ST512

Fall Semester, 2006

Quiz 4

Name:____

Directions: Answer questions as directed. Please show work. You do not need to carry out any arithmetic that is obvious for the grader. For example, if the correct answer is $62.1 + 2.6 + 28.1 + 30.3$, then you needn't carry out the addition to get 123.1 unless you'll need it for some subsequent calculation.

There are 5 problems (roughly 20 points each) on 4 pages and the quiz. After the fouproblems, there are three pages of SAS output pertaining to problems 1, 2 and 4 ("BARLEY PROBLEM", "WHEATYIELD PROBLEM" and "COOKIES PROBLEM") and then one page for a t-table.

1. Recall the study investigating variability of blood pressure measurements among $n = 25$ randomly sampled males. Each subject made 16 visits to a clinic and systolic blood pressures were measured on each visit, for a total of 400 measurements. Consider a one-factor random effects model for this study.

   (a) The variance component estimates for subject (σ_S^2) and visit (σ^2) below were obtained by equating observed mean squares to expected mean squares.

$$\begin{aligned} \hat{\sigma}_S^2 &= 198.7 \\ \hat{\sigma}^2 &= 184.5 \end{aligned}$$

   Use them to complete the ANOVA table below:

Source	*df*	Sum of Squares	Mean Square	Expected Mean square
Subjects	24	80728.8	3363.7	$\sigma^2 + 16\sigma_S^2$
Error	375	69187.5	184.5	σ^2
Total	399	149916		

   (b) The sample mean systolic blood pressure among all measurements was

$$\bar{y}_{++} = \frac{1}{400} \sum_{i=1}^{25} \sum_{j=1}^{16} y_{ij} = 149.2$$

   Report a 95% confidence interval for the mean systolic blood pressure in the population from which this sample was taken.

$$\bar{y}_{++} \pm t(.025, 24)\sqrt{\frac{1}{400} MS(Subj)}$$

   or

$$149.2 \pm 2.06\sqrt{\frac{1}{400}(3363.7)}$$

   or

$$149.2 \pm 6$$

   (c) Report the F-ratio and associated degrees of freedom for a test that the variance component for patient is 0: $H_0 : \sigma_S^2 = 0$.

$$F = \frac{MS(Subj)}{MS(E)} = \frac{3363.7}{184.5}, \quad df = 24, 375$$

2. An experiment randomized five barley varieties to six locations so that each variety was observed once in each location. Pertinent SAS code and output are given at the end of the exam under the title "BARLEY PROBLEM." Treat variety effects as fixed.

   (a) Precede each of the following True/False questions with "Regardless of whether locations are regarded as fixed or random, ..." then circle True or False:

      i. the F-test of equality of variety means is the same. (True)

      ii. the standard error for any contrast among sample variety means is the same. (True)

      iii. the standard error for any single sample variety mean is the same. (False) Prove or disprove this claim (iii).

$$\bar{y}_{i+} = \begin{cases} \mu + \alpha_i + \bar{E}_{i+} & \text{fixed} \\ \mu + \alpha_i + \bar{B}_{+} + \bar{E}_{i+} & \text{random} \end{cases}$$

$$SE(\bar{y}_{i+}) = \begin{cases} \sqrt{\frac{\sigma^2}{6}} & \text{fixed} \\ \sqrt{\frac{1}{6}(\sigma_B^2 + \sigma^2)} & \text{random} \end{cases}$$

   (b) For the following two problems, assume random location effects:

      i. Consider the observed difference between M and T, $\bar{y}_{4+} - \bar{y}_{1+}$. Obtain a 95% confidence interval for the parameter estimated by this difference.

$$127.4 - 102.6 \pm t(.025, 20)\sqrt{\frac{2}{6}MS(E)}$$

or

$$24.8 \pm 15.4$$

      ii. Consider the average of the five variety (population) means. Propose and report an estimate (and standard error) for this quantity.

$$\bar{Y}_{++} = \mu + \bar{B}_{+} + \bar{E}_{++}$$

$$\bar{y}_{++} = 109$$

$$SE(\bar{Y}_{++}) = \sqrt{\frac{1}{30}(5\sigma_B^2 + \sigma^2}$$

$$\widehat{SE}(\bar{Y}_{++}) = \sqrt{\frac{1}{30}(MS(loc)} = 2.3$$

3. An experiment with wheat investigates the effects of irrigation (A) and fertilizer (B). The design first randomized 10 fields to the two irrigation methods (furrow and sprinkler) and then splits each field into three subplots which are randomized to the three fertilizer rates ('1x','2x' and '3x'). The data can be found in the SAS code and output entitled "WHEAT PROBLEM."

