4.4 İki Kitle Ortalaması Arasındaki Farkın Önem Kontrolü

İki kitle ortalaması arasındaki karşılaştırma, uygulamada çok karşılaşılan problemdir. Örneğin eşit ısı, ışık, nem şartları altında iki tohum cinsinin verim farkı ile ilgili olarak iki bağımsız örneklemden bulunan sonuçlardan yararlanılarak kitle ortalaması arasındaki fark hakkında bilgi edinilebilir.

4.4.1 Kitle Varyansı σ_1^2 ve σ_2^2 Biliniyor

Hipotez Testi

1) Hipotez kurulur .

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 < \mu_2,\ \mu_1 > \mu_2,\ \mu_1 \neq \mu_2$

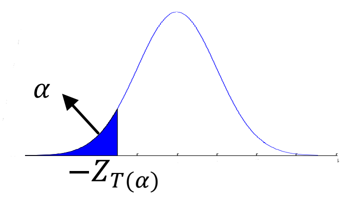
2) Test istatistiği hesaplanır.

$$Z_h = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)}}$$

3) Kritik bölgeye göre hipotez red edilir ya da red edilemez.

$H_0: \mu_1 = \mu_2$

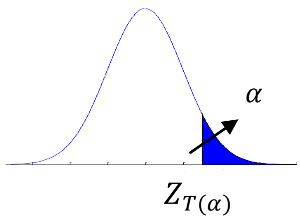
$H_1: \mu_1 < \mu_2$



$Z_H < -Z_{T(\alpha)}$ ise H_0 red edilir

$H_0: \mu_1 = \mu_2$

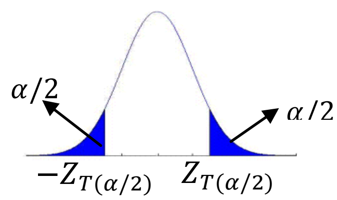
$H_1: \mu_1 > \mu_2$



$Z_H > Z_{T(\alpha)}$ ise H_0 red edilir

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2$



$Z_H < -Z_{T(\alpha/2)}$ ya da $Z_H > Z_{T(\alpha/2)}$ ise H_0 red edilir

Güven Aralığı

$$\bar{X}_1 - \bar{X}_2 \sim N\left((\mu_1 - \mu_2), \left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)\right)$$

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)}} \sim N(0,1)$$

$$P\left(-Z_{T(\alpha/2)} \le Z \le Z_{T(\alpha/2)}\right) = 1-\alpha$$

$$P\left((\bar{X}_1-\bar{X}_2) - Z_{T(\alpha/2)}\sqrt{\left(\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)} \le (\mu_1-\mu_2) \le (\bar{X}_1-\bar{X}_2) + Z_{T(\alpha/2)}\sqrt{\left(\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)}\right) = 1-\alpha$$

$1-\alpha$ güven düzeyinde $\mu_1-\mu_2$ için güven aralığı

Örnek 4.4. A ilacının ortalama etki süresinin B ilacının ortalama etki süresinden daha büyük olduğu öne sürülmektedir. A ilacı verilen hastaların etki süresine göre dağılımı normal ve varyansı 30 olarak bilinmektedir. B ilacı verilen hastaların etki süresine göre dağılımı da normal ve varyansı 36 olarak bilinmektedir. Rasgele seçilen 5 hastaya A ilacı verilmiş ve ortalama etki süresi 51 olarak bulunmuş ve yine rasgele seçilen 7 hastaya Bilacı verilmiş ve ortalama etki süresi 39.57 olarak bulunmuştur. Buna göre % 95 güven düzeyinde iddiayı test ediniz, A ve B ilaçlarının ortalama etki süreleri arasındaki fark için % 95 lik güven sınırlarını belirleyiniz.

