A / B Testing

User-level : $\dfrac{\sum X_i}{n} = \dfrac{3+1+2}{1+1+1}$

Ratio : $\dfrac{\sum X_i}{\sum Y_i} = \dfrac{3+1+2}{6+7+8}$

$$X \sim N(\mu_x, \sigma_x^2)$$

$$Y \sim N(\mu_y, \sigma_y^2)$$

$$\begin{bmatrix} X \\ Y \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, \begin{bmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{bmatrix} \right)$$

$$R = \frac{X}{Y} \qquad \frac{E(X)}{E(Y)}$$

$$Z = \frac{E(X)}{E(Y)} + \frac{1}{E(Y)}\left(X - \frac{E(X)}{E(Y)} Y \right)$$

t-test | Welch t-test

$X_1 - X_2$

Assumptions:

n_1 $\quad X_1 \sim N(\mu_1, \sigma^2)$ | $X_1 \sim N(\mu_1, \sigma_1^2)$

n_2 $\quad X_2 \sim N(\mu_2, \sigma^2)$ | $X_2 \sim N(\mu_2, \sigma_2^2)$

$H_0: \mu_1 = \mu_2$ | $H_a: \mu_1 \neq \mu_2$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{\bar{\delta}}} \sim t_{n_1+n_2-2}$$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S_{\bar{X}_1}^2 + S_{\bar{X}_2}^2}} \sim t_\nu$$

$$S_{\bar{\delta}} = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

$$S_{\bar{X}_i}^2 = \frac{S_i}{\sqrt{n_i}}$$

$$\nu \approx \frac{\left(\frac{S_1^2}{N_1} + \frac{S_2^2}{N_2}\right)^2}{\frac{S_1^4}{N_1^2 \cdot \nu_1} + \frac{S_2^4}{N_2^2 \nu_2}}$$

$n_1 = n_2 = n$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_p \cdot \sqrt{\frac{2}{n}}}$$

pooled

$\nu_i = n_i - 1$

$$s_p = \sqrt{\frac{s_1^2 + s_2^2}{2}}$$

Levene's Test:

$H_0: \sigma_1^2 = \dots = \sigma_k^2$

$H_a: \exists\, i,j \quad \sigma_i \neq \sigma_j$

$$F = \frac{\text{Var Between}}{\text{Var Within}} = \frac{\frac{1}{k-1} \sum N_i \left(z_{i1} - z_2\right)^2}{\frac{1}{N-k} \sum_{i=1}^{k} \sum_{j=1}^{N_i} \left(z_{ij} - z_{i1}\right)^2}$$

$$z_{ij} = |Y_{ij} - \bar{Y}_i|$$

z_{i1} mean group

z_2 mean pooled sample