Biology 4605/7220 Name_____
8 November 2007 Exam #2b

For each of the following situations (1 and 2):

nv = number of variables
nt = number of terms
A. score = 3nv
B. score = nt
C. score = 2nv + 2
D. score = 1

(A) Define variables in a tabular format, as follows.

name	symbol	scale

scale = nominal, ordinal, or cardinal
cardinal = interval or ratio scale.

(B) Using the symbols, write a general linear model relating the response variable to explanatory variable(s) and interaction terms (if appropriate).

(C) Complete the first two columns of the ANOVA table

source	df

(D) State the name of the analysis, from the following list.
t-test, one-way ANOVA, two-way ANOVA, three-way ANOVA
paired comparisons, randomized blocks,
hierarchical (nested) ANOVA
regression, multiple regression,
ANCOVA (at least 1 nominal and at least 1 cardinal scale explanatory variable)
none of the above.

1. Huntsberger (1967, *Elements of Statistical Inference*, p277) reported income per acre (1956 through 1965) produced by a commercial system for taking game animals. What was the rate of growth in income?

A=6 B=2 C=6 D=1

A.

name	symbol	scale
Income	I	Cardinal
Time	T	Cardinal

C.

source	df
T	1
error	8
total	9

B. $I = \beta_0 + \beta_T T + \epsilon$

D. Regression

2. Huntsberger (1967 p324) reported yields in kilograms per plot that resulted when four equally spaced levels of nitrogen N_o, N_4, N_8, N_{12} were applied to a variety of grain. Each treatment was applied to its own plot in each of four areas that differed in levels of irrigation.

A=9 B=4 C=8 D=1

A.

name	symbol	scale
Yield	Y	Cardinal
Nitrogen	N	Nominal
Area	A	Nominal

C.

source	df
Nitrogen	3
Area	3
Error	9

B. $Y = \beta_0 + \beta_N N + \beta_A A + \epsilon$

D. Randomized block (two way ANOVA)

3. (From Huntsberger, 1967, Problem 8, p 173) Given *n* random samples of the variable *Y*, which is distributed normally, the quantity *t* will follow a t-distribution with *n* – 1 degrees of freedom. The quantity F will follow the F-distribution, where $F = t^2$

$$t = \frac{\sqrt{n}\left(\overline{Y} - \mu\right)}{s}$$

Find *t*, given n =50, s = 3, $\overline{Y} - \mu = 6$ 14.142 [2]

Find s^2, given n =50, $\overline{Y} - \mu = 6$, t = 2.312 336.742 [2]

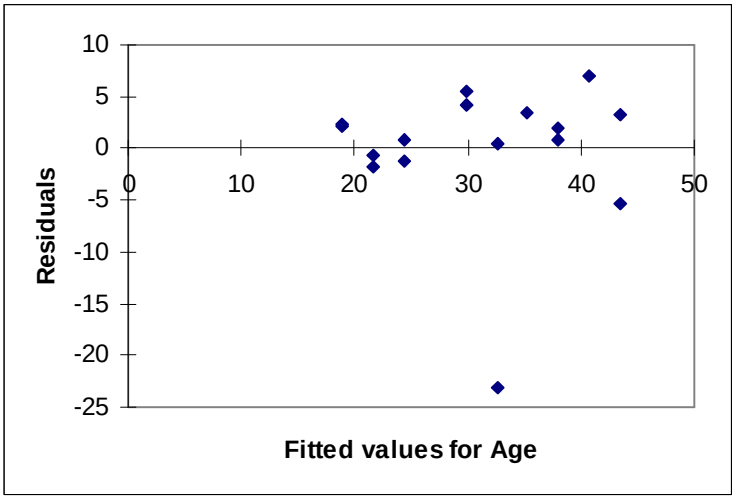
Find *F*, given n =37, s = 3, $\overline{Y} - \mu = 6$ 148 [2]

4. Huntsberger (1967 p225) reported the percent fat found in samples of two types of meat (10 samples per type). Complete the ANOVA table. [7]

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*
Types	105.8	1	105.8	3.3922	0.082
Within Types (error)	_561.4	_18	31.189		
Total	667.2	_19			

5. Huntsberger (1967 p281) reported on heights of trees in relation to age.
Here are the fitted and residual values for the regression of tree height on age.

Predicted	*Residuals*
32.56638	0.433619
29.85017	4.249828
21.70154	-0.701544
32.56638	-23.06638
29.85017	5.549828
24.41775	0.882247
18.98533	2.414666
40.71501	6.984991
21.70154	-1.701544
24.41775	-1.317753
43.43122	-5.331218
43.43122	3.168782
18.98533	2.114666
37.9988	0.901201
35.28259	3.41741
37.9988	2.001201



List and evaluate 5 assumptions for the regression, stating the evidence you used. [15]

1. Straight line. Assumption met. No bowls or arches in plot

2. Homogeneous error. Assumption not met. Outlier
Assumption met No cone [$-\frac{1}{2}$]
while this is correct (no cone) there is still heterogeneity due to the outlier

3. Independent errors. Assumption met. No pattern (++++----- or + - + - + - + -)

4. Σresiduals = 0 Assumption met. Always met for least square estimate
Assumption met. Values from graph add up to zero.

5. Normal Errors Assumption not met. Not normal due to large outlier

6. Two different methods were used to determine the fat content in different samples of meat. Both methods were used on portions of the same meat sample. The data are percentages (Huntsberger 1967, p227).

Define response and explanatory variables, with symbols. [3]

Meat Sample	Method I	Method II
1	23.1	22.7
2	23.2	23.6
3	26.5	27.1
4	26.6	27.4
5	27.1	27.4
6	48.3	46.8
7	40.5	40.4
8	25.0	24.9

Fat content F

Method M

Sample S

In the space below, arrange the data in model format, with column headings for explanatory and response variables. [6]

%fat	Method	Sample
23.1	1	1
23.2	1	2
26.5	1	3
26.6	1	4
27.1	1	5
48.3	1	6
40.5	1	7
25.0	1	8
22.7	2	1
23.6	2	2
27.1	2	3
27.4	2	4
27.4	2	5
46.8	2	6
40.4	2	7
24.9	2	8