4.6. Kitle Yüzdesinin Önem Kontrolü ve Güven Aralığı

Hipotez Testi(Önem Kontrolü)

1) Hipotez kurulur .

$H_0: p = p_0$

$H_1: p < p_0,\ p > p_0,\ p \neq p_0$

2) Test istatistiği hesaplanır.

$X_1, X_2, \dots, X_n \sim b(n,p)$

$X \sim N(np_0, np_0(1-p_0))$

$$Z_H = \frac{\left(X \mp \frac{1}{2}\right) - np_0}{\sqrt{np_0(1-p_0)}}$$

Süreklilik düzeltmesi

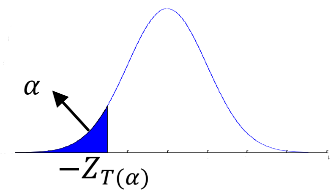
$$X > np_0 \Rightarrow X - \frac{1}{2}$$

$$X < np_0 \Rightarrow X + \frac{1}{2}$$

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

$H_0: p = p_0$

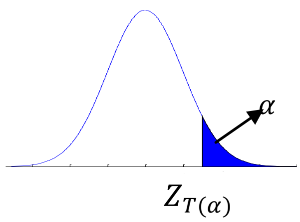
$H_1: p < p_0$



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

$H_0: p = p_0$

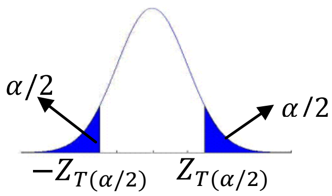
$H_1: p > p_0$



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

$H_0: p = p_0$

$H_1: p \neq p_0$



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

Güven Aralığı

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

$1-\alpha$ güven düzeyinde p için güven aralığı

$$\hat{p} = \frac{X}{n} \quad , \quad \hat{q} = 1 - \hat{p}$$

Örnek 4.8. Belli bir şehirde yaşayanlardan 20 yaşın üzerindeki nüfusun %30 dan fazlasının sigara içtiği iddia edilmektedir. İlgili şehirdeki 20 yaşın üzerinde olanlardan rasgele olarak 100 kişi seçilmiş ve bunların 45'inin sigara içtiği gözlenmiştir. $\alpha = 0.05$ anlamlılık düzeyinde iddiayı test ediniz.

1) $H_0: p = 0.30$

$H_1: p > 0.30$

2) $X = 45 > np_0 = 100(0.30) = 30 \ \Rightarrow X - \frac{1}{2}$

$$Z_H = \frac{\left(X + \frac{1}{2}\right) - np_0}{\sqrt{np_0(1 - p_0)}} = \frac{\left(45 - \frac{1}{2}\right) - 100(0.30)}{\sqrt{100(0.30)(0.70)}} = \frac{44.5 - 30}{\sqrt{21}} = 3.16$$

3) $Z_{T(\alpha=0.05)} = 1.645$

$Z_H = 3.16 > Z_{T(\alpha=0.05)} = 1.645$ olduğundan H_0 red edilir.

Yorum: Bir şehirde yaşayan 20 yaş üstü nüfusun %30'dan fazlasının sigara içtiği %95 güvenle söylenebilir.

Güven Aralığı

$\hat{p} = \frac{X}{n} = \frac{45}{100} = 0.45$, $\hat{q} = 1 - \hat{p} = 1 - 0.45 = 0.55$, $Z_{T(\alpha/2=0.05/2)}$=$Z_{T(0.025)} = 1.96$

$$\hat{p} \mp Z_{T(\alpha/2)}\sqrt{\frac{\hat{p}\hat{q}}{n}} = 0.45 \mp 1.96\sqrt{\frac{(0.45)(0.55)}{100}} \begin{cases} 0.353 \\ 0.548 \end{cases}$$

p: $[0.353;\ 0.548]$ Bu aralığın p'yi içeren aralıklardan biri olması olasılığı %95'dir.

Örnek 4.9. Bir beslenme uzmanı okul öncesi çocuklarda protein eksikliğinin 0.75 den az olduğunu düşünüyor. Yapılan araştırmada seçilen 300 okul öncesi çocuktan 206'sında protein eksikliği saptanıyor. Beslenme uzmanının tanısını %99 güvenirlikle test ediniz. Kitle yüzdesinin %99 güvenirlilikle güven sınırlarını oluşturunuz.

$\alpha = 0.01,\ n = 300,\ X = 206,\ p_0 = 0.75$

1) $H_0: p = 0.75$ (Okul öncesi çocuklarda protein eksikliği %75'dir)

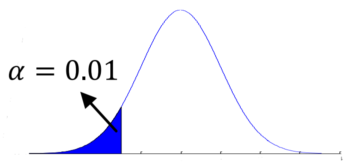
$H_1: p < 0.75$ (Okul öncesi çocuklarda protein eksikliği %75 den azdır.)

