Chapter 9 – Correlation and Regression
Chapter Outline

I. Introduction
   A. Correlation Values
      1. Ranges from –1.0 to +1.0
         a. 0 = no correlation (zero correlation)
         b. +1.0 = perfect positive correlation (as one value goes up the other goes up)
         c. –1.0 = perfect negative correlation (as one value goes up the other goes down)
   B. Assumptions and Conditions for the Pearson Correlation
      1. The two variables have a linear relationship.
      2. Scores on one variable are normally distributed for each value of the other variable and vice versa.
      3. Outliers can have a big effect on correlation.
   C. Assumptions and Conditions for Spearman Rho
      1. Data on both variables are at least ordinal.
      2. Scores on a variable are monotonically related to the other variable (as the value of one variable increases, the other should increase, but not necessarily in a linear fashion).
      3. Rho is computed by ranking the data for each variable and then computing a Pearson product moment correlation.

II. Problem 9.1: Scatterplots to Check Assumptions
   A. Features of a scatterplot.
      1. Plot or graph to show the association of two variables.
      2. High positive correlations will plot a nearly straight line (linear regression line) from the lower left corner of the plot to the upper right corner of the plot.
      3. High negative correlations will plot a nearly straight line (linear regression line) from the upper left corner of the plot to the lower right corner of the plot.
   B. Follow the directions in the book to create and interpret a scatterplot.

III. Problem 9.2: Bivariate Pearson and Spearman Correlations
   A. Pearson Product Moment Correlation
      1. Bivariate parametric statistic
      2. Used when both variables are approximately normally distributed.
   B. Spearman Rho
      1. Nonparametric equivalent of the Pearson correlation.
      2. Use with ordinal data or if assumptions for the Pearson correlation are violated.
   C. Follow the directions in the book to compute and interpret a Pearson correlation and a Spearman rho.

IV. Problem 9.3: Correlation Matrix for Several Variables

A. Follow the directions in the book to compute and interpret a correlation matrix. This is used when you have more than 2 ordinal or normally distributed variables that you want to correlate.

V. Problem 9.4: Internal Consistency Reliability with Cronbach's Alpha
   A. Cronbach's Alpha
      1. Common measure of internal consistency reliability.
      2. Useful for checking several scores that the researcher wants to add together to obtain a summated score.
      3. Based upon a correlation matrix and is interpreted similarly to other reliability measures.
         a. Alpha should be positive.
         b. Value should be greater than .70 to provide support for internal consistency reliability.
   B. Follow the directions in the book to compute and interpret Cronbach's Alpha.

VI. Problem 9.5: Bivariate or Simple Linear Regression
   A. Utilized to make predictions about the relationship between two normally distributed variables. (Also referred to as simple regression or simple linear regression).
   B. Follow the directions in the book to compute and interpret a bivariate regression.

VII. Problem 9.6: Multiple Regression
   A. Utilized to make predictions about the dependent variable from a combination of several predictor variables.
      1. Can use a combination of interval, scale, and/or dichotomous independent/predictor variables.
      2. Can look for the combination of variables that provide the best prediction.
   B. Assumptions and Conditions of Multiple Regression
      1. The relationship between each of the predictor variables and the dependent variable should be linear.
      2. Errors are normally distributed.
      3. Variance of the residuals (the difference between the actual and predicted scores) is constant.
      4. Multicollinearity (condition where there is high intercorrelations among some set of the predictor variables) is minimal.
   C. Follow the directions in the book to compute and interpret a multiple regression.