WRITTEN PORTION.

HW #16

Problem 3: Show that the F statistic for a one-way analysis of variance with $I=2$ is t^2.

$$F = \frac{SS_B/(I-1)}{SS_W/[I(J-1)]} = \frac{SS_B}{SS_W/2(J-1)}$$

$$= \frac{\sum_i J_i(\bar{Y}_{i.}-\bar{Y}_{..})^2}{\frac{\sum_i\sum_j (Y_{ij}-\bar{Y}_{i.})^2}{2(J-1)}}$$

Now, recall that $\bar{Y}_{..} = \frac{1}{IJ}\sum_i\sum_j Y_{ij}$.

$\Rightarrow$ In this case $\bar{Y}_{..} = \frac{1}{2J}[Y_{1.}+Y_{2.}]$

$\bar{Y}_{..} = \frac{1}{2}[\bar{Y}_{1.}+\bar{Y}_{2.}]$

Now we focus on simplifying the numerator.

NUMERATOR: $\sum_i J_i(\bar{Y}_{i.}-\bar{Y}_{..})^2 = J_1(\bar{Y}_{1.}-\bar{Y}_{..})^2 + J_2(\bar{Y}_{2.}-\bar{Y}_{..})^2$

$$= J_1\left(\bar{Y}_{1.}-\left(\frac{\bar{Y}_{1.}+\bar{Y}_{2.}}{2}\right)\right)^2 + J_2\left(\bar{Y}_{2.}-\left(\frac{\bar{Y}_{1.}+\bar{Y}_{2.}}{2}\right)\right)^2$$

$$= J_1\left(\frac{\bar{Y}_{1.}-\bar{Y}_{2.}}{2}\right)^2 + J_2\left(\frac{\bar{Y}_{2.}-\bar{Y}_{1.}}{2}\right)^2$$

Let $J_1 = J_2 = J \Rightarrow J\left(2\left(\frac{\bar{Y}_{1.}-\bar{Y}_{2.}}{2}\right)^2\right) = \frac{J}{2}(\bar{Y}_{1.}-\bar{Y}_{2.})^2$

DENOMINATOR: $\frac{\sum_i\sum_j (Y_{ij}-\bar{Y}_{i.})^2}{2(J-1)} = \frac{\sum_j (Y_{1j}-\bar{Y}_{1.})^2 + \sum_j (Y_{2j}-\bar{Y}_{2.})^2}{2(J-1)}$

Recall $s^2_{\bar{Y}_1-\bar{Y}_2} = \frac{\sum_j (y_{1j}-\bar{y}_{1.})^2 + \sum_j (y_{2j}-\bar{y}_{2.})^2}{J(J-1)}$

In this case $J(J-1)\, s^2 = \sum_j (y_{1j}-\bar{y}_{1.})^2 + \sum_j (y_{2j}-y_{2.})^2$.

Back to the Denominator $\quad \dfrac{J(J-1)s^2}{2(J-1)} = \dfrac{J}{2} s^2_{\bar{x}_1 - \bar{x}_2}$

Now: $\quad F = \dfrac{\frac{J}{2}(\bar{Y}_{1.} - \bar{Y}_{2.})^2}{\frac{J}{2} s^2_{\bar{x}_1 - \bar{x}_2}} = \dfrac{(\bar{Y}_{1.} - \bar{Y}_{2.})^2}{s^2_{\bar{x}_1 - \bar{x}_2}}$

$$= \left[\frac{\bar{Y}_{1.} - \bar{Y}_{2.}}{s_{\bar{x}_1 - \bar{x}_2}}\right]^2 = t^2 \quad \text{As desired}$$

Notice $s_{\bar{x}_1 - \bar{x}_2} = s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$

Problem 6:

$P(\bigcap_{i=1}^{n} A_i) \geq 1 - \sum_{i=1}^{n} P(A_i^c)$.

Recall the identity $P(A \cap B) = P(A) + P(B) - P(A \cup B)$.

Notice then that $P(A \cap B) \leq P(A) + P(B)$ and this is our base case. We now WTS this is true for $n+1$.

Assume the identity holds for n: $P(\bigcap_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} P(A_i)$

Now intersect one additional set $P(\bigcap_{i=1}^{n+1} A_i) \leq \sum_{i=1}^{n} P(A_i)$

is the same because the LHS either stays the same or gets smaller. Then, $P(\bigcap_{i=1}^{n+1} A_i) \leq \sum_{i=1}^{n} P(A_i) + P(A_{n+1})$

$$\leq \sum_{i=1}^{n+1} P(A_i)$$

Thus this holds for $n+1$ instances. So $P(\bigcap_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} P(A_i)$ due to the principal of mathematical induction.

WRITTEN PORTION.