A ilacına ilişkin	$n_1 = 5$	$\bar{X}_1 = 51$	$\sigma_1^2 = 30$
B ilacına ilişkin	$n_2 = 7$	$\bar{X}_2 = 39.57$	$\sigma_2^2 = 36$

1) $H_0: \mu_1 = \mu_2 \ (\mu_1-\mu_2 = 0) \longrightarrow$ A ve B ilacının ortalama etki süreleri arasında fark yoktur

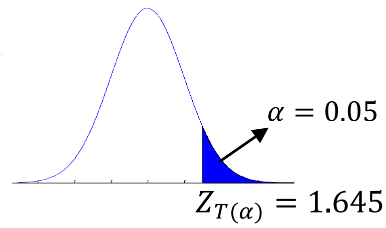
$H_1: \mu_1 > \mu_2 \ (\mu_1-\mu_2 > 0) \longrightarrow$ A ilacının ortalama etki süresi B ilacının ortalama etki süresinden fazladır

2)

$$Z_h = \frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\left(\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)}} = \frac{51-39.57}{\sqrt{\frac{30}{5}+\frac{36}{7}}} = \frac{11.43}{\sqrt{11.14}} = 3.42$$

3) $\alpha = 0.05$

$H_1: \mu_1 > \mu_2$



$Z_H > Z_{T(\alpha)}$ olduğundan H_0 red edilir

Yorum: A ilacının ortalama etki süresi B ilacının ortalama etki süresinden fazladır denilebilir.

Güven aralığı,

$$P\left((\bar{X}_1-\bar{X}_2) - Z_{T(\alpha/2)}\sqrt{\left(\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)} \le (\mu_1-\mu_2) \le (\bar{X}_1-\bar{X}_2) + Z_{T(\alpha/2)}\sqrt{\left(\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)}\right) = 1-\alpha$$

$$(\bar{X}_1-\bar{X}_2) \pm Z_{T(\alpha/2)}\sqrt{\left(\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)} = (51-39.57) \pm 1.96\sqrt{11.14} \begin{cases} \nearrow 4.89 \\ \searrow 17.97 \end{cases}$$

$\mu_1-\mu_2$: [4.89, 17.97] $\rightarrow$ Bu aralığın $\mu_1-\mu_2$'yü içeren aralıklardan biri olması olasılığı %95'dir.

4.4.2 Kitle Varyansı σ_1^2 ve σ_2^2 Bilinmiyor

a) $\boldsymbol{\sigma_1^2=\sigma_2^2=\sigma^2}$ (Birbirlerine eşit)

1) Hipotez kurulur .

$H_0: \mu_1=\mu_2$

$H_1: \mu_1<\mu_2,\ \mu_1>\mu_2,\ \mu_1\neq\mu_2$

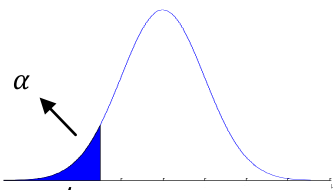
2) Test istatistiği hesaplanır.

$$t_h=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t_{(n_1+n_2-2)}, S_p=\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$$

3) Kritik bölgeye göre hipotez red edilir ya da red edilemez.

$H_0: \mu_1=\mu_2$

$H_1: \mu_1<\mu_2$



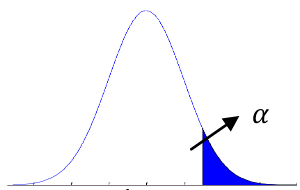
$-t_{T(\alpha,n_1+n_2-2\ )}$

$t_H<-t_{T(\alpha,n_1+n_2-2\ )}$

ise H_0 red

$H_0: \mu_1=\mu_2$

$H_1: \mu_1>\mu_2$



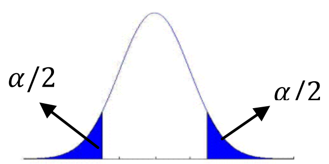
$t_{T(\alpha,n_1+n_2-2\ )}$

$t_H>t_{T(\alpha,n_1+n_2-2\ )}$

ise H_0 red

$H_0: \mu_1=\mu_2$

$H_1: \mu_1\neq\mu_2$



$-t_{T\left(\frac{\alpha}{2},n_1+n_2-2\ \right)}$ $t_{T\left(\frac{\alpha}{2},n_1+n_2-2\ \right)}$

$t_H<-t_{T\left(\frac{\alpha}{2},n_1+n_2-2\ \right)}$ ya da $t_H>t_{T\left(\frac{\alpha}{2},n_1+n_2-2\ \right)}$ ise H_0 red