2) $X = 206 < np_0 = 300(0.75) = 225 \ \Rightarrow X + \frac{1}{2}$

$$Z_H = \frac{\left(X + \frac{1}{2}\right) - np_0}{\sqrt{np_0(1 - p_0)}} = \frac{\left(206 + \frac{1}{2}\right) - 300(0.75)}{\sqrt{300(0.75)(0.25)}} = \frac{206.5 - 225}{\sqrt{56.25}} = -2.47$$

3) $H_0: p = 0.75$

$H_1: p < 0.75$



$-Z_{T(\alpha=0.01)} = -2.33$

$Z_H < -Z_{T(\alpha=0.01)}$ olduğundan H_0 red edilir.

Yorum: Okul öncesi çocuklarda protein eksikliğinin %75'den daha az olduğu %99 güvenle söylenebilir.

Güven Aralığı

$$P\left(\hat{p} - Z_{T(\alpha/2)}\sqrt{\frac{\hat{p}\hat{q}}{n}} \leq p \leq \hat{p} + Z_{T(\alpha/2)}\sqrt{\frac{\hat{p}\hat{q}}{n}}\right) = 0.99$$

$\hat{p} = \frac{X}{n} = \frac{206}{300} = 0.69$, $\hat{q} = 1 - \hat{p} = 1 - 0.69 = 0.31$, $Z_{T(\alpha/2=0.01/2)}=Z_{T(0.005)} = 2.57$

$$\hat{p} \mp Z_{T(\alpha/2)}\sqrt{\frac{\hat{p}\hat{q}}{n}} = 0.69 \mp 2.57\sqrt{\frac{(0.69)(0.31)}{300}} \quad \begin{matrix} \nearrow 0.621 \\ \searrow 0.759 \end{matrix}$$

$p: [0.621;\ 0.759]$ Bu aralığın p'yi içeren aralıklardan biri olması olasılığı %99'dur.

4.7. İki Kitle Yüzdesi Arasındaki Farkın Önem Kontrolü ve Güven Aralığı

Hipotez Testi

1) Hipotez kurulur .

$H_0: p_1 = p_2$

$H_1: p_1 < p_2,\ p_1 > p_2,\ p_1 \neq p_2$

2) Test istatistiği hesaplanır.

$$\hat{p}_1 = \frac{X_1}{n_1}, \qquad \hat{p}_2 = \frac{X_2}{n_2} \qquad \hat{p} = \frac{X_1+X_2}{n_1+n_2}$$

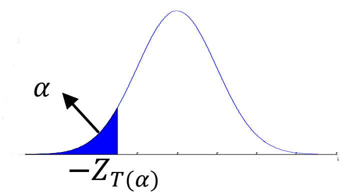
$$Z_h = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{S_{\hat{p}}}$$

$$S_{\hat{p}} = \sqrt{\hat{p}(1-\hat{p})\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}$$

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

$H_0: p_1 = p_2$

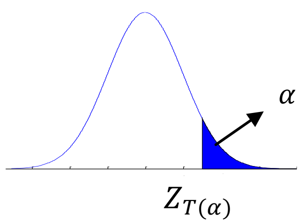
$H_1: p_1 < p_2$



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

$H_0: p_1 = p_2$

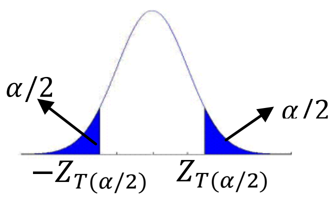
$H_1: p_1 > p_2$



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

$H_0: p_1 = p_2$

$H_1: p_1 \neq p_2$



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

Güven Aralığı

$$P\left((\hat{p}_1 - \hat{p}_2) - Z_{T(\alpha/2)}S_{\hat{p}} \leq (p_1 - p_2) \leq (\hat{p}_1 - \hat{p}_2) - Z_{T(\alpha/2)}S_{\hat{p}}\right) = 1 - \alpha$$

$1 - \alpha$ güven düzeyinde $p_1 - p_2$ için güven aralığı

Örnek 4.10. İlaçla tedavinin ameliyatla tedaviye göre daha az etkili olduğu öne sürülmektedir. Rasgele olarak seçilen 100 hasta ilaçla ve 120 hasta da ameliyatla tedavi edilmişlerdir. İlaçla tedavi edilenlerin 40'ı ve ameliyatla tedavi edilenlerin de 80'i iyileşmiştir. $\alpha = 0.10$ anlamlılık düzeyinde öne sürülen görüşe katılır mısınız? %90 güven düzeyinde $(p_1 - p_2)$ parametresi için güven aralığını oluşturunuz.

