Measures of Relationship

Data Science for Quality Management: Describing Data Numerically

with **Wendy Martin**

Learning objectives:

Discriminate between correlation & association
Calculate correlation for two variables

Measures of Relationship

Correlation and association are measures of the strength of a relationship between two variables.

Measures of Relationship

Before we calculate statistics related to relationship, we must first properly classify each variable.

- Nominal
- Ordinal
- Continuous

Correlation

- Where both variables are continuous, the statistic employed to measure the relationship may be referred to as a Coefficient of Correlation

Association

- Where both variables are nominal, the statistic employed to measure the relationship may be referred to as a Coefficient of Association

Correlation and Association

- Coefficients of Correlation and Association can vary given all possible combinations of nominal, ordinal, and continuous data that can occur

Coefficient of Correlation

- The most frequently used coefficient of correlation used is the Pearson Product-Moment Coefficient of Correlation.
- Symbols
  Population: ρ_{xy}
  Sample: r_{xy}

Coefficient of Correlation

- The most frequently used coefficients of correlation is the Pearson Product-Moment Coefficient of Correlation.
- Symbols
  Population: ρ_{xy}
  Sample: r_{xy}

Product Moment Coefficient

$$r_{xy} = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{\sqrt{\sum(X-\bar{X})^2 \sum(Y-\bar{Y})^2}}$$

Product Moment Coefficient

Two components:

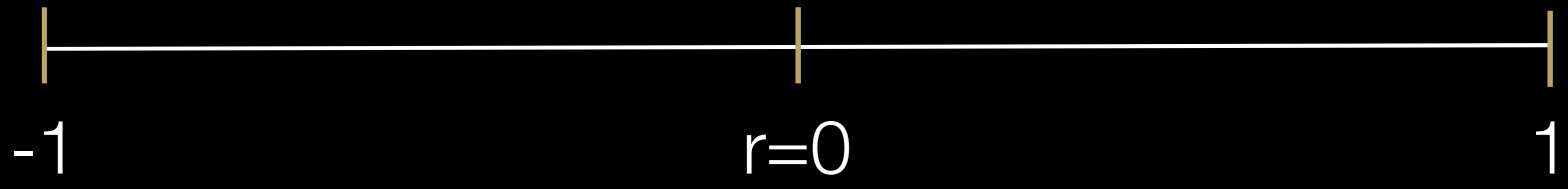
- Sign (+ or -)
- Numeric Value

Product Moment Coefficient

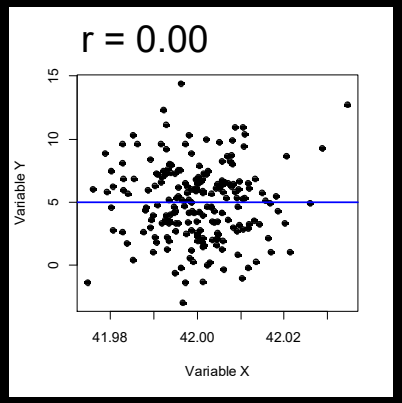
Sign (+ or -) gives the direction of the relationship

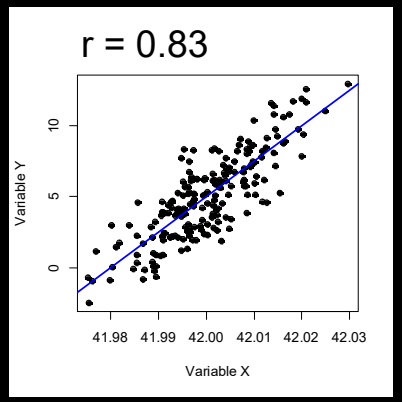
- Positive: As one variable increases in magnitude, the other variable increases
- Negative: As one variable increases in magnitude, the other variable decreases

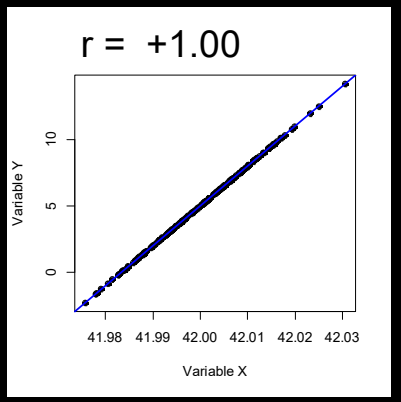
Product Moment Coefficient

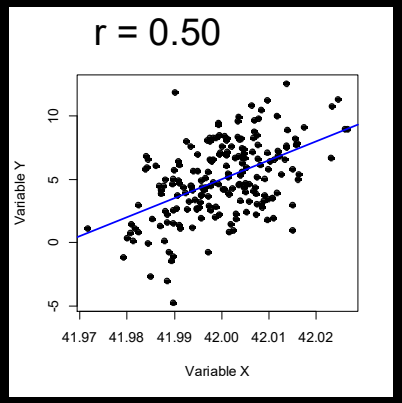


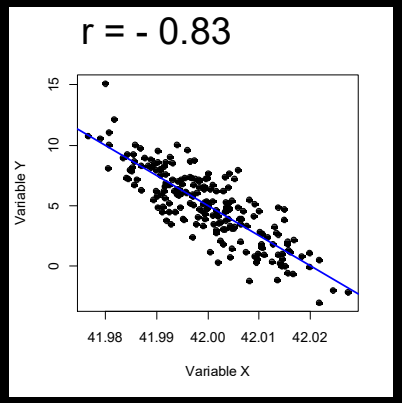
Scatterplot Examples

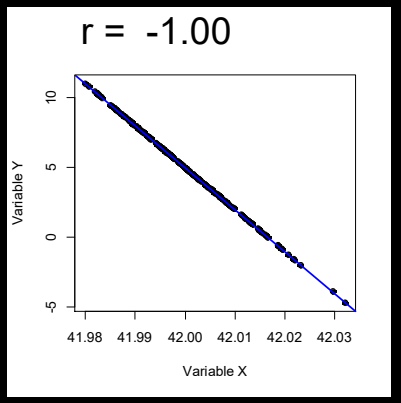












Sources

The material used in the PowerPoint presentations associated with this course was drawn from a number of sources. Specifically, much of the content included was adopted or adapted from the following previously-published material:

- Luftig, J. An Introduction to Statistical Process Control & Capability. Luftig & Associates, Inc. Farmington Hills, MI, 1982
- Luftig, J. Advanced Statistical Process Control & Capability. Luftig & Associates, Inc. Farmington Hills, MI, 1984.
- Luftig, J. A Quality Improvement Strategy for Critical Product and Process Characteristics. Luftig & Associates, Inc. Farmington Hills, MI, 1991
- Luftig, J. Guidelines for Reporting the Capability of Critical Product Characteristics. Anheuser-Busch Companies, St. Louis, MO. 1994
- Spooner-Jordan, V. Understanding Variation. Luftig & Warren International, Southfield, MI 1996
- Luftig, J. and Petrovich, M. Quality with Confidence in Manufacturing. SPSS, Inc. Chicago, IL 1997
- Littlejohn, R., Ouellette, S., & Petrovich, M. Black Belt Business Improvement Specialist Training, Luftig & Warren International, 2000
- Ouellette, S. Six Sigma Champion Training, ROI Alliance, LLC & Luftig & Warren, International, Southfield, MI 2005