   (a) Sketch a diagram with fields, numbered 1-10. Each field should be divided into subplots as well. Indicate treatment combinations for the fields. Indicate on your diagram where the three measurements, y_{215}, y_{225} and y_{235} were observed. (i, j, k denote irrigation method, fertilizer method and field, respectively.)

   diagram omitted

   (b) Write out an appropriate statistical model for data from this experiment which allows for fixed main effects for A (α_i) and B (β_j) as well as interaction ($(\alpha\beta)_{ij}$) and also a random effect for field. You don't need to specify the sum-to-zero constraints, but you do need to identify variance components.

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + S_{k(i)} + E_{ijk}$$

   Here, $S_{k(i)} \overset{iid}{\sim} N(0, \sigma_s^2)$ and $E_{ijk} \overset{iid}{\sim} N(0, \sigma^2)$, with S and E independent.

   (c) Note that $SS[fert] = 11.3$. Use it to construct the F-ratio for a test of no fertilizer effect. Also, partition $SS[Irr]$ into linear and quadratic orthogonal components.

$$F_{irr} = \frac{MS(fert)}{MS(E))} = \frac{11.3}{1} = 11.3$$

$$\begin{aligned} \theta_L &= 59.1 - 58.4 \\ SS(\theta_L) &= \frac{0.7^2}{2/10} \\ &= 2.5 \\ \theta_Q &= 59.1 - 2(59.9) + 58.4 \\ &= 2.3 \\ SS(\theta_Q) &= \frac{2.3^2}{6/10} \\ &= 8.8 \end{aligned}$$

(d) Complete the table of main estimated effects below. Beneath the table, provide expressions for each standard error.

Effect	Contrast	Estimate	standard error	confidence interval
Main effect of irrigation	$\alpha_2 - \alpha_1$			
		16.8	3.2	$16.8 \pm 2.31(3.2)$
				16.8 ± 7.5
Difference between	$\beta_2 - \beta_1$	1.5	0.44	$1.5 \pm 2.12(0.44)$
Fertilizers $1x$ and $2x$				$1.5 \pm .94$

$$
\begin{aligned}
\bar{y}_{2++} - \bar{y}_{1++} &= \alpha_2 - \alpha_1 + \bar{S}_{+(2)} - \bar{S}_{+(1)} + \bar{E}_{2++} - \bar{E}_{1++} \\
SE(\bar{y}_{2++} - \bar{y}_{1++} &= \sqrt{\frac{\sigma_s^2}{5} + \frac{\sigma^2}{15}} \\
&= \sqrt{\frac{2}{15}(3\sigma_s^2 + \sigma^2)} \\
\widehat{SE}(\bar{y}_{2++} - \bar{y}_{1++} &= \sqrt{\frac{2}{15} MS(field(irr))} \\
&= 3.2(df = 8) \\
\bar{y}_{+2+} - \bar{y}_{+1+} &= \beta_2 - \beta_1 + \bar{E}_{+2+} - \bar{E}_{+1+} \\
\widehat{SE}(\bar{y}_{+2+} - \bar{y}_{+1+} &= \sqrt{\frac{2}{10} MS(E)} \\
&= .44(df = 16)
\end{aligned}
$$

4. An experiment measures the effects of four row spacings $(1x, 1.2x, 1.4x, 1.6x)$ and two fertilizer treatments $(f1, f2)$ and their interaction on grain yield. While it is possible to divide a field into two sections and randomize the sections to the fertilizer treatments, it is not easy to do this with row spacings; the spacing must be the same for the entire field. Each of three participating farmers randomizes four of their fields to the four spacings, then divides the fields in two and randomizes the fertilizer treatments. SAS code and output may be found under the title "SPACING PROBLEM."

