# Monday, Aug 30

Normal ("Bell-Shaped") Distributions

Example: Consider the distribution of the observations of the heights of adult men and women in the United States.

1. For **men** the mean and standard deviation would be about **69** and **3** inches, respectively.

2. For **women** the mean and standard deviation would be about **63.5** and **2.5** inches, respectively.

Also those distributions would be approximately *normal* or *bell-shaped* (or sometimes *Gaussian*).
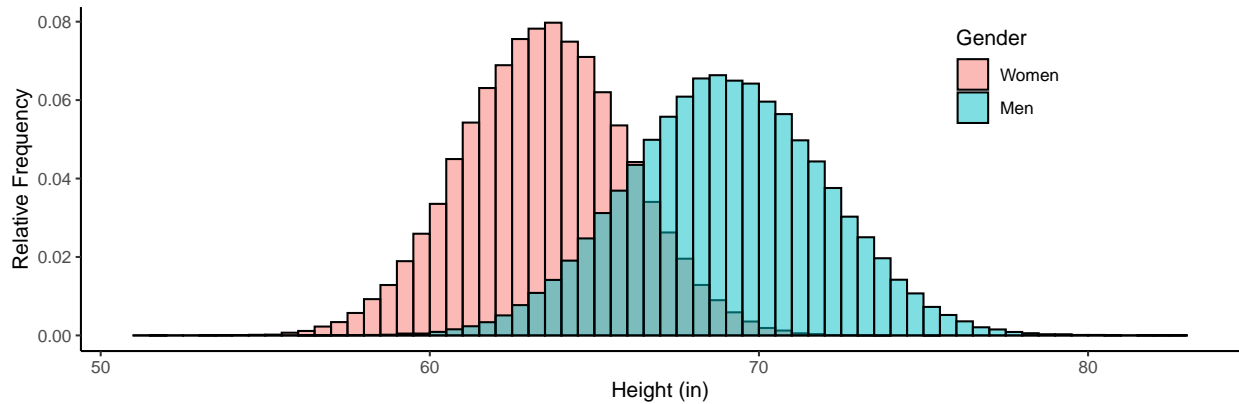
Relationship with Mean and Standard Deviation

The *shape* of a normal distribution is related to its mean and standard deviation.

1. The mean is located where the (relative) frequency is at its peak.

2. The standard deviation is the distance from the mean to the value of the variable where the (relative) frequency is a little more than half way (actually about 61%) to its maximum.

Example: Consider the distributions of the heights of adult men and women.
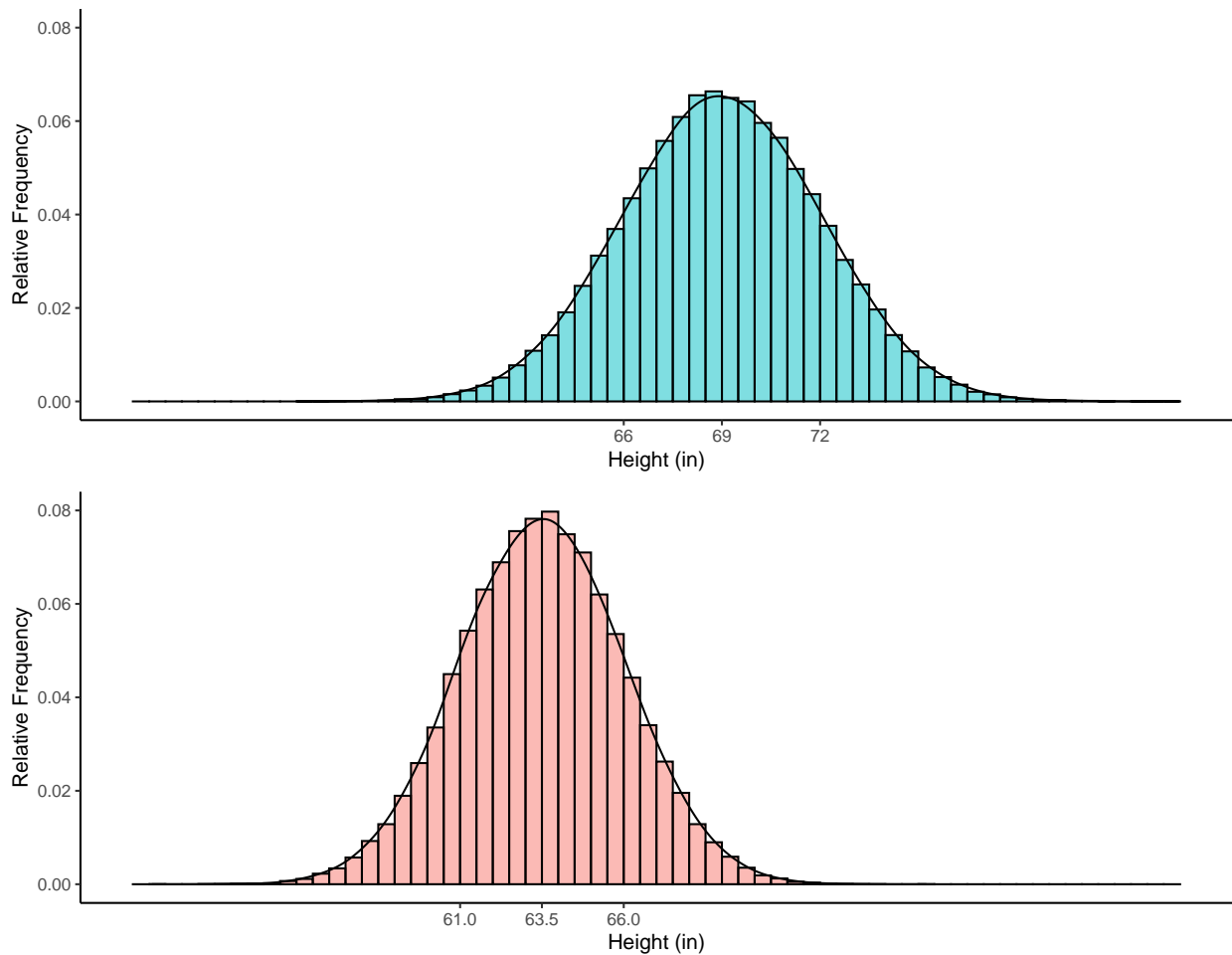
Example: Suppose that the distribution of Hobbit height is normal in shape with a mean of 107 cm and a standard deviation of 10 cm. How could we plot an outline of this distribution?