Güven Aralığı

$$t=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t_{(n_1+n_2-2)}, S_p=\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$$

$$P\left((\bar{X}_1-\bar{X}_2)-t_{T\left(\frac{\alpha}{2},n_1+n_2-2\right)}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\le(\mu_1-\mu_2)\le(\bar{X}_1-\bar{X}_2)+t_{T\left(\frac{\alpha}{2},n_1+n_2-2\right)}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)=1-\alpha$$

$1-\alpha$ güven düzeyinde $\mu_1-\mu_2$ için güven aralığı

b) $\boldsymbol{\sigma_1^2 \neq \sigma_2^2}$ (Birbirinden farklı)

Hipotez Testi

1) Hipotez kurulur .

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 < \mu_2,\ \mu_1 > \mu_2,\ \mu_1 \neq \mu_2$

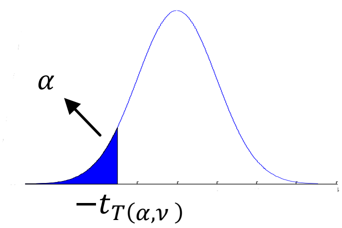
2) Test istatistiği hesaplanır.

$$t_h=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}\sim t_\nu,\ \upsilon=\frac{\left(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}\right)^2}{\left(\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1+1}+\frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2+1}\right)}-2$$

3) Kritik bölgeye göre hipotez red edilir ya da red edilemez.

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 < \mu_2$

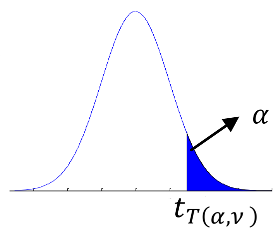


$t_H < -t_{T(\alpha,\nu)}$

ise H_0 red

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 > \mu_2$

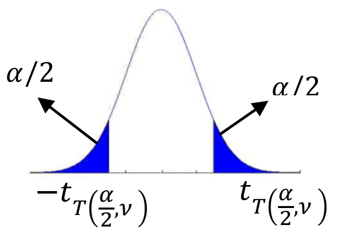


$t_H > t_{T(\alpha,\nu)}$

ise H_0 red

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2$



$t_H < -t_{T\left(\frac{\alpha}{2},\nu\right)}$ ya da

$t_H > t_{T\left(\frac{\alpha}{2},\nu\right)}$ ise H_0 red

Güven Aralığı

$$t=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}\sim t_\nu,\ \upsilon=\frac{\left(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}\right)^2}{\left(\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1+1}+\frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2+1}\right)}-2$$

$$P\left((\bar{X}_1-\bar{X}_2)-t_{T\left(\frac{\alpha}{2},\nu\right)}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}\le(\mu_1-\mu_2)\le(\bar{X}_1-\bar{X}_2)+t_{T\left(\frac{\alpha}{2},\nu\right)}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}\right)=1-\alpha$$

$1-\alpha$ güven düzeyinde $\mu_1-\mu_2$ için güven aralığı

Örnek 4.5. Yapılan bir çalışmada 12 karditli hastanın sedimantasyon hızı ortalaması $\bar{X}_1 = 87.29\,mm$ ve $S_1 = 18.4\,mm$, 15 romatizmalı hastanın sedimantasyon hızı ortalaması $\bar{X}_2 = 69.03\,mm$ ve $S_2 = 21.4\,mm$ olarak hesaplanmıştır. Karditli ve romatizmalı hastaların kitle ortalamaları arasında fark olup olmadığını %5 anlamlılık düzeyinde belirleyiniz. İki kitle ortalaması farkı için %95' lik güven sınırlarını bulunuz.

