Biology 4605/7220 Name _____Key_____
Exam #2b 5 November 2003

1a. Name the assumptions that underlie the accurate computation of p-values from F and t distributions. [4]

1. Σresiduals = 0
2. residuals independent
3. residuals homogeneous (fixed variance, homoscedastic)
4. residuals normal

1b List one assumption and state how you would check the assumption. [2]

1. Σresiduals = 0 for fitted values produced by statistical packages.
2. plot residuals against neighboring value, check for no trend (no correlation)
3. plot residuals versus fits, check for violation (unequal vertical distribution)
4. examine histogram, examine normal plot for deviations from straight line

2. A geneticist analyzes survival (in days) of the fruit fly *Drosophila subobscura* in relation to lifetime production of eggs (Prod) and egg size as measured by length (L).

2a. What units will the mean survival have ? __days____ [1]

Complete the ANOVA table. [9]

Source	df	SS	MS	F
Prod	_1_	3.340	_3.34_	_8.48_
L	_1_	21.842	_21.842_	_55.45_
Error	_22_	8.665	_0.394_	
Total	24	_33.85_		

2b. What is the variance in survival ? __33.85/24 = 1.41 days^2__ [1]

2c. If egg size L is omitted from the analysis of survival in relation to egg length, state what happens to each of the following (increase or decrease?). [3]

Error degrees of freedom __increase__

MS error __increase___

F- ratio for Prod _decrease__

3. Write the general linear model for the following analyses.
Use: Y = response variable
F1 = categorical variable with fixed levels
F2 = additional categorical variable with fixed levels
R1 = categorical variable with random levels
R2 = additional categorical variable with random levels
Z1 = variable on a cardinal scale (ratio scale or interval scale)
Z2 = additional variable on a cardinal scale (ratio scale or interval scale)

Use β_{F1} to denote parameter for F1, β_{F2} to denote parameter for F2, ϵ for normal error, *etc.*

T-test

GLM Y = _____ $\beta_o + \beta_{F1} \cdot F1 + \varepsilon$ _____ [1]

ANCOVA with interaction

GLM Y = ____ $\beta_o + \beta_{F1} \cdot F1 + \beta_{Z1} \cdot Z1 + \beta_{F1Z1} \cdot F1 \cdot Z1 + \varepsilon$ ____ [3]

Hierarchical ANOVA

GLM Y = ____ $\beta_o + \beta_{F1} \cdot F1 + \beta_{F2 \subset F1} \cdot F2 \subset F1 + \varepsilon$ ____ [2]

Two-way ANOVA no interaction term

GLM Y = ____ $\beta_o + \beta_{F1} \cdot F1 + \beta_{F2} \cdot F2 + \beta_{F1F2} \cdot F1 \cdot F2 + \varepsilon$ ____ [2]

4. For the following tests, write out the number of categorical and number of ratio scale explanatory variables. [6]

Name of test	Categorical	Ratio Scale
regression	0	1
multiple regression	0	2or more
3-way ANOVA	3	0

5. An experiment was designed to study the effects of three different drugs and three types of stressful situation in producing anxiety in adolescent subjects. The table shows the difference between the pre- and posttreatment scores of 12 subjects who participated in the experiment, for drugs A and B only. (Data from Daniel 1995 Ex 8.16 p337)

Stressful Situation (Factor A)	Drug (Factor B) A	B	Ratio
I	4	1	
	5	3	2.25
II	6	6	
	6	6	1
III	5	7	
	4	4	0.82

5a Assign symbols to variables [2]

Name	Symbol
A,Stress	D
B,Stress	S
Diff in scores	*DiffSc*

5b. Write a general linear model to analyze the data [5]

GLM DiffSc = $\beta_o + \beta_D \cdot D + \beta_S \cdot S + \beta_{D \cdot S} \cdot D \cdot S + \varepsilon$

5c. Compute the mean value in each of the 6 cells in the table, compute the ratio of mean scores of Drug A / Drug B for each stress type, and place this value in the Ratio column in the table. [3]

5d. Do your calculations suggest that there will be interactive effects of stress and drug type on scores ? Yes (No) [1]

Why or why not? Yes, because ratios differ [1]
(No, because ratios appear to differ, but true ratios do not)

5e. Complete the first two columns of the ANOVA table, for the data on 12 subjects shown above. [2]

Source	df
Drug	1
Stress	2
Drug*Stress	2
Error	6
total	11

6. A oenologist is interested in grape production (*Mgrape*, kg per $(2m)^2$ plot per year) in relation to watering and grape variety. An experiment is carried out in 4 beds with 9 plots in each bed. Each plot within each bed is assigned to one of the 9 possible combinations of variety (3 levels) and watering (3 levels).

7a Assign symbols to each explanatory variable that should be included in the analysis [3]

Name	Symbol
grape production	*Mgrape*
water	W
grape variety	V
bed	B

7b Write a general linear model to analyze the effects of water, variety, and any interactive effects of water and variety on grape production, controlled for differences among beds [3]

GLM DiffSc = $\beta_o + \beta_W \cdot W + \beta_V \cdot V + \beta_{W.V} \cdot W \cdot V + \beta_B \cdot B + \varepsilon$

df 71 = 2 2 4 3 60

7c Write in the degrees of freedom beneath each term in the model, assuming two measurements per plot. [3]

7d How would you examine whether the variety*water effect is consistent across beds? [1]

Look at 3-way interaction term W · V · B
If this term is not significant, conclude that W · V effect is consistent