

[Note that typically, part of this exam includes annotation and explanation of computer output. This has NOT been included here, but is a significant component of the exam]

1. You want to use race as a variable in a multiple regression analysis. Please recode the data below (creating new variables if needed). Use “Wht” as the reference group. (10 pts):

<u>Sub</u>	<u>Race</u>
1	Wht
2	Wht
3	Blk
4	His
5	Blk
6	His

2. What are the advantages of using a distribution-free test? (10 pts)
3. You are going to do a trend analysis in multiple regression, and therefore must use powered vectors appropriately. You decide you want to fit a cubic line to your data, so you enter the predictor variable (A) cubed. So, you create a new variable called “AAA” ($A*A*A$), and enter that single variable into the regression run. Did you do this correctly? Why? If there is a problem, what? (15 pts)
4. You have tossed a coin a number of times, and you have come out with the following sequence or clusters of heads and tails:

H H H H T T H H H H T T T H H T

Is the sequence or clustering by chance, or is it different from chance? Use the appropriate test and interpret your results (10 pts).

5. You have three variables. The correlation matrix resembles the following...use a Venn diagram to illustrate the % variance interrelationships among the variables. (10 pts)

	Var1	Var2	Var3
Var1	1	0.7	0.7
Var2	0.7	1	0.0
Var3	0.7	0.0	1

Answers:

1.

<u>Sub</u>	<u>Race</u>	Dum1	Dum2
1	Wht	0	0
2	Wht	0	0
3	Blk	1	0
4	His	0	1
5	Blk	1	0
6	His	0	1

2.

Can use non-normal data
Analyses based on medians instead of means
Less sensitive to extreme scores

3.

Although you calculated “AAA” correctly, just entering that term into the regression run by itself does not give you the cubic term. In fact, the term you calculated is contaminated by the linear and quadratic trends. To do this correctly, you would first enter the Linear term (“A”), then enter the quadratic “AA” – then enter the cubic term. Hence, you are using a block-entry approach. As each term is entered, evaluate the R-square change and its significance for the appropriate interpretations.

4.

R = 6, M = 10, N = 6
Tabled Crit range is 4/13
Nonsig finding....no clustering or sequencing

5.

First, convert your correlations to % variance overlap. Var 1&2 correlate at .70 (49%), and 1&3 correlate at .70 (49%). There is no correlation between 2&3. Draw the circles for 1&2 first, then add variable 3 to the graph. 2&3 should not touch, while most of the variance in Var 1 should be overlapped. Play with it a bit, and you’ll get the picture.