Notice that `sex` is displayed on the x-axis and therefore follows the modeling ˜.

3.5.3 Displays of Categorical Variables

For categorical variables, it makes no sense to compute descriptive statistics such as the mean, standard deviation, or variance. Instead, look at the number of cases at each level of the variable.

```
> tally( ˜ sex, data=galton)

   F    M Total
 433  465   898
```

Proportions can be found in a similar way:

```
> tally ( ˜ sex, data=galton, format="proportion")
    F     M Total
0.482 0.518 1.000
```

Note that the variable, in this case `sex`, follows the ˜. That's because the data set is being broken down by the variable.

Reading Questions

1. What is the disadvantage of using a 100% coverage interval to describe variation?

2. In describing a sample of a variable, what is the relationship between the variance and the standard deviation?

3. What is a residual?

4. What's the difference between "density" and "frequency" in displaying a variable with a histogram?

5. What's a normal distribution?

6. Here is the graph showing boxplots of height broken down according to sex as well as for both males and females together.