Biology 4605/7220/ST4581 Name ____________________________

Final Examination 11 December 1992

1. The general linear equation is a compact summary of many forms of parametric statistical analysis. The equation has the form:

Response = Explanatory + residual

$$Y = \beta_o + \beta X + \epsilon$$

The systematic part of the model (treatment or explanatory component) can consist of several terms ($\beta_1 X_1$ $\beta_2 X_2$ etc) composed of parameters ß and variables X. Write a general linear equation for each of the following designs.

nested ANOVA

t-test

Regression

Analysis of Covariance

Underline all interval or ratio scale variables. Circle all classification (nominal scale) variables. Don't forget to underline or circle both explanatory and response variables.

2. An entomologist observes the number of scale insects (n_t) on a tree in 9 successive days. Assuming an invariant mortality rate of $m = 0.2$ day^{-1} the expected number of scale insects ($nhat_t$) is

$$nhat_t = n_{t=0} e^{-mt}$$

Which column shows the residuals
(difference between observed and expected) from this model ? ______________

MTB > print c1-c7

ROW	t=days	n_t	nhat	C4	C5	C6
1	0	100	100.000	1.00000	0.00000	-0.04136
2	4	46	44.933	1.02375	1.06710	0.99750
3	5	38	36.788	1.03295	1.21206	1.21098
4	6	32	30.119	1.06244	1.88058	1.84663
5	7	23	24.660	0.93270	-1.65970	-1.70339
6	8	18	20.190	0.89155	-2.18965	-2.25062
7	9	15	16.530	0.90745	-1.52989	-1.55144

Based on the residuals does mortality appear to be invariant over time ? __________

State why or why not.

3. According to Schreck (1982 p 200 in Air Pollution--Physiological Effects) the Stokes number ST can be used to predict the tendency for deposition of particles in lung airways:

$$St = \frac{\rho\, d^2\, v}{18\ MU\ \ R}$$

where ρ is the density of particles (gram/litre),
d is the diameter of particles ($\mu m = 10^{-6} m$),
v is the velocity of particles (μm/second),
MU is absolute viscosity of air ($g\ m^{-1}\ second^{-1}$)
and R is the airway radius (cm).

Schreck (1982 p 207) reports that deposition in volunteer subjects varied in accordance with an anatomical Stokes number ST*

$$St^* = \frac{4\, d^2\, F}{18\ MU\ \pi\ (BDS)^3}$$

where F is flow rate (litres/minute) and
BDS is bronchial deposition size ($cm = 10^{-2} m$).
Which Stokes number (St or ST*) does not have dimensions (*i.e.* is correct) ? __________
Justify your answer by showing the dimensions for the incorrect Stokes Number
Show dimensions as exponents for M L T (Mass, Length, Time)

4. One goal of this course in quantitative methods was to increase your skill in moving among 3 forms of generalization: verbal, graphical, and formal models. Draw a picture showing the relation between data (or quantities) and these 3 forms of generalization.

Circle the forms of generalization that you would use in presenting your research in a paper or thesis.

Redraw the diagram and circle the form of generalization you would use in presenting your results to a newspaper reporter.

5. List two ways of judging whether a set of residuals are normally distributed.

List two solutions to the problem of non-normal residuals.
(List the better solution first)

Justify your choice of which solution to list first.

6. An ethologist records the turning movements of a planarian worm as follows:
R L L R L L L R R L R R L L L R R L L R R

The number of left turns is L = ________________

If turning is random, what is the expected number of R(ight) turns immediately after a L(eft) turn ? ___________

What is the observed number of R turns immediately after an L turn? _______

State a null Hypothesis H_o:______________________________

H_A:________________

Describe a randomization test with this data, to evaluate whether planarian worms turn at random.

7. A geneticist (Sewall Wright 1968 p 203) reported the mean weights (W = grams) of guinea pigs.

Three analyses of the data are attached (A, B, and C) below, as Minitab outfiles.

