Part 1

Refer to **Grade point average** Problem

a. Obtain a 95 percent interval estimate of the mean freshman GPA for students whose ACT test score is 28. Interpret your confidence interval.

b. Mary Jones obtained a score of 28 on the entrance test. Predict her freshman GPA using a 95 percent prediction interval. Interpret your prediction interval.

c. Is the prediction interval in part (b) wider than the confidence interval in part (a)? Should it be?

d. Determine the boundary values of the 95 percent confidence band for the regression line when $X_h = 28$. Is your confidence band wider at this point than the confidence interval in part (a)? Should it be?

Part 2

Refer to **Grade point average** Problem

a. Set up the ANOVA table.

b. What is estimated by *MSR* in your ANOVA table? by *MSE*? Under what condition do *MSR* and *MSE* estimate the same quantity?

c. Conduct an F test of whether or not $\beta_1 = 0$. Control the α risk at .01. State the alternatives, decision rule, and conclusion.

d. What is the absolute magnitude of the reduction in the variation of Y when X is introduced into the regression model? What is the relative reduction? What is the name of the latter measure?

e. Obtain r and attach the appropriate sign.

f. Which measure, R^2 or r, has the more clear-cut operational interpretation? Explain.

Part 3

Refer to **Muscle mass** Problem

a. Obtain a 95 percent confidence interval for the mean muscle mass for women of age 60. Interpret your confidence interval.

b. Obtain a 95 percent prediction interval for the muscle mass of a woman whose age is 60. Is the prediction interval relatively precise?

c. Determine the boundary values of the 95 percent confidence band for the regression line when $X_h = 60$. Is your confidence band wider at this point than the confidence interval in part (a)? Should it be?

Part 4

Refer to **Muscle mass** Problem 1.27.

a. Plot the deviations $Y_i - \hat{Y}_i$ against X_i on one graph. Plot the deviations $\hat{Y}_i - \bar{Y}$ against X_i on another graph, using the same scales as in the first graph. From your two graphs, does *SSE* or *SSR* appear to be the larger component of *SSTO*? What does this imply about the magnitude of R^2?

b. Set up the ANOVA table.

c. Test whether or not $\beta_1 = 0$ using an F test with $\alpha = .05$. State the alternatives, decision rule, and conclusion.

d. What proportion of the total variation in muscle mass remains "unexplained" when age is introduced into the analysis? Is this proportion relatively small or large?

e. Obtain R^2 and r.

Part 5

Refer to **Grade point average** Problem

a. Would it be more reasonable to consider the X_i as known constants or as random variables here? Explain.

b. If the X_i were considered to be random variables, would this have any effect on prediction intervals for new applicants? Explain.

Part 6

Water flow. An engineer, desiring to estimate the coefficient of correlation ρ_{12} between rate of water flow at point A in a stream (Y_1) and concurrent rate of flow at point B (Y_2), obtained $r_{12} = .83$ in a sample of 147 cases. Assume that bivariate normal model (2.74) is appropriate.

a. Obtain a 99 percent confidence interval for ρ_{12}.

b. Convert the confidence interval in part (a) to a 99 percent confidence interval for ρ_{12}^2.

Part 7

Refer to **Muscle mass** Problem Assume that the normal bivariate model is appropriate.

a. Compute the Pearson product-moment correlation coefficient r_{12}.

b. Test whether muscle mass and age are statistically independent in the population; use $\alpha = .05$. State the alternatives, decision rule, and conclusion.

c. The bivariate normal model assumption is possibly inappropriate here. Compute the Spearman rank correlation coefficient, r_S.

d. Repeat part (b), this time basing the test of independence on the Spearman rank correlation computed in part (c) and test statistic (2.101). Use $\alpha = .05$. State the alternatives, decision rule, and conclusion.

e. How do your estimates and conclusions in parts (a) and (b) compare to those obtained in parts (c) and (d)?

Assignment 2

- Submit your responses in Blackboard in a **single** pdf file by Thursday, September 28, midnight.
- Use the following file name: LASTNAME_FIRSTNAME_ASUID_ASSIGNMENTNUMBER
- Prepare your pdfs carefully; each week some of you will present their work.
- Include the `R` (or another statistical software you prefer to use) commands you used in your pdf.