STAT 516 Fall 2004 Quiz 4-6 Answers

QUIZ 4

Questions 1-4 refer to the one-way ANOVA and data presented below.

1) On the list of the data and means, indicate the following values: y_{11}, y_{23}, $\bar{y}_{3\bullet}$, and $\bar{y}_{\bullet\bullet}$.

	Variety A	Variety B	Variety C	Variety D	
	934 = y_{11}	880	987	992	
	1041	963	951	1143	
	1028	**924** = y_{23}	976	1140	
	935	946	840	1191	
means	984.50	928.25	**938.50** = $\bar{y}_{3\bullet}$	1116.50	**991.9375** = $\bar{y}_{\bullet\bullet}$

2) Complete the above ANOVA table.

The GLM Procedure

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	**4-1=3**	89931.1875	29977.0625	**MSB/MSE=7.21**	0.0050
Error	**15-3=12**	**TSS-SSB=49875.75**	**SSE/df=4156.313**		
Corrected Total	15	139806.9375			

Brown and Forsythe's Test for Homogeneity of yield Variance
ANOVA of Absolute Deviations from Group Medians

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
variety	3	1538.2	512.7	0.25	0.8607
Error	12	24740.8	2061.7		

3) At α=0.05 do we believe the variety means are equal? or that at least one is different? **We reject the null hypothesis in the ANOVA table that they are all equal as the p-value of 0.0050 is less than α=0.05. It appears that at least one is different.**

4) Which ANOVA assumption is being tested by the p-value of 0.8607? **Does the assumption appear to be met? The modified Levene's test tests the null hypothesis that the variances of the errors are equal in each group. We fail to reject the null hypothesis that they are equal, and so we accept that they are.**

5) Say that five tests of hypotheses result in the p-values of 0.043, 0.002, 0.009, 0.54, and 0.37. Using α_T=0.05 and Bonferroni's method, circle the p-values (if any) that would result in a rejection. **0.002 < 0.05/5=0.01 so reject that one. 0.009<0.05/4=0.0125 so reject that one. 0.043 is not < 0.05/3=0.0167 and so we fail to reject that and the others.**

QUIZ 5

Questions 1 through 4 deal with the following scenario. A one-way ANOVA is performed to determine the difference in cleaning ability of four leading detergents and a generic detergent. The five types of detergent are entered in SAS in the following order: *Brand A*, *Ultra-Brand A*, *Brand B*, *Ultra-Brand B*, and *Megalo-Clean*. The average cleaning power for each detergent could be represented by the parameters: μ_A, μ_{UA}, μ_B, μ_{UB}, and μ_{MC} respectively, where a higher value is better.

```
DATA detergent;
INPUT brand $ power @@
CARDS;
A      3.1    A      3.6    A      2.7
UA     3.0    UA     3.2    UA     4.0
B      3.8    B      4.2    B      3.9
UB     4.0    UB     3.9    UB     3.7
MC     2.8    MC     3.4    MC     2.8
;
```

```
PROC GLM DATA=detergent ORDER=DATA;
CLASS brand;
MODEL power=brand;
CONTRAST 'c1' brand  0.5   0.5   -0.5   -0.5  0;
CONTRAST 'c2' brand  0.25 0.25   0.25   0.25 -1;
CONTRAST 'c3' brand  0.5 -0.5    0.5   -0.5  0;
RUN;
```

and the output for testing the null hypotheses that the three contrasts were equal to zero is:

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
c1	1	1.26750000	1.26750000	9.46	0.0117
c2	1	0.84016667	0.84016667	6.27	0.0312
c3	1	0.02083333	0.02083333	0.16	0.7016

1) Using the Holm procedure and a family wise error rate of α_T=0.05, which contrasts do we reject the null hypotheses for? (Show your work!) **Putting the p-values in order we compare `0.0117` to α/3=0.0167 and reject the null hypothesis. Compare 0.0312 to α/2=0.025 and fail to reject that null hypothesis and the following one.**

2) In terms of either the problem description or the parameters, describe what comparison is being made in contrast 1. **Contrast one is comparing brand A on average (ultra and non-ultra) to brand B on average (ultra and non-ultra).** $\frac{\mu_A + \mu_{UA}}{2} - \frac{\mu_B + \mu_{UB}}{2}$

3) Give the values you would need to enter it to SAS for the contrast $\frac{\mu_A + \mu_B}{2} - \mu_{MC}$. **0.5 0 0.5 0 -1**

4) Demonstrate whether or not contrast 2 and contrast 3 are orthogonal or not.

```
0.25 0.25  0.25  0.25 -1 = 0.125 - 0.125 +0.125 - 0.125 -0 = 0
0.5 -0.5   0.5  -0.5  0
```

so they are orthogonal

Questions 5 and 6 concern the complete version of the concrete data we saw in class.

