Chapter 10
Correlation Analysis

J.C. Wang

Western Michigan University
Department of Statistics
Goal and Objective

To learn about the correlation coefficient as measure of strength of the \textit{linear} association between two variables.
Outline

1. Correlation Analysis
   - Purpose of Correlation Analysis
   - Correlation Coefficient
   - Sample Correlation Coefficient
Outline

1 Correlation Analysis
   - Purpose of Correlation Analysis
   - Correlation Coefficient
   - Sample Correlation Coefficient
The correlation measures the strength of the *linear* relationship between numerical variables, for example, the height of men and their shoe size or height and weight. In these situations the goal is not to use one variable to predict another but to show the strength of the *linear* relationship that exists between the two numerical variables.
Correlation Coefficient

definition

The strength of \textit{linear} association between two numerical variables in a population is determined by the correlation coefficient, $\rho$, whose range is $-1$ to $+1$.

\begin{center}
\begin{tabular}{cccccc}
strong & $\leftarrow$ & weak & no & weak & $\rightarrow$ & strong \\
$-$ & $-$ & $-$ & 0 & $+$ & $+$ & $+$ \\
$-1$ & $-0.65$ & $-0.35$ & 0 & $0.35$ & $0.65$ & $+1$
\end{tabular}
\end{center}

The sign (positive or negative) of the correlation = the sign of the slope of straight line.
Graphical Examples

\[ r = \text{sample correlation} \]
Graphical Examples
continued
Sample Correlation Coefficient

Since our interest is the regression analysis, the sample correlation coefficient \( r \) is derived from the coefficient of determination \( R^2 \), to be discussed in Chapter 11.

\[
R^2 = \frac{\text{regressionSumOfSquares}}{\text{totalSumOfSquares}} = \frac{SSR}{SST}
\]

The correlation coefficient \( r = \pm \sqrt{R^2} \) where \( r \) takes the sign of the slope.
Sample Correlation Coefficient

continued

Correlation coefficient can also be calculated by

$$r = \frac{\sum (x_i - \bar{x})(y_j - \bar{y})}{\sqrt{(x_i - \bar{x})^2 \sqrt{(y_j - \bar{y})^2}}}$$
Example

\[ r = 0.9526 \]

| Data |  
|-----|------|------|------|------|------|  
| \( x \) | -2 | 2 | 5 | -1 | 6 |  
| \( y \) | 0 | 3 | 10 | 1 | 15 |  

STAT \rightarrow\ filename L_1 \text{ and } L_2

STAT \rightarrow\ TESTS \downarrow LinRegTTest
iClicker Question 10.1