   (a) Report an F-ratio for a test for no interaction between spacing and fertilizer.

$$F = \frac{2.46/3}{2.25} = 0.36$$

   (b) Report a 95% confidence intervals for

      i. the main effect of the fertilizer treatment on yield.

      ii. the mean difference in yield between the two extreme spacing treatments $1.0x$ and $1.6x$ after averaging over fertilizer.

$$\begin{aligned}
\hat{\theta}_f &= \bar{y}_{+2+} - \bar{y}_{+3+} \\
&= 27.50 - 27.42 \\
&= 0.08 \\
SE(\hat{\theta}_f) &= \sqrt{\frac{2}{12} MS(E)} (df = 8) \\
&= 0.61 \\
t(.025, 8) &= 2.31
\end{aligned}$$

95% confidence interval is $0.08 \pm 2.31(0.61)$ or $.08 \pm 1.41$

$$\begin{aligned}
\hat{\theta}_2 &= \bar{y}_{1++} - \bar{y}_{4++} \\
&= 31.2 - 23 \\
&= 8.2 \\
SE(\hat{\theta}_2) &= \sqrt{\frac{2}{5} MS(spacing \times farmer)(df = 6)} \\
&= 1.45 \\
t(.025, 6) &= 2.45
\end{aligned}$$

95% confidence interval is $8.2 \pm 2.45(1.45)$ or 8.2 ± 3.55

BARLEY PROBLEM

```
proc glm data=one;
   class loc variety;
   model yield =loc variety;
   means variety;
run;
```

The SAS System 1

The GLM Procedure

Class	Levels	Values
loc	6	C D GR M UF W
variety	5	M P S T V

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
loc	5	17829.84667	3565.96933	21.89	<.0001
variety	4	2756.62467	689.15617	4.23	0.0121
Error	20	3257.74333	162.88717		
Corrected Total	29	23844.21467			

R-Square	Coeff Var	Root MSE	yield Mean
0.863374	11.70391	12.76273	109.0467

Level of variety	N	------------yield------------ Mean	Std Dev
M	6	102.583333	25.9640842
P	6	109.750000	21.3401734
S	6	102.033333	26.0269604
T	6	127.400000	36.6710785
V	6	103.466667	32.6467558

SPACING PROBLEM

```
proc glm;
   class spacing fertilizer farmer;
   model yield=spacing|fertilizer farmer farmer*spacing;
   lsmeans spacing*fertilizer;
run;
```

The SAS System
The GLM Procedure

Class Level Information

Class	Levels	Values
spacing	4	1.2x 1.4x 1.6x 1x
fertilizer	2	f1 f2
farmer	3	1 2 3

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	__	2767.958333	__________	82.01	<.0001
Error	__	18.000000	________		
Corrected Total	23	2785.958333			

R-Square	Coeff Var	Root MSE	yield Mean
0.993539	5.462822	1.500000	27.45833

Source	DF	Type III SS	Mean Square	F Value	Pr > F
spacing	_	260.791667			
fertilizer	_	0.041667			
spacing*fertilizer	_	2.458333			
farmer	_	2466.583333			
spacing*farmer	_	38.083333			

Least Squares Means

spacing	fertilizer	yield LSMEAN
1.2x	f1	27.6666667
1.2x	f2	28.6666667
1.4x	f1	26.3333333
1.4x	f2	26.6666667
1.6x	f1	23.3333333
1.6x	f2	22.6666667
1x	f1	32.3333333
1x	f2	32.0000000

WHEAT PROBLEM

The Mixed Procedure

Class	Levels	Values
field	5	1 2 3 4 5
irr	2	furrow sprinkler
fert	3	1x 2x 3x

Type 3 Analysis of Variance

Source	DF	Sum of Squares	Mean Square
irr	1	2116.800000	2116.800000
fert	2	11.266667	5.633333
irr*fert	2	0.200000	0.100000
field(irr)	8	627.333333	78.416667
Residual	16	15.866667	0.991667

Covariance Parameter Estimates

Cov Parm	Estimate
field(irr)	25.8083
Residual	0.9917

Least Squares Means

Effect	irr	fert	Estimate	Standard Error	DF	t Value	Pr > |t|
irr	furrow		50.7333	2.2864	8	22.19	<.0001
irr	sprinkler		67.5333	2.2864	8	29.54	<.0001
fert		1x	58.4000	1.6371	16	35.67	<.0001
fert		2x	59.9000	1.6371	16	36.59	<.0001
fert		3x	59.1000	1.6371	16	36.10	<.0001