Varyanslar bilinmiyor. Buna göre,

a) İlk olarak σ_1^2=$\sigma_2^2=\sigma^2$ birbirlerine eşit kabul edilsin,

1) $H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2$

2)

$n_1 = 12$	$\bar{X}_1 = 87.29$	$S_1 = 18.4$
$n_2 = 15$	$\bar{X}_2 = 69.03$	$S_2 = 21.4$

$$S_p = \sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$$

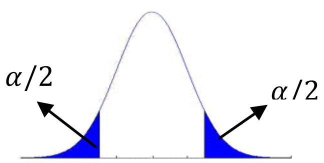
$$= \sqrt{\frac{(12-1)(18.4)^2+(15-1)(21.4)^2}{12+15-2}} = \sqrt{\frac{10135.6}{25}} = \sqrt{405.424} = 20.14$$

$$t_h = \frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{87.29-69.03}{20.14\sqrt{\frac{1}{12}+\frac{1}{15}}} = 2.34$$

3) Kritik bölgeye göre hipotez red edilir ya da red edilemez.

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2$



-2.060 $\quad t_{T(0.025,25)} = 2.060$

$t_H = 2.34 > t_{T\left(\frac{\alpha}{2}, n_1+n_2-2\right)}$ olduğundan H_0 red edilir. Yani, iki kitle ortalaması arasında fark vardır.

Güven aralığı,

$$P\left((\bar{X}_1 - \bar{X}_2) - t_{T\left(\frac{\alpha}{2}, n_1+n_2-2\right)} S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \le (\mu_1 - \mu_2) \le (\bar{X}_1 - \bar{X}_2) + t_{T\left(\frac{\alpha}{2}, n_1+n_2-2\right)} S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right) = 1 - \alpha$$

$$(\bar{X}_1 - \bar{X}_2) \pm t_{T\left(\frac{\alpha}{2}, n_1+n_2-2\right)} S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

$$(87.29 - 69.03) \pm 2.060 \times 20.14 \times \sqrt{\frac{1}{12}+\frac{1}{15}} \; \begin{cases} 3.52 \\ 33.00 \end{cases}$$

$\mu_1 - \mu_2 : [3.52, 33]$ → Bu aralığın $\mu_1 - \mu_2$'yü içeren aralıklardan biri olması olasılığı %95'dir.

b) $\sigma_1^2 \neq \sigma_2^2$ birbirinden farklı kabul edilsin

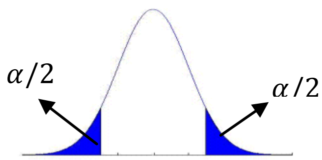
1) $H_0: \mu_1 = \mu_2$
$H_1: \mu_1 \neq \mu_2$

2)Test istatistiği hesaplanır.

$$t_h = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}} = \frac{87.29 - 69.03}{\sqrt{\frac{18.4^2}{12}+\frac{21.4^2}{15}}} = \frac{18.26}{7.66} = 2.38$$

$$\upsilon = \frac{\left(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}\right)^2}{\left(\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1+1}+\frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2+1}\right)} - 2 = \frac{\left(\frac{18.4^2}{12}+\frac{21.4^2}{15}\right)^2}{\frac{\left(18.4/12\right)^2}{13}+\frac{\left(21.4^2/15\right)^2}{15}} - 2 = 26.88 \cong 27$$

3) $H_0: \mu_1 = \mu_2$
$H_1: \mu_1 \neq \mu_2$



$-t_{T(0.025,27)} = -2.052$ $\quad t_{T(0.025,27)} = 2.052$

$t_H = 2.38 > t_{T(0.025,27)} = 2.052$ olduğundan H_0 red edilir. Yani, iki kitle ortalaması arasında fark vardır.

Güven Aralığı

$$P\left((\bar{X}_1-\bar{X}_2)-t_{T\left(\frac{\alpha}{2},\nu\right)}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}\le(\mu_1-\mu_2)\le(\bar{X}_1-\bar{X}_2)+t_{T\left(\frac{\alpha}{2},\nu\right)}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}\right)=1-\alpha$$

$$(\bar{X}_1-\bar{X}_2)\pm t_{T\left(\frac{\alpha}{2},\nu\right)}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}=(18.26)\pm 2.052\sqrt{\frac{18.4^2}{12}+\frac{21.4^2}{15}}$$

2.54

33.99

$\mu_1-\mu_2$: $[2.54, 33.99]$ $\longrightarrow$ Bu aralığın $\mu_1-\mu_2$' yü içeren aralıklardan biri olması olasılığı %95'dir.