HW 3 Due 10/11/2022

a) $H_0: M_1 = M_2 = M_3$ → Means of all three brands are the same

$H_A: M_1 \neq M_2 \neq M_3$ → One or more means are not equal with each other

Brand 1 mean = 95.2, Brand 2 mean = 79.4, Brand 3 mean = 100.4

→ Mean of mean of brands : $\frac{95.2 + 79.4 + 100.4}{3} = 91.667$

Find SSR $(n\Sigma(\bar{x}_k - \bar{\bar{x}})^2)$

→ $SSR = 5(95.2_1 - 91.667)^2 + 5(79.4_2 - 91.667)^2 + 5(100.4_3 - 91.667)^2$

$= \boxed{1196.13}$

Find SSE $(\Sigma(x_{ik} - \bar{x}_k)^2)$

→ $\Sigma((100-95.2)^2 + (96-95.2)^2 + \ldots + (92-95.2)^2)$

$+ (76-79.4)^2 + (80-79.4)^2 + \ldots + (82-79.4)^2$

$+ (108-100.4)^2 + (100-100.4)^2 + \ldots + (100-100.4)^2)$

$= 44.8 + 59.2 + 83.2 = \boxed{187.2}$

Since df of brands = K-1 = 2 and df of error = n-K = 12

MST of brand = $(1196.13 + 187.2)/2 = 578.06$

MSE of error = $187.2/12 = 15.6$

→ F statistic = $578.06/15.6 = 38.34$

→ $P = 0.0001 < \alpha = 0.05$

→ we reject the null, and conclude that mean lives of batteries from each brand is distinct

b) The data indicates a skewed, and thus not normally distributed, residual plot

c) For a 95% CI for brand 2: $M_{brand 2} \pm t_{0.025}(df\ error)\left(\sqrt{\frac{MSE}{n}}\right)$

$\hookrightarrow 79.4 \pm 2.1788\left(\sqrt{\frac{15.6}{5}}\right) = 79.4 \pm 3.85$

95% CI of brand 2 = $\boxed{(75.55, 83.25)}$

d) For a 99% CI for the difference $M_{brand 3} - M_{brand 2}$:

$$M_{brand 3} - M_{brand 2} \pm t_{0.005}(df\ error)\left(\sqrt{\frac{2 \cdot MSE}{n}}\right)$$

$$= 100.4 - 79.4 \pm 3.428\left(\sqrt{\frac{2 \cdot 15.6}{5}}\right) = 21 \pm 8.563$$

99% CI of difference $M_{brand 3} - M_{brand 2}$ = (12.44, 29.56

e) based on prior calculations, brand 3 has the highest mean battery life, as well as the highest values for each of the randomly selected batteries among all 15 total batteries.

f) $85 - 100.4 / \sqrt{\frac{15.6}{5}} = -8.72$ at df = 12, which indicates that for brand 3, it is very unlikely that the company would have to replace any batteries.

2) F statistic $= \frac{MST}{MSE} = \frac{SST}{a-1} \Big/ \frac{SSE}{N-a} \to \frac{SST}{2-1} \Big/ \frac{SSE}{N-2}$

$\to MST = \frac{SST}{2-1} = \sum n(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2$

$\to \sum n_1(\bar{y}_{1\cdot} - \bar{y}_{\cdot\cdot})^2 + n_2(\bar{y}_{2\cdot} - \bar{y}_{\cdot\cdot})^2$

$\to MSE = \frac{SSE}{N-2} = \left[\frac{\sum(y_1 - \bar{y}_{1\cdot})^2 + \sum(y_2 - \bar{y}_{2\cdot})^2}{N-2}\right] \to$ USE $S_i = \frac{\sum(y_i - \bar{y}_{i\cdot})^2}{n_i - 1}$

$= \frac{(n_1-1)\frac{\sum(y_1-\bar{y}_{1\cdot})^2}{(n_1-1)} + (n_2-1)\frac{\sum(y_2-\bar{y}_{2\cdot})^2}{n_2-1}}{\underbrace{n_1+n_2}_{\to N = n_1+n_2} - 2}$

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

$\to$ Since $SST = \sum n_1(\bar{y}_{1\cdot} - \bar{y}_{\cdot\cdot})^2 + n_2(\bar{y}_{2\cdot} - \underbrace{\bar{y}_{\cdot\cdot}}_{\to \text{global mean}})^2$

$= \frac{n_1\bar{y}_1 + n_2\bar{y}_2}{N}$

$\therefore SST = \underbrace{n_1\left(\bar{y}_{1\cdot} - \frac{n_1\bar{y}_1 + n_2\bar{y}_2}{N}\right)^2}_{} + n_2\left(\bar{y}_{2\cdot} - \frac{n_1\bar{y}_1 + n_2\bar{y}_2}{N}\right)^2$

$\to n_1\left(\frac{N\bar{y}_{1\cdot} - n_1\bar{y}_1 + n_2\bar{y}_2}{N}\right)^2 = n_1\left(\frac{\bar{y}_{1\cdot}(N-n_1) + n_2\bar{y}_2}{N}\right)^2$

$= n_1\left(\frac{n_2\bar{y}_{1\cdot} + n_2\bar{y}_2}{N}\right)^2$

Repeat same process for $n_2\left(\bar{y}_{2\cdot} - \frac{n_1\bar{y}_1 + n_2\bar{y}_2}{N}\right)^2$

$SST = n_1\left(\frac{n_2\bar{y}_{1\cdot} - n_2\bar{y}_2}{N}\right)^2 + n_2\left(\frac{n_1\bar{y}_2 - n_1\bar{y}_{1\cdot}}{N}\right)^2$

$= \frac{n_1 n_2^2}{N^2}(\bar{y}_{1\cdot} - \bar{y}_{2\cdot})^2 + \frac{n_2 n_1^2}{N^2}(\bar{y}_{2\cdot} - \bar{y}_{1\cdot})^2$

$= \left(\frac{n_1 n_2^2}{N^2} + \frac{n_2 n_1^2}{N_2}\right)(\bar{y}_{1\cdot} - \bar{y}_{2\cdot})$

$$SST = \left(\frac{n_1 n_2^2}{N^2} + \frac{n_2 n_1^2}{N^2}\right)(\bar{y}_{1.} - \bar{y}_{2.})^2$$

$$= \frac{n_1 n_2 (n_1 + n_2)}{N^2}(\bar{y}_{1.} - \bar{y}_{2.})^2$$

$$= \frac{n_1 n_2 N}{N^2}(\bar{y}_{1.} - \bar{y}_{2.}) = \frac{1}{\frac{N}{n_1 n_2}}(\bar{y}_{1.} - \bar{y}_{2.})^2 = \frac{1}{\frac{1}{n_1} + \frac{1}{n_2}}(\bar{y}_{1.} - \bar{y}_{2.})^2$$

$$\frac{SSE}{N-2} = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} = \text{Pooled variance estimate} = S_{pooled}$$

Since F statistic $= {}^{SST}/_{2-1} \Big/ {}^{SSE}/_{N-2}$

$$\hookrightarrow \boxed{\frac{(\bar{y}_1 - \bar{y}_2)^2}{S_{pooled}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = t^2}$$