ST512 - Quiz 2
Fall Semester, 2007 - solution

1. An experiment randomized $N = 25$ turkeys to 5 diets, correspondingk to 5 amounts of supplementary methionine, as a percentage of diet: (0,.04,.08,.12,.16) The treatment and error mean squares from a one-way ANOVA are $MS[Trt] = 115034$ and $MS[E] = 302$, respectively. (The response is weightgain.) The linear orthogonal polynomial contrast is given by

$$\hat{\theta}_1 = -2\bar{y}_{1+} - \bar{y}_{2+} + 0\bar{y}_{3+} + \bar{y}_{4+} + 2\bar{y}_{5+} = 734.4$$

   (a) Find the sum of squares for the estimated contrast $\hat{\theta}_1$.

$$SS(\theta) = \frac{\hat{\theta}^2}{\sum c_i^2/n} = \frac{734.4^2}{10/5} = 269672$$

   (b) Find the standard error of the estimated contrast $\hat{\theta}_1$.

$$SE(\hat{\theta}) = \sqrt{MS(E)\sum c_i^2/n} = \sqrt{302(10/5)} = 24.6$$

   (c) Give the F-ratio for a lack-of-fit test of the simple linear regression model, along with the corresponding degrees of freedom. ($df = 3, 20$)

$$\begin{aligned} F_{LOF} &= \frac{SS(LOF)/(4-1)}{MS(E)} \\ &= \frac{(SS(Trt) - SS(SLR))/3}{MS(E)} \\ &= \frac{(4 * 115034 - 269672)/3}{302} = 210.2 \end{aligned}$$

2. True or false: circle one

Consider a one-way analysis of variance with $t = 5$ treatments each observed on n experimental units.

~~True~~	(False)	(i) For multiple contrasts among the means, the comparisonwise error rate is defined as the probability of at least one type I error
(True)	~~False~~	(ii) For tests derived from multiple confidence intervals with simultaneous confidence level .95, the probability of no type I errors is at least .95.
(True)	~~False~~	(iii) For the family of all pairwise comparisons of means, Tukey's procedure has a lower type II error rate than the Bonferroni procedure (with $k = t(t-1)/2 = 10$).
(True)	~~False~~	(iv) A design like this in which an equal number of units is randomized to each treatment is called balanced

3. An experiment randomly assigns 12 pigs to 6 treatments. Initial weights (z) are measured, along with the response of interest, growth rate (`grate`). The SAS output on the next page, entitled "PIGS PROBLEM", may be used in an analysis of covariance.

   (a) Report the difference between the adjusted means for treatments "cm" and "cf", along with a standard error. $\hat{\beta}_{cf} = 0.93, SE = 0.25$

   (b) Report the adjusted mean growth rates for two treatments: "bf" and "bm".

$$\begin{aligned} \hat{\beta}_0 + \hat{\beta}_{bf} + \hat{\beta}_z(\bar{z}) &= 6.77 + 0.74 + 0.048(39.2) \\ &= 9.4 \\ \hat{\beta}_0 + \hat{\beta}_{bm} + \hat{\beta}_z(\bar{z}) &= 6.77 - 0.18 + 0.048(39.2) \\ &= 8.5 \end{aligned}$$

   (c) Which estimate from part (b) has a larger standard error? "bm" because it makes a bigger "adjustment" (since $|37 - \bar{z}| > |38 - \bar{z}|$)

   (d) Which is larger, the adjusted mean for treatment "cm", or the unadjusted mean? The adjusted mean adjusts back to $\bar{z}$ and is therefore less than $\bar{y}_6$

   (e) Obtain the type III sum of squares for initial weight, z. If F denotes the F-ratio for a test of $\beta_z = 0$,

$$SS(\beta_z) = F \times MS(E) = t^2 MS(E) = 1.2$$

```
proc glm data=pigs;
   class treat;
   model grate=treat;
   means treat;
```


