.05/2;12-1} = t_{0.025;11} = 2.201$$

$$P\left(\bar{X} - t_{\alpha/2, n-1}\frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{\alpha/2, n-1}\frac{S}{\sqrt{n}}\right) = 1 - \alpha$$

$$P\left(223.58 - 2.201\frac{38.533}{\sqrt{12}} < \mu < 223.58 + 2.201\frac{38.533}{\sqrt{12}}\right) = 0.95$$

$$P(199.10 < \mu < 248.07) = 0.95$$

2) Aşağıda 15 kişiye ait yaşların dağılımı verilmiştir. Kitle varyansı için %95 güven aralığını hesaplayınız.

Yaş	28	46	48	49	32	36	50	51	52	55	56	57	58	59	47

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{28 + \cdots + 47}{15} = 48.27$$

$$S^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{(28-48.27)^2 + \cdots + (47-48.27)^2}{15-1} = 89.21$$

$n = 15$

$$1 - \alpha = 0.95 \rightarrow \alpha = 0.05 \rightarrow \chi^2_{0.05/2;15-1} = \chi^2_{0.025;14} = 26.11895$$

$$\chi^2_{1-(0.05/2);15-1} = \chi^2_{0.975;14} = 5.62873$$

$$P\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2;n-1}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\alpha/2;n-1}}\right) = 1 - \alpha$$

$$P\left(\frac{(15-1)89.21}{26.11895} < \sigma^2 < \frac{(15-1)89.21}{5.62873}\right) = 0.95$$

$$P(47.817 < \sigma^2 < 221.887) = 0.95$$

Yorum: Hesaplanan güven aralığının kitle varyansını kapsama olasılığı %95'tir. Yaş değişkenine ait varyans [47.817, 221.887] aralığındadır.

3) Bir bölgede guatr hastalığına rastlanma oranının tahmini için 250 kişi üzerinde yapılan bir araştırmada 156 kişide guatr hastalığına rastlandığı saptanmıştır. Kitle oranı için %95 ve %90 güven aralığını hesaplayınız.

$$p = \frac{156}{250} = 0.624$$

$1 - \alpha = 0.95 \rightarrow \alpha = 0.05 \rightarrow Z_{0.05/2} = Z_{0.025} = 1.96$

$1 - \alpha = 0.90 \rightarrow \alpha = 0.10 \rightarrow Z_{0.10/2} = Z_{0.05} = 1.645$

$$P\left(p - Z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} < P < p + Z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}\right) = 1 - \alpha$$

$$P\left(0.624 - 1.96\sqrt{\frac{0.624(1-0.624)}{250}} < P < 0.624 + 1.96\sqrt{\frac{0.624(1-0.624)}{250}}\right) = 0.95$$

$$P(0.564 < P < 0.684) = 0.95$$

Yorum: Hesaplanan güven aralığının kitle oranını kapsama olasılığı %95'tir. Guatr hastalığına yakalananların oranı [0.564, 0.684] aralığındadır.

$$P\left(0.624 - 1.645\sqrt{\frac{0.624(1-0.624)}{250}} < P < 0.624 + 1.645\sqrt{\frac{0.624(1-0.624)}{250}}\right) = 0.90$$

$$P(0.574 < P < 0.674) = 0.90$$

Yorum: Hesaplanan güven aralığının kitle oranını kapsama olasılığı %90'dır. Guatr hastalığına yakalananların oranı [0.574, 0.674] aralığındadır.

4) Bağımsız bir çevreci kuruluşun atmosferdeki karbondioksit oranına ilişkin hazırladığı bir rapora göre son 20 yıllık atmosferdeki karbondioksit ölçümlerinin ortalamasının 1.453 ppm olduğu belirtilmektedir. Daha önce incelenen uzun dönemli verilerden atmosferdeki karbondioksit miktarının varyansının 0.745 olduğu bilinmektedir. Atmosferdeki karbondioksit miktarı için %95 güven aralığını bulunuz.

Kitle varyansı (σ^2) bilindiği için Z dağılımından yararlanılır.

σ^2=0.745

n =20

$\bar{X}$=1.453

$1 - \alpha = 0.95 \rightarrow \alpha = 0.05 \rightarrow Z_{0.05/2} = Z_{0.025} = 1.96$

$$P\left(\bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

$$P\left(1.453 - 1.96\sqrt{\frac{0.745}{20}} < \mu < 1.453 + 1.96\sqrt{\frac{0.745}{20}}\right) = 0.95$$

$$P(1.075 < \mu < 1.831) = 0.95$$

Yorum: Hesaplanan güven aralığının kitle ortalamasını kapsama olasılığı %95'tir. Atmosferdeki karbondioksit miktarının ortalaması [1.075, 1.831] aralığındadır.