Aggregate Type	Static	Regular	Low	Very Low
	Compaction Method			
Basalt	68	126	93	56
	63	128	101	59
	65	133	98	57
Silicious	71	107	63	40
	66	110	60	41
	66	116	59	44

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Stat	Pr > F
Model	7	19122.5000	2731.7857	287.56	<.0001
Error	16	152.0000	9.5000		
C Total	23	19274.5000			

Type III Tests

Source	DF	Sum of Squares	Mean Square	F Stat	Pr > F
agg	1	1734.0000	1734.0000	182.53	<.0001
comp	3	16243.5000	5414.5000	569.95	<.0001
agg*comp					<.0001

5) The row for the Type III tests has been deleted. Provide the four missing values.
The df and SS in the Type III table must add up to the model line in the ANOVA table. So the df are 7-1-3=3 and the SS are 19122.5-1734.0-16243.5 = 1145.0 The MS=SS/df so 1145.0/3=381.67. The F is the MSagg*comp/MSerror=381.67/9.5=40.18.

6) Briefly say why this design is Balanced. **Each combination of factor levels has the same number of observations (3).**

QUIZ 6

Questions 1 and 2 deal with the same scenario as questions 1-4 on Quiz 5.

1) In terms of either the problem description or the parameters, describe what comparison is being made in the contrast: `0 1 0 -1 0`.
This contrast compares the two types of ultra product. μ_{UA}-μ_{UB}.

2) Give the values you would need to enter it to SAS for the contrast $\frac{\mu_A + \mu_B + \mu_{MC}}{3} - \frac{\mu_{UA} + \mu_{UB}}{2}$. Remember that SAS won't take a 1/3rd and so you will need to specify a divisor. **We want `+1/3 -1/2 +1/3 -1/2 +1/3`, so `2 -3 2 -3 +2 / divisor=6`**

Questions 3 and 4 concern the complete version of the concrete data as used on Quiz 5, except with different Type III values missing.

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Stat	Pr > F
Model	7	19122.5000	2731.7857	287.56	<.0001
Error	16	152.0000	9.5000		
C Total	23	19274.5000			

Type III Tests

Source	DF	Sum of Squares	Mean Square	F Stat	Pr > F
agg					<.0001
comp	3	16243.5000	5414.5000	569.95	<.0001
agg*comp	3	1145.0000	381.6667	40.18	<.0001

3) The row for `agg` in the Type III tests has been deleted. Provide the four missing values.
The df and SS in the Type III table must add up to the model line in the ANOVA table. So the df are 7-3-3=1 and the SS are 19122.5-16243.5-1145.0 = 1734.0 The MS=SS/df so 1734/1=1734. The F is the MS_{agg}/MS_{error}=1734/9.5=182.53.

4) Briefly say why this design has replications. **Each combination of factor levels has more than one observation.**

Questions 5 and 6 deal with the following scenario. Seven different types of material (factor γ, labeled A-F) were sent out to a sample of 13 laboratories (factor α, labeled 1-13) for stress testing (since different laboratories use different testing methods). The `PROC GLM` code used was:

```
PROC GLM;
CLASS lab material;
MODEL stress = lab material lab*material;
RANDOM lab lab*material;
RUN;
```

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	90	322913.2482	3587.9250	177.01	<.0001
Error	273	5533.5800	20.2695		
Corrected Total	363	328446.8282			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
lab	12	30328.0547	2527.3379	124.69	<.0001
material	6	268778.0771	44796.3462	2210.03	<.0001
lab*material	72	23807.1165	330.6544	16.31	<.0001

Source	Type III Expected Mean Square
lab	Var(Error) + 4 Var(lab*material) + 28 Var(lab)
material	Var(Error) + 4 Var(lab*material) + Q(material)
lab*material	Var(Error) + 4 Var(lab*material)

5) In terms of the σ^2_{lab}, $\gamma_{matA}=\gamma_{matB}=\ldots=\gamma_{matF}$, and/or $\sigma^2_{lab*material}$, what null hypothesis is the `lab` line in the Type III table (the F-value of 124.69) testing?

The F-value on the lab line is the ratio of the lab MS and the error MS, so looking at:

```
lab          Var(Error) + 4 Var(lab*material) + 28 Var(lab)
error        Var(Error)
```

we see that it must be testing the null hypothesis that both the σ^2_{lab}=0 and $\sigma^2_{lab*material}$=0.

6) Find the value of the F statistic for testing that σ^2_{lab} =0 against σ^2_{lab} >0.

Looking at the E(MSlab) line we see that we need to cancel out both the error and interaction term to have only the lab term remaining. So we need to use the lab*material MS as the denominator (check the expected mean-squares!). So the F is 124.69/16.31=7.65 and the degrees of freedom are12 and 72.