Biology 4605/7220 Name Key

6 November 2002 Exam #2b

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

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

name	symbol	scale

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

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

(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. Height is frequently named as a good predictor variable of weight among people of the same age and gender. Roberts (*American Journal of Clinical Nutrition* 54:499) measured the heights (cm) and weights (kg) of 14 males between the ages of 19 and 26 years of age. Does weight depend on height ? A=6 B=3 C=6 D=1

A.

name	symbol	scale
weight	W	cardinal
height	H	cardinal

C.

source	df
H	1
error	12
total	13

B. W = $\beta_o + \beta_H H + \epsilon$ [3]

D. regression [1]

2. Skinner and Allison (*J. Agric. Res.* 23:433-445) studied the effect of date of planting and amount of fertilizer (borax) on cotton growth, measured in pounds. Amount of borax was 0, 5, or 10 pounds. Three methods of borax application were (borax in drill & seed planted immediately, borax in drill & seed planted one week later, or borax broadcast). The experiment was carried out on 3 dates. When the analysis is carried out, all of the interaction terms were found to be non significant, with p-values of 0.173 or more. Write the model with no interaction terms.

A=12 B=5 C=10 D=1

A.

name	symbol	scale
cotton growth	ΔM	cardinal
Borax amount	B	nominal
Application method	App	nominal
Date	T	nominal

C.

source	df
B	2
App	2
T	2
error	20
total	26

B. $\Delta M = \beta_o + \beta_B B + \beta_{App} App + \beta_T T + \epsilon$ [3]

A. Three way ANOVA [1]

3a. Define a symbol for scutum width (units of μm) of ticks on rabbit #1 (Sokal and Rohlf, 1995, p 210), then define a symbol for the observed (sample) mean and the true (population) mean . . . . [3]

W = scutum width   observed mean $\overline{W}$

population mean   E(W), or μ_W

3b. For the data on scutum width (8 values below) write the observed mean.

____$\overline{W}$___ = __372.25__[1]

(Symbol) = (Value)

3c. Write a probability statement for the 95% confidence limits around the true mean . . . . [2]

$P\{L_1 < \mu_W < L_2\} = 1\text{-}\alpha = 95\%$

3d. What value of the t-distribution should you use for the 95% limits ? __________[1]

```
MTB > invcdf c1;   SUBC> t 7.
     0.0100   -2.9980
     0.0250   -2.3646 <--2.5% in each tail to obtain 95% limit
     0.0500   -1.8946
     0.1000   -1.4149
     0.9000    1.4149
     0.9500    1.8946
     0.9750    2.3646 <--2.5% in each tail to obtain 95% limit
     0.9900    2.9980
```

3e. Compute the 95% confidence limits . . . . [2]

```
MTB > print c2

 ScWidth   380    376    360    368    372    366    374    382

MTB > describe c2

              N     MEAN    MEDIAN   TRMEAN    STDEV   SEMEAN
     ScWidth  8    372.25   373.00   372.25     7.36     2.60
```

$\frac{7.36}{\sqrt{8}} = 2.60$

$L_1 = 372.25 - 2.3646 \cdot 2.6 = 366.098$

$L_2 = 372.25 + 2.3646 \cdot 2.6 = 378.398$

4a. Construct an ANOVA table for which the total Sum of Squares is 100, 15% of this variability is due to regression, and the sample size is 10. Be sure to compute MS and F-ratio . . . . [12]

Source	df	SS	MS	F
regression	1	15	15	1.411765
error	8	85	10.625	
total	9	100		

4b. Explain how you would compute a p-value for the F-ratio in the table you have constructed, if the residuals were heterogeneous and non normal . . . . [2]

randomize the response variable with respect to the regression variable, to render the null hypothesis true

compute F-ratio

do this at least 1000 times to obtain F-distribution when null hypothesis is true

compute % of distribution above F = 1.41, this is the p-value by randomization

4c. Circle the effect (increase/decrease) of doubling the sample size, in the ANOVA table you constructed (or any ANOVA table for regression) . . . . [3]

increase **[decrease]** in MS error

[increase] decrease in F-ratio

increase **[decrease]** in p-value