The "Empirical Rule" for Normal Distributions
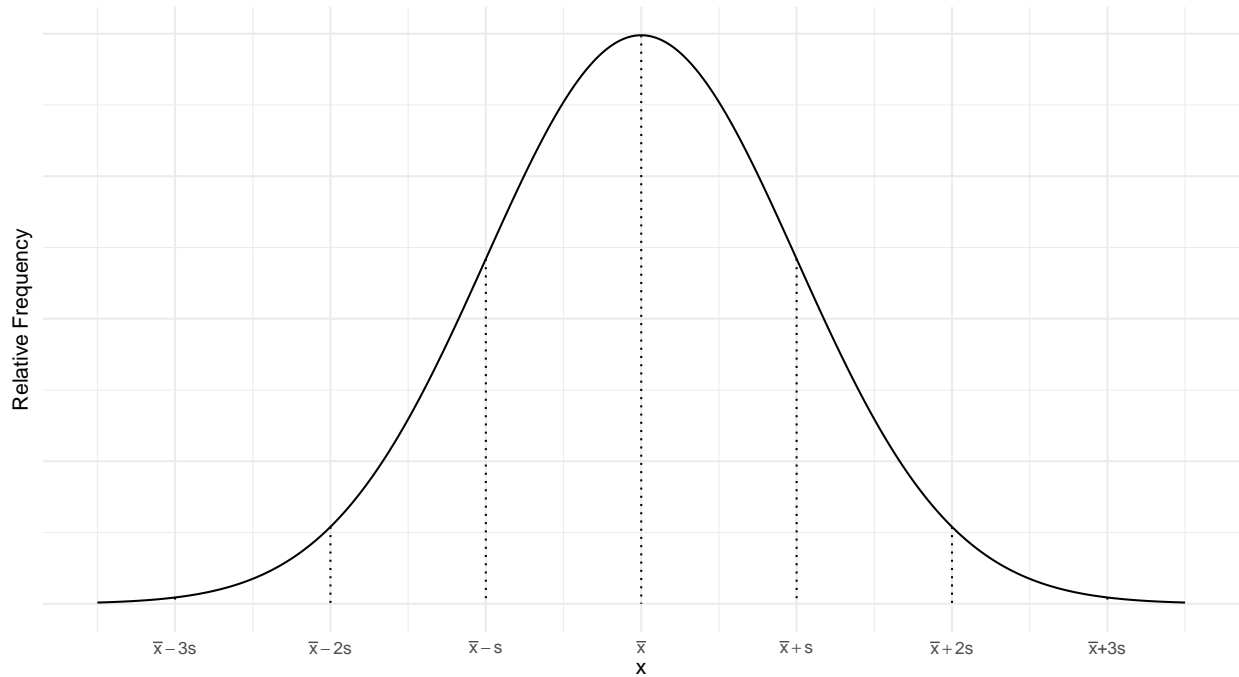


If a distribution is normal then we know the following.