The GLM Procedure

Class	Levels	Values
treat	6	af am bf bm cf cm

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	4.16234076	0.69372346	11.56	0.0084
Error	5	0.30015091	0.06003018		
Corrected Total	11	4.46249167			

R-Square	Coeff Var	Root MSE	grate Mean
0.932739	2.651869	0.245011	9.239167

Source	DF	Type III SS	Mean Square	F Value	Pr > F
treat	5	2.97601155	0.59520231	9.92	0.0125
z	1	xxxxxxxxxx	xxxxxxxxxx	xxxx	0.0065

Parameter		Estimate		Standard Error	t Value	Pr > |t|
Intercept		6.769942803	B	0.48656385	13.91	<.0001
treat	af	0.904709247	B	0.24646604	3.67	0.0144
treat	am	1.145476644	B	0.24969549	4.59	0.0059
treat	bf	0.740476644	B	0.24969549	2.97	0.0313
treat	bm	-0.181639657	B	0.25197690	-0.72	0.5033
treat	cf	0.929709247	B	0.24646604	3.77	0.0130
treat	cm	0.000000000	B	.	.	.
z		0.047883699		0.01069823	4.48	0.0065

NOTE: The X'X matrix has been found to be singular, and a generalized inverse was used to solve the normal equations. Terms whose estimates are followed by the letter 'B' are not uniquely estimable.

Level of treat	N	grate Mean	grate Std Dev	z Mean	z Std Dev
af	2	9.59000000	0.49497475	40.0000000	11.3137085
am	2	9.73500000	0.30405592	38.0000000	0.0000000
bf	2	9.33000000	0.94752309	38.0000000	14.1421356
bm	2	8.36000000	0.21213203	37.0000000	2.8284271
cf	2	9.61500000	0.19091883	40.0000000	11.3137085
cm	2	xxxxxxxxxx	0.43133514	42.5000000	7.7781746

4. An experiment is conducted to compare the accuracy of two rifles; one is a model called "varmint" and the other is called "fireball." The two rifles are mounted and stationary and a drawstring is used to pull the trigger. Each is fired in five sessions (at different geographical locations) under each of two operating conditions, ("clean" and "fouled") for a total of 20 sessions. For each session, accuracy is measured as average distance from target (ADT), in hundredths of an inch, with the results below.

```
Level of      Level of              -------------ADT-------------
gunmodel      condition     N               Mean          Std Dev

Fireball      clean         5         110.000000       14.1421356
Fireball      fouled        5         164.000000       15.1657509
varmint       clean         5          74.000000       11.4017543
varmint       fouled        5         124.000000       11.4017543
```

The variance between the four treatment means is 1384. The average of the four within-treatment-group variances is $MS[E] = 172$.

(a) Obtain an ANOVA table for a two-factor analysis that partitions the total sum of squares into four sources of variability. Include a column for SS, MS and df.

Source		df		SS		MS
treatments		3		20760		
	gunmodel		1		7220	7220
	condition		1		13520	13520
	interaction		1		20	20
error		16		2752		172
total		19		23512		

(b) Report either a t-statistic or an F-ratio for a test of the hypothesis of no interaction between `gunmodel` and `condition`. Give the df for the statistic.

$$F = \frac{20}{172} = 0.12 (df = 1, 16)$$

(c) Conduct a test that there is no difference in the mean accuracy between these two guns, averaging over condition.

$$F = \frac{7220}{172} = 41.9 (df = 1, 16)$$

(This exceeds the critical value $F(.05, 1, 16) = 4.49$ given below, so that the observed difference is statistically signficant, and there appears to be a difference in accuracy between the two rifles/models.)

(d) Sketch an interaction plot and briefly characterize the effects of `gunmodel` and `condition`. (If you want to declare significance, you may use the critical value $F(.05, 1, 16) = 4.49$.) An interaction plot reveals parallel lines, with additive effects of `gunmodel` and `condition` that are significant. The effect of the gun being dirty is more pronounced than the superior accuracy of the rifle that is of the "varmint" model.