Table of critical values ($t(df, \alpha)$) from t-distributions:

df	$\alpha = 0.2$	$\alpha = 0.15$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.025$	$\alpha = 0.01$	$\alpha = 0.005$	$\alpha = 0.001$
1	1.37638	1.96261	3.07768	6.31375	12.7062	31.8205	63.6567	318.309
2	1.06066	1.38621	1.88562	2.91999	4.3027	6.9646	9.9248	22.327
3	0.97847	1.24978	1.63774	2.35336	3.1824	4.5407	5.8409	10.215
4	0.94096	1.18957	1.53321	2.13185	2.7764	3.7469	4.6041	7.173
5	0.91954	1.15577	1.47588	2.01505	2.5706	3.3649	4.0321	5.893
6	0.90570	1.13416	1.43976	1.94318	2.4469	3.1427	3.7074	5.208
7	0.89603	1.11916	1.41492	1.89458	2.3646	2.9980	3.4995	4.785
8	0.88889	1.10815	1.39682	1.85955	2.3060	2.8965	3.3554	4.501
9	0.88340	1.09972	1.38303	1.83311	2.2622	2.8214	3.2498	4.297
10	0.87906	1.09306	1.37218	1.81246	2.2281	2.7638	3.1693	4.144
11	0.87553	1.08767	1.36343	1.79588	2.2010	2.7181	3.1058	4.025
12	0.87261	1.08321	1.35622	1.78229	2.1788	2.6810	3.0545	3.930
13	0.87015	1.07947	1.35017	1.77093	2.1604	2.6503	3.0123	3.852
14	0.86805	1.07628	1.34503	1.76131	2.1448	2.6245	2.9768	3.787
15	0.86624	1.07353	1.34061	1.75305	2.1314	2.6025	2.9467	3.733
16	0.86467	1.07114	1.33676	1.74588	2.1199	2.5835	2.9208	3.686
17	0.86328	1.06903	1.33338	1.73961	2.1098	2.5669	2.8982	3.646
18	0.86205	1.06717	1.33039	1.73406	2.1009	2.5524	2.8784	3.610
19	0.86095	1.06551	1.32773	1.72913	2.0930	2.5395	2.8609	3.579
20	0.85996	1.06402	1.32534	1.72472	2.0860	2.5280	2.8453	3.552
21	0.85907	1.06267	1.32319	1.72074	2.0796	2.5176	2.8314	3.527
22	0.85827	1.06145	1.32124	1.71714	2.0739	2.5083	2.8188	3.505
23	0.85753	1.06034	1.31946	1.71387	2.0687	2.4999	2.8073	3.485
24	0.85686	1.05932	1.31784	1.71088	2.0639	2.4922	2.7969	3.467
25	0.85624	1.05838	1.31635	1.70814	2.0595	2.4851	2.7874	3.450
26	0.85567	1.05752	1.31497	1.70562	2.0555	2.4786	2.7787	3.435
27	0.85514	1.05673	1.31370	1.70329	2.0518	2.4727	2.7707	3.421
28	0.85465	1.05599	1.31253	1.70113	2.0484	2.4671	2.7633	3.408
29	0.85419	1.05530	1.31143	1.69913	2.0452	2.4620	2.7564	3.396
30	0.85377	1.05466	1.31042	1.69726	2.0423	2.4573	2.7500	3.385
40	0.85070	1.05005	1.30308	1.68385	2.0211	2.4233	2.7045	3.307
60	0.84765	1.04547	1.29582	1.67065	2.0003	2.3901	2.6603	3.232
80	0.84614	1.04320	1.29222	1.66412	1.9901	2.3739	2.6387	3.195
100	0.84523	1.04184	1.29007	1.66023	1.9840	2.3642	2.6259	3.174
200	0.84342	1.03913	1.28580	1.65251	1.9719	2.3451	2.6006	3.131
100000	0.84162	1.03644	1.28156	1.64487	1.9600	2.3264	2.5759	3.090