## Pharm 3011 - Fall 2019 - Assignment 1 solutions

(out of 40 points)

1. An analysis of variance is to be carried out comparing the mean forced expiratory volume in patients with coronary artery disease. Patients were sampled at each of three centres, leading to the following summary data:

| location | i | sample size <br> $n_{i}$ | sample mean <br> $\bar{x}_{i .}$ | sample variance <br> $s_{i}^{2}$ |
| :--- | :---: | :---: | :---: | :---: |
| Johns Hopkins | 1 | 10 | 2.7 | .25 |
| Rancho Los Amigos | 2 | 10 | 3.0 | .30 |
| St. Louis | 3 | 5 | 2.9 | .25 |

Using these summary statistics:
(a) Calculate the grand mean $\bar{x}$..
$>\operatorname{xbar}=(10 * 2.7+10 * 3.0+5 * 2.9) / 25$
> xbar
[1] 2.86
(b) Calculate the treatment sum of squares $S S_{T r}$.
$>\operatorname{SSTr}=10 *(2.7-x b a r)^{\wedge} 2+10 *(3.0-x b a r)^{\wedge} 2+5 *(2.9-x b a r)^{\wedge} 2$
> SSTr
[1] 0.46
(c) Calculate the error sum of squares $S S E$.
$>S S E=(10-1) * .25+(10-1) * .30+(5-1) * .25$
> SSE
[1] 5.95
2. Fill in the blanks in the following analysis of variance table. Evaluate the p-value as accurately as possible using the F tables provided on the course website. (For example $.01<$ p-value $<.05$ )

| Source | SS | df | MS | F | p-value |
| :--- | ---: | ---: | :--- | :--- | :--- |
| Treatment | $\mathbf{3 2 0}$ | 4 | $\mathbf{8 0}$ | $\mathbf{7 . 1 1}$ | $\mathbf{. 0 0 1}<\mathbf{p}$-value $<\mathbf{. 0 1}$ |
| Error | 180 | $\mathbf{1 6}$ | $\mathbf{1 1 . 2 5}$ |  |  |
| Total | 500 | 20 |  |  |  |

(6 points - 1 for each of: treatment SS, error df, MSTr, MSE, F, and p-value. Subtract one point for each error, but do not penalize for accumulated errors.)
3. A remotivation team in a psychiatric hospital conducted an experiment to compare five methods for remotivating patients. Patients were grouped according to level of intial motivation (IM). Patients in each group were randomly assigned to the five methods (METH). At the end of the experimental period the patients were evaluated by a team composed of a psychiatrist, a psychologist, a nurse, and a social worker, none of whom was aware of the method to which patients had been assigned. The team assigned each patient a composite score as a measure of his or her level of motivation. The results were as follows:

| Level of |  |  |  |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: |
| initial | Remotivation method |  |  |  |  |
| motivation | A | B | C | D | E |
| Nil | 68 | 68 | 60 | 68 | 64 |
| Very low | 62 | 70 | 65 | 80 | 69 |
| Low | 67 | 78 | 68 | 81 | 70 |
| Average | 70 | 81 | 70 | 89 | 74 |

A partial ANOVA table is shown below:

| Source | DF | SS | MS |
| :--- | ---: | ---: | ---: |
| IM | 3 | 346.2 |  |
| METH |  | 525.8 |  |
| Error | 12 | 137.8 | 11.483 |
| Total |  |  |  |

Assess whether there is an effect of initial motivation (IM)
(a) State the hypotheses.
$H_{0}$ : no effect of IM
$H_{A}$ : there is an effect of IM
(b) What is the mean square for initial motivation?
$M S_{I M}=S S_{I M} / d f_{I M}=346.2 / 3=115.4$.
(c) Calculate the test statistic, and state the degrees of freedom.

$$
\begin{equation*}
F_{o b s}=M S_{I M} / M S E=115.4 / 11.483 \approx 10.05 \tag{2}
\end{equation*}
$$

There are 3 and 12 degrees of freedom (1 point for each of the degrees of freedom)
(d) Put bounds on the P value.

The p-value is $P\left(F_{3,12}>F_{\text {obs }}\right)=P\left(F_{3,12}>10.05\right)$.
From the tables used in class, $.001<P\left(F_{3,12}>10.05\right)<.01$.
(e) Give a conclusion at level $\alpha=.01$.

Reject $H_{0}$ at level . 01 .
4. The following table gives data for a two factor experiment with two replications.

|  |  | FACTOR B |  |  |
| :---: | :---: | :---: | :---: | :---: |
|  |  | 1 | 2 | 3 |
| FACTOR A | 1 | $3.1,4.0$ | $4.6,4.2$ | $6.4,7.1$ |
|  | 2 | $5.9,5.3$ | $2.9,2.2$ | $3.3,2.5$ |

(a) A partial ANOVA table is given below. FILL IN THE MISSING ENTRIES.

| Source | DF | SS | MS | F |
| :--- | ---: | ---: | ---: | ---: |
| A |  | 4.44 |  |  |
| B |  | 4.13 |  |  |
| INTERACTION | 2 | 18.00 | 9.00 | 36 |
| ERROR | 6 |  | .25 |  |
| TOTAL |  | 28.05 |  |  |


| Source | DF | SS | MS | F |
| :--- | ---: | ---: | ---: | ---: |
| A | 1 | 4.44 | 4.44 | 17.76 |
| B | 2 | 4.13 | 2.065 | 8.26 |
| INTERACTION | 2 | 18.00 | 9.00 | 36 |
| ERROR | 6 | 1.5 | .25 |  |
| TOTAL | 11 | 28.05 |  |  |

(8 points, 1 each for $\operatorname{df}(\mathrm{A}), \operatorname{df}(\mathrm{B})$, SSE, MSA, MSB, MSE and each of the two F values ). SSE could also be reported as 1.48 depending on calculation method.
(b) Test for interaction.
i. State the null and alternative hypotheses.
$H_{0}$ : there is no interaction.
$H_{A}$ : there is an interaction.
ii. What is the observed value of the test statistic?
$F_{o b s}=36$
iii. What are the numerator and denominator degrees of freedom? Numerator degrees of freedom is 2 , denominator degrees of freedom is 6 . (1 point for each of the degrees of freedom.)
iv. Bound the p-value as accurately as possible. Using the class tables, the most accurate statement that can be made is $p$-value $<.001$.
v. What is your conclusion when testing at level .01 ?

Reject $H_{0}$, conclude that there is interaction.
(c) Based on your conclusion when testing for interaction, does it make sense to test for the main effects of factors A and B? Why?
No.
Why? If there is interaction, then there are main effects of each factor, but the effect of one factor depends on the level of the other factor. (1 point for some statement similar to this.)