Calculate the observed difference in mean weight between the two strains
D = _______________ grams

You have been asked to determine whether this difference is statistically significant. Using the response variable W, write a General Linear Model for each of the 3 ANOVA tables (analyses).

Analysis A ____

Analysis B ____

Analysis C ____

Underline all interval or ratio scale variables. Circle all classification (nominal scale) variables. Don't forget to underline or circle both explanatory and response variables.

Which analysis is better A or B ? __________________

Why ?

Is analysis C an improvement over either A or B ? ________

Why or why not ?

8. Xenon lighting is less expensive than other electrical lighting, but emits light within a narrow range of wavelengths, compared to other artificial lighting. You have been asked to determine whether xenon lighting reduces growth rates of cucumbers in a greenhouse, compared to other sources of lighting.

Name your response quantity (including symbols, units, and procedural statement).

Name at least two explanatory variables that should be included in your investigation (be sure to write a symbol and procedural statement for each named explanatory variable).

Write a general linear model (GLM) to use in testing the effects of lighting on cucumber growth.

State a pair of hyptheses (H_o/H_A) for each term in your model.

Construct an ANOVA table (Source and df columns only) for your analysis. To do this you will need to decide how many observations to make, as well as how many levels to use for each classification variable.

```
MTB > print c1-c6

 ROW   year  strain13  strain35  weight  strain     yr

   1   1916     190.7     195.4   190.7      13   1916
   2   1917     206.3     206.7   206.3      13   1917
   3   1918     239.3     223.0   239.3      13   1918
   4   1919     240.4     227.2   240.4      13   1919
   5   1920     259.0     232.4   259.0      13   1920
   6   1921     228.6     207.1   228.6      13   1921
   7   1922     234.2     230.3   234.2      13   1922
   8   1923     256.3     245.9   256.3      13   1923
   9   1924     257.0     244.8   257.0      13   1924
  10                              195.4      35   1916
  11                              206.7      35   1917
  12                              223.0      35   1918
  13                              227.2      35   1919
  14                              232.4      35   1920
  15                              207.1      35   1921
  16                              230.3      35   1922
  17                              245.9      35   1923
  18                              244.8      35   1924

MTB > anova 'weight' = 'strain'  . . . . . . . . . . . . . . ANALYSIS A

Factor     Type Levels Values
strain    fixed      2    13    35

Analysis of Variance for weight

Source      DF          SS          MS        F       P
strain       1       544.5       544.5     1.28   0.275
Error       16      6823.4       426.5
Total       17      7367.9       433.4

MTB > anova 'weight' = 'strain' 'yr' . . . . . . . . . . . . . ANALYSIS B

Factor     Type Levels Values
strain    fixed      2    13    35
yr        fixed      9  1916 1917 1918 1919 1920 1921 1922 1923 1924

Analysis of Variance for weight

Source      DF          SS          MS        F       P
strain       1       544.5      544.50    10.69   0.011
yr           8      6415.8      801.98    15.74   0.000
Error        8       407.6       50.95
Total       17      7367.9      433.41
```

```
MTB > plot 'weight' by 'yr'

weight  -
        -                             *
        -                                                 *      *
     250+
        -                                                 *      *
        -                *     *
        -                                          *
        -                             *      *     *
     225+                *     *
        -
        -
        -
        -         2                          *
     200+
        -  *
        -  *
        -
          --------+---------+---------+---------+---------+--------yr

             1917.0    1918.5    1920.0    1921.5    1923.0

MTB > ancova weight = strain;  . . . . . . . . . . . . . . . . ANALYSIS C
SUBC> covariates yr.

Factor   Levels Values
strain        2    13    35

Analysis of Covariance for weight

Source      DF     ADJ SS          MS        F       P
Covariates   1     4117.2      4117.2    22.82   0.000
strain       1      544.5       544.5     3.02   0.103
Error       15     2706.2       180.4
Total       17     7367.9       433.4

Covariate      Coeff      Stdev    t-value        P
yr             5.858       1.23      4.777    0.000

MTB > stop
*** Minitab Release 6.1.1 *** Minitab, Inc. ***
Storage available 16179
```