1. Approximately 68% of observations are within *one* standard deviation of the mean.
2. Approximately 95% of observations are within *two* standard deviations of the mean.
3. Approximately 100% of observations are within *three* standard deviations of the mean.

Example: Where would we find the middle 68%, 95%, and almost all of the observations of Hobbit height if the distribution of Hobbit height is normal in shape with a mean of 107 cm and a standard deviation of 10 cm?

Example: Where are the middle 95% of observations of the heights of men if the distribution of the heights of men is approximately normal with a mean of 69 and a standard deviation of 3?

Example: Where are the middle 95% of observations of the heights of women if the distribution of the heights of women is approximately normal with a mean of 63.5 and a standard deviation of 2.5?

The percents stated in the empirical rule are based on percentiles.
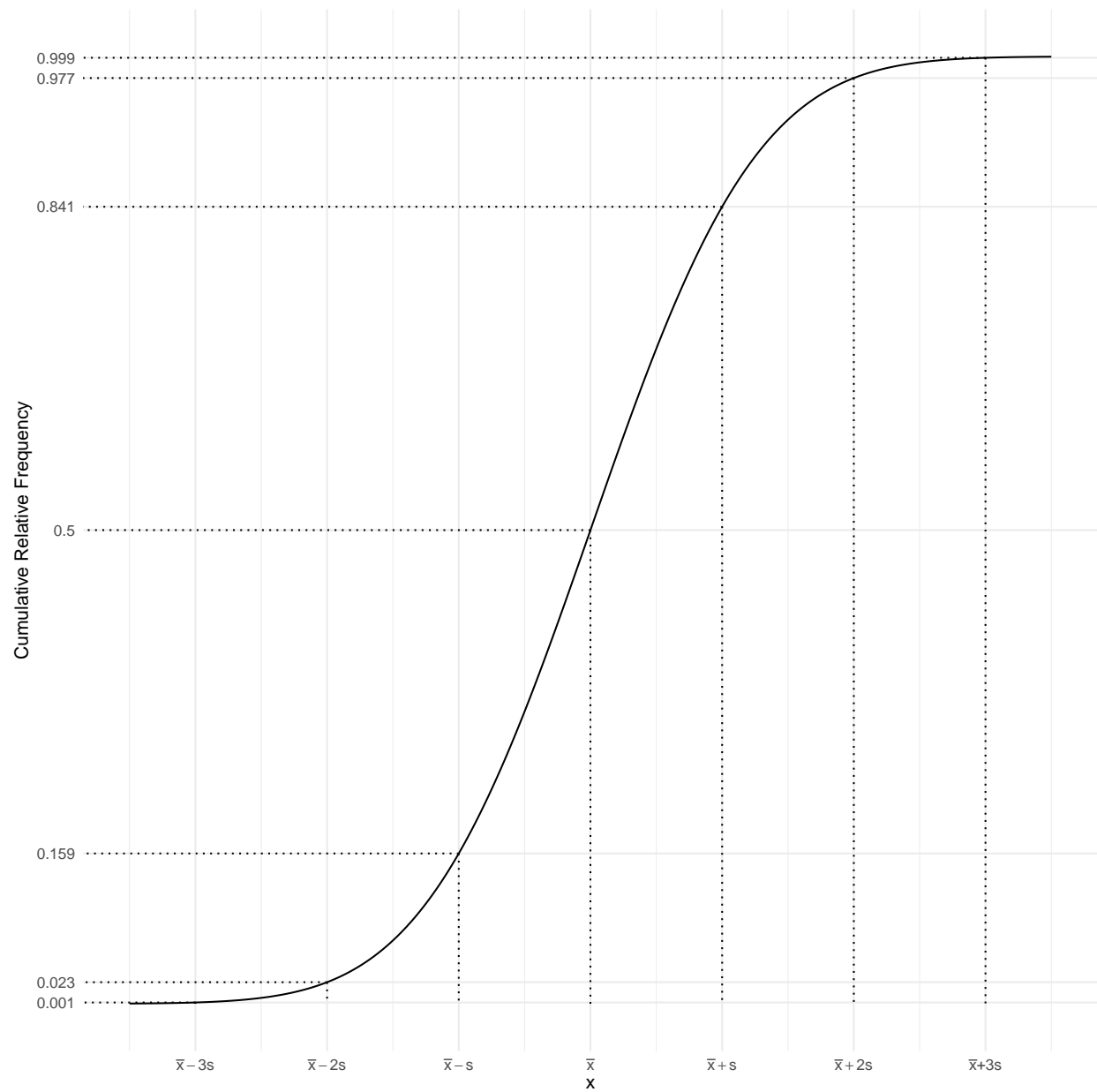
Example: Suppose I flipped a coin 200 times and observed the number of times it came up heads. Now suppose I did this many thousands of times to get a sample of observations of the number of times the coin comes up heads when flipped 200 times. The plot below shows the distribution of the number of times the coin comes up heads out of 200 flips.