5) 25 basketbol oyuncusu üzerinde yapılan bir çalışmada basketbolcuların sıçrama ortalaması 46 cm, standart sapması 4 cm'dir. Federasyona bağlı tüm basketbolcuların sıçrama ortalaması için %95 güven aralığını bulunuz.

Kitle varyansı (σ^2) bilinmediği ve n<30 olduğu için t dağılımından yararlanılır.

$n = 25$

$S = 4$

$\bar{X} = 46$

$1 - \alpha = 0.95 \rightarrow \alpha = 0.05 \rightarrow t_{0.05/2;25-1} = t_{0.025;24} = 2.064$

$$P\left(\bar{X} - t_{\alpha/2,n-1}\frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}\right) = 1 - \alpha$$

$$P\left(46 - 2.064\frac{4}{\sqrt{25}} < \mu < 46 + 2.064\frac{4}{\sqrt{25}}\right) = 0.95$$

$$P(44.35 < \mu < 47.651) = 0.95$$

Yorum: Hesaplanan güven aralığının kitle ortalamasını kapsama olasılığı %95'tir. Federasyona bağlı tüm basketbolcuların sıçrama ortalaması [44.345, 47.651] aralığındadır.

6) Bir araştırmacı Akdeniz bölgesinde mevsimlik işçi olarak çalışanların ortalama ücretini tahmin etmek için rasgele seçtiği 81 birimlik örneklemde ortalama ücret 52TL ve standart sapma 3 TL olarak bulunmuştur. Mevsimlik işçilerin ücret ortalaması için %94 güven aralığını bulunuz.

Kitle varyansı (σ^2) bilinmediği ve n>30 için Z dağılımından yararlanılır.

n =81

S=3

$\bar{X}$=52

$1 - \alpha = 0.94 \rightarrow \alpha = 0.06 \rightarrow Z_{0.06/2} = Z_{0.03} = 1.88$

$$P\left(\bar{X} - Z_{\alpha/2}\frac{S}{\sqrt{n}} < \mu < \bar{X} + Z_{\alpha/2}\frac{S}{\sqrt{n}}\right) = 1 - \alpha$$

$$P\left(52 - 1.88\frac{3}{\sqrt{81}} < \mu < 52 + 1.88\frac{3}{81}\right) = 0.94$$

$$P(51.37 < \mu < 52.63) = 0.94$$

Yorum: Hesaplanan güven aralığının kitle ortalamasını kapsama olasılığı %94'tür. Mevsimlik işçilerin ücret ortalaması [51.37, 52.63] aralığındadır.

7) Bir bölgede sağlık taraması sonuçlarına göre 400 kişinin 200'ünde sarılık hastalığına rastlanmıştır. Kitle oranı için %90 güven aralığını bulunuz.

$p = \frac{x}{n} = \frac{200}{400} = 0.50$

$1 - \alpha = 0.90 \rightarrow \alpha = 0.1 \rightarrow Z_{0.1/2} = Z_{0.05} = 1.645$

$$P\left(p - Z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} < P < p + Z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}\right) = 1 - \alpha$$

$$P\left(0.5 - 1.645\sqrt{\frac{0.5(1-0.5)}{400}} < P < 0.5 + 1.645\sqrt{\frac{0.5(1-0.5)}{400}}\right) = 0.90$$

$$P(0.459 < P < 0.541) = 0.90$$

Yorum: Hesaplanan güven aralığının kitle oranını kapsama olasılığı %90'dır. Sarılık hastalığına yakalananların oranı [0.459, 0.541] aralığındadır.

8) 12 öğrencinin notlarının varyansı 4.8'dir. Kitle varyansının %90 güven aralığını bulunuz.

$S^2 = 4.8$

$n = 12$

$1 - \alpha = 0.90 \rightarrow \alpha = 0.10 \rightarrow \chi^2_{0.10/2;12-1} = \chi^2_{0.05;11} = 19.675$

$$\chi^2_{1-(0.10/2);12-1} = \chi^2_{0.95;11} = 4.575$$

$$P\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2;n-1}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\alpha/2;n-1}}\right) = 1 - \alpha$$

$$P\left(\frac{(12-1)4.8}{19.675} < \sigma^2 < \frac{(12-1)4.8}{4.575}\right) = 0.90$$

$$P(2.683 < \sigma^2 < 11.541) = 0.90$$

Yorum: Hesaplanan güven aralığının kitle varyansını kapsama olasılığı %90'dır. Öğrencilerin notlarına ait varyans [2.683, 11.541] aralığındadır.