I slutet av förra veckan lärde vi oss att ta fram konfidensintervall och utföra hypotestest för skillnaden mellan två väntevärden $\mu_X - \mu_Y$. Vi utgick då från normalfördelningen, dvs centrala gränsvärdessatsen, dvs vi ville att stickprovsstorlekarna, n_1 och n_2, skulle vara stora. Vad gör vi om n_1 och n_2 inte är stora?

Om populationerna är <u>normalfördelade</u> och om dess <u>varianser kan antas vara lika stora</u> så använder vi följande (sida 433).

Ett $(1-\alpha)$%-igt konfidensintervall för $\mu_X - \mu_Y$ ges av

$$L = (\bar{X} - \bar{Y}) - t_{\frac{\alpha}{2}, n_1+n_2-2} \cdot \sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$U = (\bar{X} - \bar{Y}) + t_{\frac{\alpha}{2}, n_1+n_2-2} \cdot \sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

där S_p^2 är den poolade stickprovsvariansen:

$$S_p^2 = \frac{(n_1-1)S_X^2 + (n_2-1)S_Y^2}{n_1+n_2-2}$$

Vi kan även göra ett hypotestest

$$H_0: (\mu_X - \mu_Y) = D_0$$

$$H_1: (\mu_X - \mu_Y) \neq D_0$$

och använder då teststatistikan

$$T = \frac{(\bar{X} - \bar{Y}) - D_0}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

som är t-fördelad med $n_1 + n_2 - 2$ frihetsgrader. Vi jämför alltså den observerade teststatistikan med de kritiska värdena från t-tabellen.

Gör uppgift 12! ~~[illegible]~~

~~[illegible]~~

Om varianserna inte kan antas vara lika stora men om stickproven är lika stora, dvs om $n_1 = n_2$, använder vi följande formel för KI. $(n = n_1 = n_2)$

$$L = (\bar{X} - \bar{Y}) - t_{\frac{\alpha}{2}, n_1 - n_2 - 2} \cdot \sqrt{\frac{S_X^2 + S_Y^2}{n}}$$

$$U = (\bar{X} - \bar{Y}) + t_{\frac{\alpha}{2}, n_1 - n_2 - 2} \cdot \sqrt{\frac{S_X^2 + S_Y^2}{n}}$$

Om vi då vill testa en hypotes, t. ex.

$$H_0: (\mu_X - \mu_Y) = D_0$$
$$H_1: (\mu_X - \mu_Y) \neq D_0$$

använder vi teststatistikan

$$\cancel{T = \frac{(\bar{X} - \bar{Y}) - D_0}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}} \quad T = \frac{(\bar{X} - \bar{Y}) - D_0}{\sqrt{\frac{S_X^2 + S_Y^2}{n}}}$$

som är t-fördelad med $2 \cdot (n-1)$ frihetsgrader.

Gör uppgift 15!

Jämförelse av parade stickprov. (kap 9.3)

Ofta vill man jämföra två stickprov där observationerna ~~[illegible]~~ är parade mellan stickproven. Ett vanligt exempel är "före" och "efter" något. T. ex. blodtryck hos patienter före och efter tagen medicin, vikt hos personer före och efter bantningskur, osv. Om man vill analysera skillnaden mellan de parade observationerna ser vi då dessa skillnader som ett stickprov:

$$D_1 = X_1 - Y_1, \ D_2 = X_2 - Y_2, \ \ldots, \ D_n = X_n - Y_n.$$

OBS: Eftersom det är parade stickprov så vet vi att $n_1 = n_2$, så vi använder n istället.

Vi kan nu göra konfidensintervall och hypotestest för ~~denna~~ ~~skillnad~~ väntevärdet μ_d för denna skillnad.

(4)

Ett $(1-\alpha)\%$-igt konfidensintervall för μ_d ges av

om n stort och σ_d känt

$$L = \bar{D} - z_{\frac{\alpha}{2}} \cdot \frac{\sigma_d}{\sqrt{n}}$$

$$U = \bar{D} + z_{\frac{\alpha}{2}} \cdot \frac{\sigma_d}{\sqrt{n}}$$

om n stort och σ_d okänt

$$L = \bar{D} - z_{\frac{\alpha}{2}} \cdot \frac{s_d}{\sqrt{n}}$$

$$U = \bar{D} + z_{\frac{\alpha}{2}} \cdot \frac{s_d}{\sqrt{n}}$$

om n litet och populationerna normalfördelade

$$L = \bar{D} - t_{\frac{\alpha}{2}, n-1} \cdot \frac{s_d}{\sqrt{n}}$$

$$U = \bar{D} + t_{\frac{\alpha}{2}, n-1} \cdot \frac{s_d}{\sqrt{n}}$$

För att testa en hypotes om μ_d, t. ex.

$H_0: \mu_d = D_0$ eller $H_0: \mu_d \geq D_0$

$H_1: \mu_d \neq D_0$ $H_1: \mu_d < D_0$

använde vi oss av teststatistikorna

om n stort och σ_d känt

$$Z = \frac{\bar{D} - D_0}{\sigma_d/\sqrt{n}}$$ som är $N(0,1)$-fördelad

om n stort och σ_d okänt

$$Z = \frac{\bar{D} - D_0}{s_d/\sqrt{n}}$$ som är $N(0,1)$-fördelad

om n litet och populationerna normalfördelade

$$T = \frac{\bar{D} - D_0}{s_d/\sqrt{n}}$$ som är t-fördelad med n-1 frihetsgrader