1) $H_0: p_1 = p_2$

$H_1: p_1 < p_2$

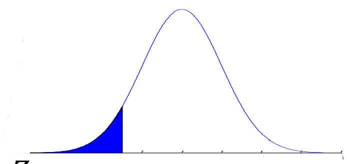
2) $\hat{p}_1 = \frac{X_1}{n_1} = \frac{40}{100} = 0.40$, $\quad \hat{p}_2 = \frac{X_2}{n_2} = \frac{80}{120} = 0.67,$ $\quad \hat{p} = \frac{X_1+X_2}{n_1+n_2} = \frac{40+80}{100+120} = 0.545$

$$S_{\hat{p}} = \sqrt{\hat{p}(1-\hat{p})\left[\frac{1}{n_1} + \frac{1}{n_2}\right]} = \sqrt{(0.545)(1-0.545)\left[\frac{1}{100} + \frac{1}{120}\right]} = 0.067$$

$$Z_h = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{S_{\hat{p}}} = \frac{(0.40 - 0.67) - 0}{0.067} = -4.03$$

3) $H_0: p_1 = p_2$

$H_1: p_1 < p_2$



$-Z_{T(\alpha=0.10)=-2.33}$

$Z_H = -4.03 < -Z_{T(0.10)} = -2.33$ olduğundan H_0 red edilir.

Yorum: İlaçla tedavinin ameliyatla tedaviye göre daha etkin olduğu %90 güvenle söylenebilir.

Güven Aralığı

$$P\left((\hat{p}_1 - \hat{p}_2) - Z_{T(\alpha/2)}S_{\hat{p}} \le (p_1 - p_2) \le (\hat{p}_1 - \hat{p}_2) - Z_{T(\alpha/2)}S_{\hat{p}}\right) = 1 - \alpha$$

$$(\hat{p}_1 - \hat{p}_2) \mp Z_{T(\alpha/2)}S_{\hat{p}} = (0.40 - 0.67) \mp 1.645(0.067) \begin{array}{l} \nearrow -0.380 \\ \searrow -0.160 \end{array}$$

$p_1 - p_2: [-0.380, -0.160]$ Bu aralığın $p_1 - p_2$'yi içeren aralıklardan biri olması olasılığı %90'dır.

Örnek 4.11. Grip aşısının etkisini görebilmek için bir çalışma yapılmıştır. Rasgele seçilmiş 3000 kişiye grip aşısı yapılmış ve daha sonra bu gruptan 130 kişinin grip olduğu gözlenmiştir. Kontrol grubu olarak 2500 kişi seçilmiş ve bunlara aşı yapılmamıştır. Daha sonra aşı yapılmayan 170 kişinin grip olduğu gözlenmiştir. Grip aşısı olan ve olmayan kişi yüzdeleri arasındaki farkın önem kontrolünü %95 güven düzeyinde sınayınız. %95 güven düzeyinde $(p_1 - p_2)$ parametresi için güven aralığını oluşturunuz.

1) $H_0: p_1 = p_2$

$H_1: p_1 < p_2$

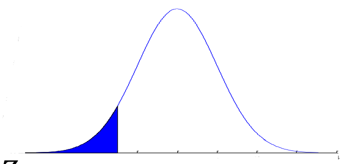
2) $\hat{p}_1 = \frac{X_1}{n_1} = \frac{130}{3000} = 0.043$, $\quad \hat{p}_2 = \frac{X_2}{n_2} = \frac{170}{2500} = 0.068,$ $\quad \hat{p} = \frac{X_1+X_2}{n_1+n_2} = \frac{130+170}{3000+2500} = 0.055$

$$S_{\hat{p}} = \sqrt{\hat{p}(1-\hat{p})\left[\frac{1}{n_1} + \frac{1}{n_2}\right]} = \sqrt{(0.055)(1-0.055)\left[\frac{1}{3000} + \frac{1}{2500}\right]} = 0.0062$$

$$Z_h = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{S_{\hat{p}}} = \frac{(0.043 - 0.068) - 0}{0.0062} = -4.032$$

3) $H_0: p_1 = p_2$

$H_1: p_1 < p_2$



$-Z_{T(\alpha=0.05)=-1.645}$

$Z_H = -4.032 < -Z_{T(0.05)} = -1.645$ olduğundan H_0 red edilir.

Yorum: Bir kişi grip aşısı olduğunda gribe yakalanma olasılığının daha küçük olduğu %95 güvenle söylenebilir.

Güven Aralığı

$$P\left((\hat{p}_1 - \hat{p}_2) - Z_{T(\alpha/2)}S_{\hat{p}} \le (p_1 - p_2) \le (\hat{p}_1 - \hat{p}_2) - Z_{T(\alpha/2)}S_{\hat{p}}\right) = 1 - \alpha$$

-0.0372

$(\hat{p}_1 - \hat{p}_2) \mp Z_{T(\alpha/2)}S_{\hat{p}} = (0.043 - 0.068) \mp 1.96(0.0062)$

-0.0128

$p_1 - p_2: [-0.0372, -0.0128]$ Bu aralığın $p_1 - p_2$'yi içeren aralıklardan biri olması olasılığı %95'dir.