Now, $P(\bigcap_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} P(A_i) \Rightarrow -P(\bigcap_{i=1}^{n} A_i) \geq -\sum_{i=1}^{n} P(A_i)$

$$\Rightarrow 1 - P(\bigcap_{i=1}^{n} A) \geq 1 - \sum_{i=1}^{n} P(A_i)$$

$$P(\bigcap_{i=1}^{n} A) \geq -P(\bigcap_{i=1}^{n} A_i) \geq 1 - \sum_{i=1}^{n} P(A_i) - 1$$

$$\geq \sum_{i=1}^{n} P(A_i) \geq -\sum_{i=1}^{n} P(A_i)$$

So $P(\bigcap_{i=1}^{n} A_i) \geq \sum_{i=1}^{n} P(A_i)$

Thus $P(\bigcap_{i=1}^{n} A_i) \geq 1 - \sum_{i=1}^{n} P(A_i^c)$

So what is A_i and A_i^c in terms of simultaneous CI?

A_i is the collection of intersecting closed sets, where $\bigcap_{i=1}^{n} A_i$ is where the closed sets connect.

A_i^c is the area outside of the closed sets.

What the above statement is saying is that landing in the region where these sets intersect will result in a p-value greater than or equal to the p-value of landing outside these bounds.

$\bar{Y}_{1.} = 4.062$
$\bar{Y}_{4.} = 3.92$
$\bar{Y}_{5.} = 3.957$
$\bar{Y}_{6.} = 3.955$

$S_p = \sqrt{MS_w} = \sqrt{0.0037} = 0.06083$

$\alpha = 0.05$

In this case $I=7,\ J=10.$

$K = \binom{7}{2} = 21$

Problem 8: Form simultaneous confidence intervals for the difference of the mean of lab 1 and those of labs 4, 5, and 6 in Example A Section 12.2.2.1

$$\bar{Y}_{i_1.} - \bar{Y}_{i_2.} \pm t_{I(J-1),\ \alpha/2k} \sqrt{\frac{2}{J}}\, S_p$$

Bonferroni Method:

Labs 1 & 4

$$4.062 - 3.92 \pm t_{63,\ 1/420} \cdot \sqrt{\frac{1}{5}}\,(0.06083)$$

$$(0.142) \pm 2.926567\,(0.027204) \Rightarrow (0.0623,\ 0.2216)$$

Labs 1 & 5

$$4.062 - 3.957 \pm 2.92657\,(0.027204) \Rightarrow (0.0254,\ 0.1846)$$

Labs 1 & 6

$$4.062 - 3.955 \pm 2.9266\,(0.0272) \Rightarrow (0.0274,\ 0.1866)$$

Problem 10 : Suppose a one-way layout has 10 treatments and 7 observations. What is the ratio of the length of a simultaneous confidence interval for CI made w/ Tukey's method vs. Bonferroni's. How do they both compare in length to intervals based on the t-dist?

Givens : $I = 10$

$J = 7$

$K = \binom{10}{2} = 45$.

Bonferroni's CI : $\bar{Y}_{i_1.} - \bar{Y}_{i_2.} \pm t_{I(J-1), \frac{\alpha}{2k}} \sqrt{\frac{2}{J}}\, s_p$

Tukey's CI : $\bar{Y}_{i_1.} - \bar{Y}_{i_2.} \pm q_{I, I(J-1)}(\alpha) \sqrt{\frac{1}{J}}\, s_p$

WTFind : $\frac{\text{Tukey length}}{\text{Bonferroni length}} = \frac{q_{I, I(J-1)}(\alpha)\sqrt{\frac{1}{J}}\,\cancel{s_p}}{t_{I(J-1), \frac{\alpha}{2k}}\sqrt{\frac{2}{J}}\,\cancel{s_p}}$

$$= \left(\frac{\sqrt{10}}{\sqrt{2}}\right) * \frac{1.426924}{3.426019} \approx 0.9313.$$

* So Bonferroni's CI is 1.0737 times larger than Tukey's CI

Computing the conf int for difference between two means

$\hookrightarrow (\bar{X}_1 - \bar{X}_i) \pm t_{\frac{\alpha}{2}, 2J-2} * s_p * \sqrt{\frac{2}{J}}$

WTFind ① $\frac{\text{t-dist}}{\text{Tukey}}$; ② $\frac{\text{t-dist}}{\text{Bonferroni}}$.

① $\frac{\text{t-dist}}{\text{Tukey}} = \frac{(2.179)\,\cancel{s_p}\sqrt{\frac{2}{7}}}{(1.427)\,\cancel{s_p}\sqrt{\frac{10}{7}}} = \sqrt{\frac{1}{5}}\,(1.527) \approx 0.683$

* So Tukey's CI is 1.464 times larger than the t-dist.

② $\frac{\text{Bonferroni}}{\text{t-dist}} = \frac{(3.426)\sqrt{\frac{2}{7}}\,\cancel{34}}{(2.179)\sqrt{\frac{2}{7}}\,\cancel{34}} \approx 1.572$

* Bonferroni's CI are 1.572 times larger than the t-dist's.