

There are two main questions on this exam, A and B. Each has sub-questions and each of these are worth three points.

Question A

This question is loosely based on the research of Tara Forbis, an EPO Biology student who was in last year's class. She conducts research on alpine tundra ecology. Alpine plants often create dense covers which make it difficult for new seedlings to germinate. A source of natural soil disturbance on the tundra is the northern pocket gopher (*Thomomys talpoides*). Burrowing activities of gophers create large mounds of open soil. Another source of soil disturbance is human activity such as off-trail hiking. The seed germination rate is measured along with other characteristics at a number of randomly selected sites on Niwot Ridge. The following variables are available for each site.

ELEVATION — altitude in feet of the site

DISTURB — type of soil disturbance : Gopher, Human, or None

AGE — the age of the disturbance in years

SUN — typical number of daily hours of sunlight on the plot during the summer

RATE — seedling germination rate

Use the above variables to specify Model A/C comparisons you would use to answer the following research questions. If you construct any new variables, be sure to specify their definitions. You may also use the first letter of each variable name as an abbreviation.

1. Do sites at higher elevations tend to have lower germination rates?
2. A previous researcher estimated that germination rates decrease by .1 for each additional 1000 feet of elevation. Is that the case for these data?
3. As a group, do elevation, age of the disturbance, and hours of sunlight predict germination rates?
4. Is elevation related to age of the disturbance and hours of sunlight as a set?
5. Controlling for age of disturbance and hours of sunlight, is there a relationship between elevation and germination rates?
6. Assume that elevation, age of disturbance, and hours of sunlight are all related to germination rate. Suppose that for seedlings to be successful the germination rate must

be .3 (i.e., 30% of the seeds germinate). Is there evidence that sites at 11,000 ft, that were disturbed one year ago, and that get 9 hours of sunlight have germination rates exceeding .3?

7. Do the mean germination rates differ for the three types of disturbance?
8. In the context of the previous question, do mean germination rates differ according to whether the source of disturbance was gopher or human?
9. Is there still a difference between the germination rates for sites disturbed by gophers versus humans when controlling for elevation?
10. Elevation and sunlight are enduring characteristics of a site. Do the more transitory states (type of disturbance and age of disturbance) predict germination rates over and above the enduring characteristics?

Question B

The following data set is from D.S. Moor and G.P McCabe (1989) Introduction to the practice of statistics. Data were collected in a study of the effect of three different methods of instruction on reading comprehension in children. The three methods of instruction are referred to as "Basal," "DRTA," and "Strat" in the dataset. Sixty-six children were randomly assigned to one of the three teaching methods (with 22 children in each). RCB (Reading Comprehension Before) is a measure of reading comprehension taken before the instruction. RCA (Reading Comprehension After) is a measure of reading comprehension taken after the instruction.

To analyze these data, two contrast coded predictors were created as follows:

$x_1 = 1$ if Basal or DRTA; -2 if Strat;
 $x_2 = 1$ if Basal; -1 if DRTA; 0 if Strat;

In light of these codes and the output on the following pages, answer the following questions.

1. Is reading comprehension prior to instruction related to reading comprehension after instruction, ignoring which instruction condition children have been assigned to? (Give PRE, F^* , $pa-pc$, and $N-pa$.)
2. What are the mean levels of reading comprehension for the three instruction groups on RCA?
3. Does method of instruction make a difference (i.e., are there any differences among the means that you reported in response to question 2?) (Give PRE, F^* , $pa-pc$, and $N-pa$.)

4. Provide an interpretation for the slope associated with X1 in the model where RCA is regressed on X1 and X2.
5. Are the RCA means from the Basal and DRTA instruction groups significantly different from each other? (Give PRE, F*, pa-pc, and N-pa and a substantive conclusion.)
6. What is the confidence interval for the mean RCA difference between the Basal and DRTA instruction groups?
7. Are the RCA means from the DRTA and Strat instructions groups significantly different from each other? (Give PRE, F*, pa-pc, and N-pa and a substantive conclusion.)
8. Are there any differences in reading comprehension between the three instruction conditions on the measure of reading comprehension taken before the instruction began, i.e., on RCB? (Give PRE, F*, pa-pc, and N-pa).
9. In light of what you have learned about multiple regression, discuss why the regression coefficient for X2 differs between Models 2 and 3. Provide a substantive interpretation for the regression coefficient for X2 in Model 3.

Model: MODEL1
 Dependent Variable: RCA

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	1	239.74272	239.74272	30.152	0.0001
Error	64	508.87849	7.95123		
C Total	65	748.62121			

Root MSE	2.81979	R-square	0.3202
Dep Mean	8.07576	Adj R-sq	0.3096
C.V.	34.91675		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > T
INTERCEP	1	1.852427	1.18531598	1.563	0.1230
RCB	1	0.635820	0.11579202	5.491	0.0001

Model: MODEL2

Dependent Variable: RCA

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	2	108.12121	54.06061	5.317	0.0073
Error	63	640.50000	10.16667		
C Total	65	748.62121			
Root MSE		3.18852	R-square	0.1444	
Dep Mean		8.07576	Adj R-sq	0.1173	
C.V.		39.48262			

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > T
INTERCEP	1	8.075758	0.39247981	20.576	0.0001
X1	1	0.151515	0.27752514	0.546	0.5870
X2	1	-1.545455	0.48068764	-3.215	0.0021

Squared Partial					
Variable	DF	Type II SS	Corr Type II	Tolerance	
INTERCEP	1	4304.378788	.	.	
X1	1	3.030303	0.00470887	1.00000000	
X2	1	105.090909	0.14094983	1.00000000	

Model: MODEL3

Dependent Variable: RCA

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	3	383.19051	127.73017	21.671	0.0001
Error	62	365.43070	5.89404		
C Total	65	748.62121			
Root MSE		2.42777	R-square	0.5119	
Dep Mean		8.07576	Adj R-sq	0.4882	
C.V.		30.06238			

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > T
INTERCEP	1	1.290925	1.03715746	1.245	0.2179
X1	1	-0.074296	0.21387951	-0.347	0.7295
X2	1	-1.813277	0.36809306	-4.926	0.0001
RCB	1	0.693187	0.10146965	6.831	0.0001

Squared Partial					
Variable	DF	Corr Type II	Tolerance		
INTERCEP	1	.	.		
X1	1	0.00194247	0.97611512		
X2	1	0.28129963	0.98865644		
RCB	1	0.42946026	0.96530404		