5. Proportions

When your observations consist of *numeric* data the *numerical* summaries used are the mean and the standard deviation. When your observations consist of *attribute* data, the numerical summary is the *proportion*.

Sometimes it will help to think of membership in a category as a “yes" or “no" answer to the question: “Is this observation in this category?"

The most obvious example of this is opinion polling: you ask each member of the sample, for instance, whether or not they are a registered Republican. Instead of a number for each observation you have a “yes” or a “no” response.

What you do is then place your data in categories – some of the subjects respond with a “yes” and some respond with a “no.” You can then compute a proportion. For census data this looks like

$$\begin{aligned}\text{population proportion} &= p \\ &= \frac{\text{number answering yes}}{\text{population size}}\end{aligned}$$

while, for sample data, it is

$$\begin{aligned}\text{sample proportion} &= \hat{p} \\ &= \frac{\text{number answering yes}}{\text{sample size}}\end{aligned}$$

The symbol p is used for the population proportion and the symbol $\hat{p}$ (p-hat) is used for the sample proportion.

Note that the proportion (sample or census) must always be between 0 and 1; the corresponding percentages are

$$p \times 100\% = \text{population percent}$$

and for sample data

$$\hat{p} \times 100\% = \text{sample percent}$$

We will generally use the proportions rather than percentages because the formulae used later in the course rely on proportions, not percentages.

Usually the “yes" response is thought of as “success" and the “no" response is thought of as “failure." Usually you set up your problem so that whatever you are studying is the “yes" or “success" response. When this terminology is used, “Success" and “Failure" have no intrinsic meaning beyond being a “yes" or “no" response.

5.1. Example.

Mechanics R Us *hires two classes of employees: mechanics and managers. Last year, the franchise in Oklahoma hired a total of 425 male mechanics out of a total of 465 male applicants for mechanic positions. What proportion of male applicants for mechanic positions was hired?*

Solution. The total number of observations is

$$n = 465.$$

We are studying “being hired" so the question is

- “Was the subject hired as a mechanic?"

The “yes" respondents to this question are the “successes." There are

$$\boldsymbol{k} = 425$$

successes.
Thus the hiring rate for male mechanics is

$$\boldsymbol{p} = \frac{425}{465} = 0.9139$$

The proportion is 0.9139; the percentage is 91.39%.