

AMS 394 FINAL

11:15 AM - 1:45 PM

Please submit a word document

Question 1 (35 points)

Given the following data

X	Y	Z
30	70	25
29	68	27
31	72	17
47	93	20
40	84	10
27	65	38
46	91	36
50	96	29

- (a) Write a SAS program and compute a correlation matrix; that is, the correlation coefficient between each variable versus every other variable. And conclude if the correlation is significantly different from 0 at significance level 0.05.
- (b) Compute a new variable LX which is the natural logs of the original values of variable X and compute the correlation between Y and LX and test if the correlation is significantly different from 0 at significance level 0.05.
- (c) Compute a regression line (Y on LX). Y is the dependent variable and LX is the independent variable. (LX is the natural log of X) Report the regression equation and coefficient of determination.
- (d) What are the estimates of the slope and intercept? Are they significantly different from zero at significance level 0.10?
- (e) Generate a plot containing the Prediction (fitted values) and 90% confidence interval versus LX of the model in (c) with the REG statement.

Question 2 (15 points)

A friend gives you some summary data on the relationship between socioeconomic status (SES) and asthma as follows:

Asthma	Yes	No
Low SES	40	100
High SES	30	130

- (a) Conduct a chi-square test to test if Asthma and SES are independent with significance level 0.01.
- (b) Report the test statistics and degrees of freedom of the chi-square test.

Question 3 (25 points)

Given the following data

X1	X2	X3	X4	Y
34	32	44	43	283
2	20	11	9	54
1	36	17	46	77
33	45	18	34	175
22	19	16	30	135
10	50	9	49	72
16	26	2	15	61
36	41	30	23	232
42	2	48	44	322
44	44	15	11	196

- (a) Compute a multiple regression of Y on X1, X2, X3, X4 (without intercept) and report the regression equation.
- (b) Compute a multiple regression of Y on X1, X2, X3, X4 (with intercept) and report the regression equation.
- (c) Conduct best subset regression on the model with intercept and find the best model with 2 variables according to R square.

Question 4 (25 points)

Four different methods of preparing for a college entrance exam were compared. They are labeled 'A', 'B', 'C' and 'D'. The following exam scores were obtained for each of the four programs:

A	570	530	540	535	585	537	590
B	555	512	510	520	510	512	570
C	512	518	555	502	510	520	516
D	505	508	512	520	543	523	517

- (a) Conduct a one-way ANOVA test to analyze if the mean scores of different methods are the same with significance level 0.10.
- (b) Conduct a DUNCAN multiple range test to compare the mean score of these four methods and reach a conclusion with significance level 0.10.
- (c) Create a contrast of A versus B and C and reach a conclusion with significance level 0.05.