n = 30'dan az ise Student's t Test 4/11

* sample sayısı az old zamanlarda normal dist app. çalışmıyor. Student's t test kullanılıyor bu durumlarda.

confidence interval → significance level → $\alpha$ : check !

(Z test, normal dist app)

Hypothesis Testing Using p-value

p-value is the lowest level of significance at which the observed value of a test statistic is significant

Reject $H_0$ if $p < \alpha$

Ex: A batch of 100 resistors have an average of 101.5 Ω

Standard deviation of the population is 5.

Test whether the population mean is 100Ω at a level of significance 0.05. $\alpha = 0.05$

$H_0: \mu = 100$

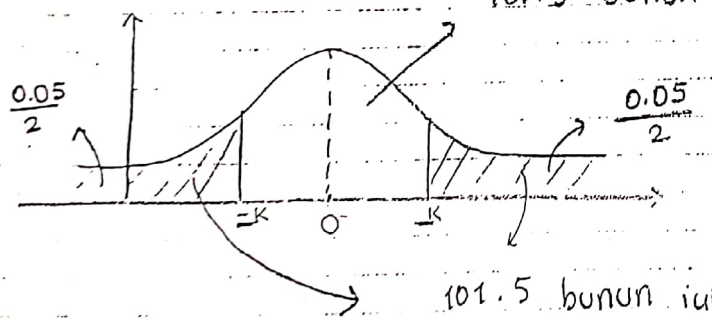
$H_1: \mu \neq 100$ } two-sided.

p → critical value → $\alpha$

$\alpha$ → critical value →

Test statistic : $\bar{X}$ sample mean

101.5 bunun içindeyse accept $H_0$




$F(-K) \rightarrow 0.025$ $F(K) \rightarrow 1 - \frac{\alpha}{2} = 0.975$

- bu değerleri bulmak istiyoruz

101.5 bunun içindeyse reject $H_0$

Reject $H_0$ if $\pm 1.96$

$\bar{X} < 100 - z_{0.025} \frac{\sigma}{\sqrt{n}}$ $\quad \bar{X} < 99.02$

$\bar{X} > 100 + z_{0.025} \frac{\sigma}{\sqrt{n}}$ $\quad \bar{X} > 100.98$

$\bar{X} = 101.5$ → so reject $H_0$.

$0.025 = \frac{\alpha}{2}$ $\quad \alpha$'i daha da küçültebilirim.

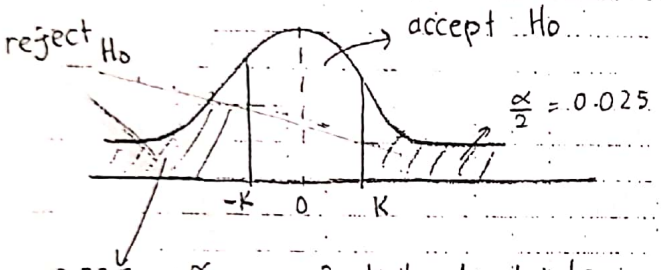
Confidence interval = $100\ (1-\alpha)\ \%$

$\alpha$ : 0.1 olsaydı reject $H_0$ daha da kesinleşecekti ama confidence interval azalacaktı.

$\alpha$ : 0.01 olsaydı belki de accept $H_0$ olacaktı.

$\mu = 100$, $\sigma = 5$ olan bir topluluktan 100 sample alıyoruz.

Sample mean = 101.5 olmuş $\alpha = 0.05$

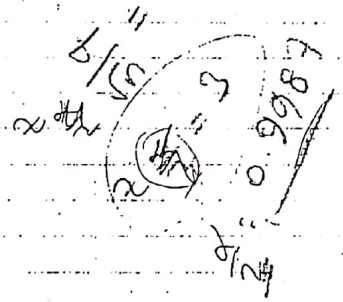


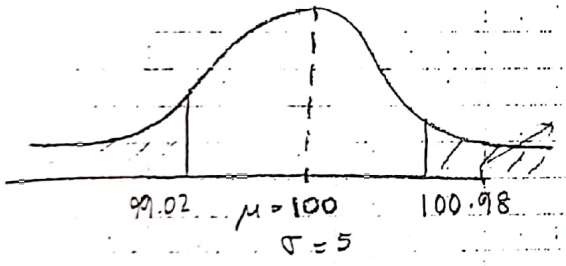

$0.025 = \frac{\alpha}{2}$ → z. dist. da böyle

$\mu = 0$ $\sigma^2 = 1$

K için = $1 - \frac{\alpha}{2}$

-K için = $\frac{\alpha}{2}$




101.5 => reject $H_0$

$\alpha$'ı küçültürsek sonuç değişmez 101.5 için hesaplanan değerin 2 katına kadar azaltırsak $\alpha$'yı sonuç değişmiyor hâlâ reject $H_0$.

101.5 için K değeri → 0.0026

$\alpha$ : 0.05 } reject $H_0$

$\alpha$ : 0.0026

hâlâ reject edebileceğimiz en küçük $\alpha$ değeri

en yüksek Confidence Interval

$n > 30$ but population variance is NOT known:

sample variance: $s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$ , $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$

$n-1$ helps to have an unbiased estimation of sample variance (instead of $n$)

z dist için population variance'a ihtiyaç var. (sample variance'dan estimation ediyoruz) ?

## Student's t distribution and t-test

The student t-dist. with $\nu$ degree of freedom:

$$f(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{t^2}{\nu}\right)^{\frac{-\nu+1}{2}} , \quad -\infty \leq t \leq \infty$$

$T = \frac{Z}{\sqrt{Y/\nu}}$ → normal Gaussian $(0,1)$

↳ chi-square distribution

$\nu \to \infty$ , $T \to$ Gaussian

If $n \geq 30$, t-dist gives similar values to z-dist.

z-test and t-test provide similar results.

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \qquad z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

↓ either $\sigma$ is not known or $n < 30$

example:

If $n = 30$, the critical values for the student's t-dist. are 2.045 for $\alpha = 0.25$ and 2.756 for $\alpha = 0.005$ which are close to the corresponding values from the z-dist. as 1.96 for $\alpha = 0.025$ and 2.576 for $\alpha = 0.005$.

$n\uparrow$ artar bu yakınlık, $n\downarrow$, $\downarrow$ yakınlık

Ex: the safety level : 200

175 190 215 198 184

207 210 193 196 180

Test at $\alpha = 0.01$ level (sample sayısı ↓ tolerans ↑)

$H_0 : \mu > 200$

$H_1 : \mu \leq 200$

one tailed

$n = 10$

$\nu$, degree of freedom : $10 - 1 = 9$ $(n-1)$

$(1-\alpha)\ 100\ \%$ → conf. interval

$(1 - 0.01)\ 100 = 99\ \%$

$T \leq -2.821$ gives $\alpha = 0.01$

$\bar{x} = 194.8$, $s^2 = 172.66$ → sample variance

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{194.8 - 200}{\sqrt{172.66/10}} = -1.25$$

sample standard deviation

computed

(t değeri T değerinden büyük çıktı.)

$t > T$

$\alpha$ need to be satisfied with $\alpha$ given in the test.

critical value for $\alpha$ given in the test

-2.821

We've failed to reject $H_0$. Samples support mean value of bacteria is within the safety level.