Biology 4605/7220 Name______________________________
6 November 2002 Exam #2a

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. 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

C. source df

B. ________ = ____ + ϵ [3]

D. [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

C. source	df

B .

________ = ____ + ϵ [3]

D. [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]

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

________ = ________[1]
(Symbol) = (Value)

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

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
    0.0500   -1.8946
    0.1000   -1.4149
    0.9000    1.4149
    0.9500    1.8946
    0.9750    2.3646
    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
```

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]

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]

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

Biology 4605/7220 Name________________________
6 November 2002 Exam #2

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. Daniel (*Biostatistics* 1995, p234) reported cell diameters (µm) of 40 lymphocytes and 50 tumor cells obtained from biopsies of tissue from patients with melanoma. Do cancerous and non-cancerous cells differ in diameter ? A=6 B=3 C=6 D=1

A. name symbol scale

C. source df

B. ________ = ________________________________ + ϵ [3]

D. [1]

2. Does birth weight depend on maternal smoking, controlled for gestation period and maternal weight ? Selvin (*Practical Biostatistical Methods*, 1995, Duxbury Press) reported birth weights of first infants (grams), gestation period (weeks), maternal smoking (0, 10-20, or ≥40 cigarettes per day), and maternal weight (kg) for 48 women over 40 years old. (Assume no interactive effects of explanatory variables on the response variable, as in multiple regression). A=12 B=5 C=10 D=1

A. name symbol scale

C. source df

B. ________ = ____ + ϵ [3]

D. [1]

3a. Construct a one-way ANOVA table for which the total Sum of Squares is 100, 15% of this variability is due to treatment effects (control vs one treatment), and the sample size is 100. Be sure to compute MS and F-ratio . . . . [12]

3b. Explain how you would compute a p-value for the F-ratio in the table you have constructed, if the residuals were normal and independent, with fixed variance . . . . . [2]

3c. Circle the effect (increase/decrease) of <u>halving</u> the sample size, in the ANOVA table you constructed (or any ANOVA table with the same model structure) . . . . [3]

increase	decrease	in MS error
increase	decrease	in F-ratio
increase	decrease	in p-value

4a. Define a symbol for scutum width, in microns, of tick larvae on rabbit #4 (Sokal and Rohlf, 1995, p 208), then define a symbol for the observed (sample) mean and the true (population) mean . . . . [3]

4b. For the data on scutum width (6 values below) write the observed mean.

________ = ________[1]
(Symbol) = (Value)

4c. Write a probability statement for the 95% confidence limits around the true mean . . . . [2]

4d. What value of the t-distribution should you use for the 95% limits ? __________[1]

```
MTB > invcdf c10;   SUBC> t 5.
    0.0100   -3.3649
    0.0250   -2.5706
    0.0500   -2.0151
    0.1000   -1.4759
    0.9000    1.4759
    0.9500    2.0151
    0.9750    2.5706
    0.9900    3.3649
```

4e. Compute the 95% confidence limits . . . . [2]

```
C6
   376    344    342    372    374    360

MTB > describe c6

                N     MEAN   MEDIAN   TRMEAN    STDEV   SEMEAN
C6     ScWidth  6   361.33   366.00   361.33    15.27     6.23
```