STAT3119-10 Quiz#5 (25 mins, total 100 points) **Class: 11/21/2019**

Name:___________________________________ GWID: ______________________________

1. In a two-factor study, the investigator was interested in the effects of a child's gender (factor A) and bone development (factor B) on the rate of growth induced by hormone administration. A child's bone development was classified into one of three categories: severely depressed, moderately depressed, mildly depressed. Three children were randomly selected for each gender-bone development group. The response variable (Y) of interest was the difference between the growth rate during growth hormone treatment and the normal growth rate prior to the treatment, expressed in centimeters per month. Four of the 18 children were unable to complete the year-long study, thus creating unequal treatment sample sizes. Data are shown in the table below: (Note: Y_{ijk} *are the observed data, n=14; and* $\bar{Y}_{ij}$ are the treatment means)

Gender (factor A) i	Bone Development (factor B) j: Severely Depressed (B_1)	Moderately Depressed (B_2)	Mildly Depressed (B_3)
Male (A_1)	1.4 (Y_{111}) 2.4 (Y_{112}) 2.2 (Y_{113})	2.1 (Y_{121}) 1.7 (Y_{122})	.7 (Y_{131}) 1.1 (Y_{132})
Mean	2.0 ($\bar{Y}_{11\cdot}$)	1.9 ($\bar{Y}_{12\cdot}$)	.9 ($\bar{Y}_{13\cdot}$)
Female (A_2)	2.4 (Y_{211})	2.5 (Y_{221}) 1.8 (Y_{222}) 2.0 (Y_{223})	.5 (Y_{231}) .9 (Y_{232}) 1.3 (Y_{233})
Mean	2.4 ($\bar{Y}_{21\cdot}$)	2.1 ($\bar{Y}_{22\cdot}$)	.9 ($\bar{Y}_{23\cdot}$)

1. (10 points) Please write down the full two-factor ANOVA factor effects model with main effects and interaction terms.

$$Y_{ijk} = \mu_{\cdot\cdot} + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

2. (10 pts) Since the study data is unbalanced, please specify the equivalent full regression model that we can use to estimate the model parameters:

$$Y_{ijk} = \mu_{\cdot\cdot} + \underbrace{\alpha_1 X_{ijk1}}_{\text{A main effect}} + \underbrace{\beta_1 X_{ijk2} + \beta_2 X_{ijk3}}_{\text{B main effect}}$$
$$+ \underbrace{(\alpha\beta)_{11} X_{ijk1} X_{ijk2} + (\alpha\beta)_{12} X_{ijk1} X_{ijk3}}_{\text{AB interaction effect}} + \varepsilon_{ijk} \qquad \text{Full model}$$

(10 pts) And specify the three indicator functions for factor A and B below

$$X_1 = \begin{cases} 1 & \text{if case from level 1 for factor A} \\ -1 & \text{if case from level 2 for factor A} \end{cases}$$

$$X_2 = \begin{cases} 1 & \text{if case from level 1 for factor B} \\ -1 & \text{if case from level 3 for factor B} \\ 0 & \text{otherwise} \end{cases}$$

$$X_3 = \begin{cases} 1 & \text{if case from level 2 for factor B} \\ -1 & \text{if case from level 3 for factor B} \\ 0 & \text{otherwise} \end{cases}$$

3. (10 pts) How do you test the interaction between the factors A and B using regression models ?

Set and run a reduced regression model without interaction terms

$$Y_{ijk} = \mu_{..} + \underbrace{\alpha_1 X_{ijk1}}_{\text{A main effect}} + \underbrace{\beta_1 X_{ijk2} + \beta_2 X_{ijk3}}_{\text{B main effect}} + \varepsilon_{ijk} \qquad \text{Reduced model}$$

Then construct a F-test by comparing the SSE from the full and reduced models.

4. (20 points) We run the full regression model and obtain the following regression output:

```
                        Estimate Std. Error t value Pr(>|t|)
(Intercept)                  1.7      0.116  14.609    0.000
IndicatorA1                 -0.1      0.116  -0.859    0.415
IndicatorB1                  0.5      0.178   2.813    0.023
IndicatorB2                  0.3      0.158   1.904    0.093
IndicatorA1:IndicatorB1     -0.1      0.178  -0.563    0.589
IndicatorA1:IndicatorB2      0.0      0.158   0.000    1.000
```

MSE= 0.16 with 8 degree of freedom, the critical value is ***t*(0.975, 8)=2.3**

Given the interaction is not significant, we would like to make inference for the factor level means. Please generate the point estimate, its standard error estimate (square root of the variance estimate), and 95% CI for each of the following quantities (multiple comparison adjustment is not needed here).

Parameter θ	$\hat{\theta}$	$\widehat{SE}(\hat{\theta})$	95% confidence interval
Mean response for male $\mu_{1.}$	1.6	0.154	(1.25, 1.95)
Mean response for female $\mu_{2.}$	1.8	0.172	(1.40, 2.20)
Mean difference by gender $\mu_{1.} - \mu_{2.}$	-0.2	0.23	(-0.73, 0.33)

Note: please find out the values and fill in the table using the treatment means, MSE from the data table and the regression output. Showing the correct formula and your work below will only get partial credit without the correct final answer.