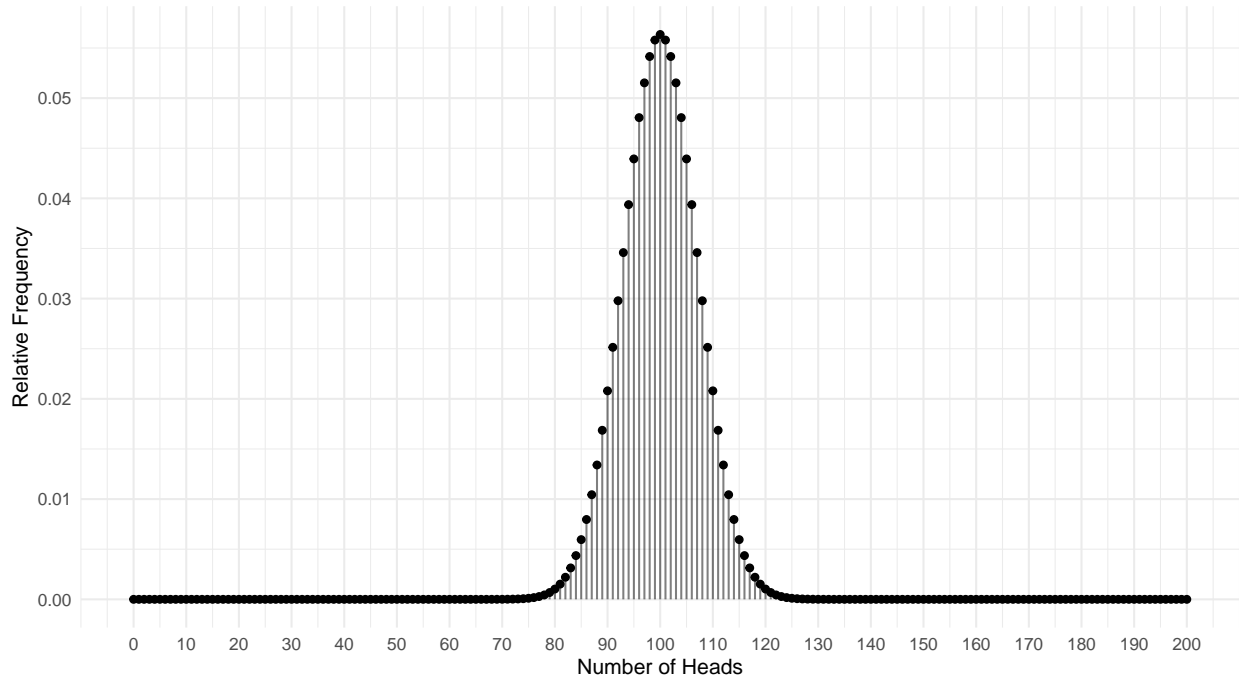
This distribution is clearly normal in shape, and it has a mean of 100 and a standard deviation of approximately 7. What does the empirical rule tell us about where we would expect (or not expect) about the number of heads out of 200 flips?

Standard Scores (i.e., z-Scores)

A z-score is computed as

$$z = \frac{x - \bar{x}}{s}.$$

It is the number of standard deviations x is above (if $z > 0$) or below (if $z < 0$) the mean.

Example: Tim is 74 inches tall. Kristine is 70 inches tall. What are their z-scores? Is Tim taller for a man or is Kristine taller for a woman?

Example: A man is five inches above average (for men) in height. A woman is also five inches above average (for women) in height. Who is taller *relative* to the distribution of adults of the same gender?

Example: Consider again the sample of coin flips, but now suppose that we look at the *proportion* of flips that come up heads, and we do so out of 100 flips and again out of 400 flips. The distribution for 100 flips has a mean of 0.5 and a standard deviation of 0.05, while the distribution of 400 flips has a mean of 0.5 and a standard deviation of 0.025. What is the z-score for each distribution if we observe that the coin comes up heads on 55% of the trials (i.e., a proportion of 0.55)?