5) Given $D_1 = \{x_1, x_2, \dots, x_n\}$ be iid from Normal(μ_1, σ_1^2)

$D_2 = \{y_1, y_2, \dots, y_m\}$ be iid from Normal (μ_2, σ_2^2)

x's and y's are independent and $\mu_1, \sigma_1^2, \mu_2, \sigma_2^2$ are unknown

then the hypothesis is

$H_0: \mu_1 > \mu_2$ $\qquad$ $H_1: \mu_1 <= \mu_2$

using unpaired T-test with thresold value $\delta > 0$ to Check hypothesis

$$T = \frac{\bar{D}}{\sqrt{\frac{S_x^2}{m} + \frac{S_y^2}{n}}} \qquad \text{where } D = \frac{\sum_{i=1}^{n} x_i}{n} - \frac{\sum_{i=1}^{m} y_i}{m} = \bar{x} - \bar{y}$$

Since this is a one-Sided test, to Validate null hypothesis we need

$$T > -\delta$$

Assume n & m are large

Using CLT we get $\bar{x} \sim Nor\left(\mu_1, \frac{\sigma_1^2}{n}\right)$ and

$\bar{y} \sim Nor\left(\mu_2, \frac{\sigma_2^2}{m}\right)$

$$\therefore \bar{D} = \bar{x} - \bar{y} \sim Nor\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n} - \frac{\sigma_2^2}{m}\right)$$

As σ_1^2 and σ_2^2 are unknown. we can replace them with their plugin estimators.

$$\bar{D} \sim Nor\left(\mu_1 - \mu_2, \frac{S_x^2}{n} + \frac{S_y^2}{m}\right)$$

$$Pr(\text{Type I error}) = Pr(\text{reject } H_0 / H_0 \text{ true})$$

$$= P(T < -\delta)$$

$$= Pr\left(\frac{\bar{D}}{\sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}} < -\delta\right)$$

$$= Pr\left(\bar{D} < -\delta\sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}\right)$$

$$= Pr\left(\bar{D} - (\mu_1 - \mu_2) < -\delta\sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}} - (\mu_1 - \mu_2)\right)$$

$$= Pr\left(\frac{\bar{D} - (\mu_1 - \mu_2)}{\sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}} < -\delta - \frac{(\mu_1 - \mu_2)}{\sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}}\right)$$

$$= \overline{\Phi}\left(-\delta - \frac{(\mu_1 - \mu_2)}{\sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}}\right)$$

$$Pr(\text{Type I error}) = \Phi\left(-\delta - \frac{(\mu_1 - \mu_2)}{\sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}}\right)$$

$$Pr(\text{Type II error}) = Pr(\text{Accept } H_0 \mid H_1 \text{ true})$$

$$= Pr(T > -\delta)$$

$$= Pr\left(\frac{\bar{D}}{\sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}} > -\delta\right)$$

Applying same steps as of type I error.

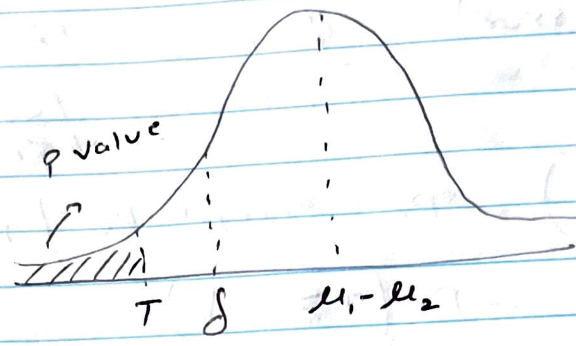
$$Pr(T > -\delta) = Pr\left(\frac{\bar{D} - (\mu_1 - \mu_2)}{\sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}} > -\delta - \frac{(\mu_1 - \mu_2)}{\sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}}\right)$$

$$= 1 - \Phi\left(-\delta - \frac{(\mu_1 - \mu_2)}{\sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}}\right)$$

5b) For the unpaired test we get T-Statistic as

$$T = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}}$$

If we reject H_0, then $T < -\delta$



$$\text{Pvalue} = \Pr\left(\frac{\bar{D}}{\sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{n}}} < \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}} \right)$$

$$= \Pr\left(\bar{D} < \bar{X} - \bar{Y}\right)$$

$$= \Pr\left(\frac{\bar{D} - \mu_1 - \mu_2}{\sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}} < \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}} \right)$$

$$= \Pr\left(Z < \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}} \right)$$

$$\text{P-value} = \Phi\left(\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}} \right)$$