Due to round off values, if the formula is correct and the values are approximately correct, it will be OK

Formula for unbalanced study as in Table 23.5 and page 960 of textbook

Approach 1: $\hat{\mu}_{i.} = (\overline{Y}_{i1} + \overline{Y}_{i2} + \overline{Y}_{i3})/3$

$$\sqrt{\hat{var}(\hat{\mu}_{i.})} = \sqrt{\frac{MSE}{3^2}\left(\frac{1}{n_{i1}} + \frac{1}{n_{i2}} + \frac{1}{n_{i3}}\right)}$$

Therefore

$\widehat{\mu 1}$ = (2.0+1.9+0.9)/3= 1.6; $\widehat{\mu 2}$ = (2.4+ 2.1+0.9)/3= 1.8;

SE($\widehat{\mu 1}$)= sqrt(0.16/9*(1/3+1/2+1/2)) = 0.154

SE($\widehat{\mu 2}$)= sqrt(0.16/9*(1+1/3+1/3)) = 0.172

For $\mu_{1.} - \mu_{2.}$, $$\sqrt{\hat{var}(\hat{\mu}_{1.} - \mu_{2.})} = \sqrt{\frac{MSE}{3^2}\left(\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{13}} + \frac{1}{n_{21}} + \frac{1}{n_{22}} + \frac{1}{n_{23}}\right)}$$

Approach 2: A much easier approach is from the regression output for indicatorA:

$$\hat{\mu}_{1.} = \hat{\mu}_{..} + \hat{\alpha}_1 = 1.7 + (-0.1) = 1.6$$

$\hat{\mu}_{1.} - \hat{\mu}_{2.} = 2\hat{\alpha}_1$ =2*(-0.1) =-0.2

$$\hat{\mu}_{2.} = \hat{\mu}_{..} + \hat{\alpha}_2 = \hat{\mu}_{..} - \hat{\alpha}_1 = 1.7 - (-0.1) = 1.8$$

$\hat{\sigma}(\hat{\mu}_{1.} - \hat{\mu}_{2.}) = 2\hat{\sigma}(\hat{\alpha}_1)$=2*(0.116) =0.23

We plug in CI formula CI:$\hat{\theta}$ +/- ***t*(0.975, 8)*SE**

2. In a three-factor study, the effects of gender of subject (factor A), body fat of subject (measured in percent, factor B), and smoking history of subject (factor C) on exercise tolerance (Y) were studied in 24 participants. Each factor has 2 level and data is balanced.

	Smoking History	
	$k=1$ Light	$k=2$ Heavy
$j=1$ Low fat:		
$i=1$ Male	24.1 (Y_{1111})	17.6 (Y_{1121})
	29.2 (Y_{1112})	18.8 (Y_{1122})
	24.6 (Y_{1113})	23.2 (Y_{1123})
$i=2$ Female	20.0 (Y_{2111})	14.8 (Y_{2121})
	21.9 (Y_{2112})	10.3 (Y_{2122})
	17.6 (Y_{2113})	11.3 (Y_{2123})
$j=2$ High fat:		
$i=1$ Male	14.6 (Y_{1211})	14.9 (Y_{1221})
	15.3 (Y_{1212})	20.4 (Y_{1222})
	12.3 (Y_{1213})	12.8 (Y_{1223})
$i=2$ Female	16.1 (Y_{2211})	10.1 (Y_{2221})
	9.3 (Y_{2212})	14.4 (Y_{2222})
	10.8 (Y_{2213})	6.1 (Y_{2223})

a) (10 points) Please write down the full 3-factor ANOVA factor effects models with all interactions terms.

____

b) (30 points) The ANOVA table can be calculated as follows:

```
                        Df    SS      MS     F value   P-value
Gender                   1 176.58  176.58  18.915 0.000497 ***
Body_Fat                 1 242.57  242.57  25.984 0.000108 ***
Smoking                  1  70.38   70.38   7.539 0.014357 *
Gender:Body_Fat          1  13.65   13.65   1.462 0.244143
Gender:Smoking           1  11.07   11.07   1.186 0.292299
Body_Fat:Smoking         1  72.45   72.45   7.761 0.013221 *
Gender:Body_Fat:Smoking  1   1.87    1.87   0.200 0.660434
Residuals               16 149.37    9.34
```

Please interpret the result using the ANOVA table (which effect is significant or not):

1) Three-way interactions:____ p=0.66 , not significant____________________________
2) Two-way interactions: three 2-way interactions, only the interaction between body fat and smoking was significant, p=0.013; the other two were not.
3) Main effects: All the main effect terms were significant. But because of significant interactions between body fat and smoking, we should look at the treatment levels for body fat and smoking combined effect on response .