Topic 2.1

Introducing Statistics: Are Variables Related?

Are Variables Related?

Unit 2 expands concepts learned in Unit 1 to explore relationships between categorical variables and between quantitative variables in two-variable datasets.

In this unit, we will use two-way tables, graphical representations, and summary statistics to describe the association between two categorical variables. Many of the skills we develop while working with two categorical variables will become building blocks for calculating probabilities in Unit 4, and for making statistical inferences about categorical data in Unit 6 and Unit 8.

We will also learn to describe the form, direction, strength, and unusual features of a relationship (association) between two quantitative variables. The correlation coefficient *r* will be used and interpreted (if appropriate) to describe these associations. For linear relationships, we will learn to use a linear model to predict values of a quantitative response variable from values of a quantitative explanatory variable. We will interpret the least-squares regression line in context, analyze residuals (errors in prediction), and explore departures from a linear pattern. The skills we will develop while working with scatterplots, the correlation coefficient r, and the least-squares regression line will become building blocks for making statistical inferences about the slope of a regression line in Unit 9.

To prepare for the AP exam, we suggest focusing on developing the following skills:

- For categorical variables be able to:
  - Represent and describe the association between two variables using frequency tables and relative frequency tables.
  - Calculate joint, marginal, and conditional relative frequencies from a two-way table.

- For quantitative variables be able to:
  - Represent the association (relationship) between variables using scatterplots.
  - Describe a scatterplot in terms of shape, direction, strength, and unusual features.
  - Evaluate and interpret the correlation coefficient for a linear relationship.
  - Identify the explanatory and the response variables.
  - Identify and interpret coefficients of the least-squares regression line model.
  - Describe the relationship of two variables using residual plots.
  - Calculate a predicted response value using a linear regression model.