

Lecture Notes & Examples 2.2 Part B

Section 2.2 Normal Distributions (continued)

****Normal Distribution Calculations****

We will use the previous procedures to answer questions about observations in *any* Normal distribution by *standardizing* and then using the standard Normal table.

4-Step Process

1. *State*:

2. *Plan*:

3. *Do*:

•

•

4. *Conclude*:

Example: In the 2008 Wimbledon tennis tournament, Rafael Nadal averaged 115 miles per hour on his first serves. Assume that the distribution of his first serves is Normal with a mean of 115 mph and a standard deviation of 6 mph. About what proportion of his first serves would you expect to exceed 120 mph?

1. *State*:

2. *Plan*:

3. *Do*:

4. *Conclude*:

Example continued: What percent of Rafael Nadal's first serves are between 100 and 110 mph?

1. State:
2. Plan:
3. Do:
4. Conclude:

Example: According to the Centers for Disease Control (CDC), the heights of three-year-old females are approximately Normally distributed with a mean of 94.5 cm and a standard deviation of 4 cm. What is the third quartile of this distribution?

1. State:
2. Plan:
3. Do:
4. Conclude:

*****Normal Distribution Calculations with Technology*****

Example: Nadal $N(115, 6)$. Find the proportion of first serves we expect to exceed 120 mph.

Example: What percent of Rafael Nadal's first serves are between 100 and 110 mph?

Example: 3-year-olds $N(94.5, 4)$. What is the third quartile of this distribution?

Check Your Understanding. Use the 4-Step Process for each of these. Include a properly labeled diagram.

1. Cholesterol levels in 14-year-old boys is approximately Normally distributed with a mean of 170 mg/dl of blood and standard deviation 30 mg/dl. What percent of 14-year-old boys have more than 240 mg/dl of cholesterol?

2. What percent of 14-year-old boys have blood cholesterol between 200 and 240 mg/dl?

3. What level of cholesterol would represent the 80th percentile?

Assessing Normality

The Normal distributions provide good models for some distributions of real data. In the latter part of this course, we will use various statistical inference procedures to try to answer questions important to us. These tests involve sampling individuals and analyzing data to gain insights about populations. Many of these procedures are based on the assumption that the population is *approximately Normally distributed*. Because of this we need to develop a strategy for assessing Normality.

Procedure.

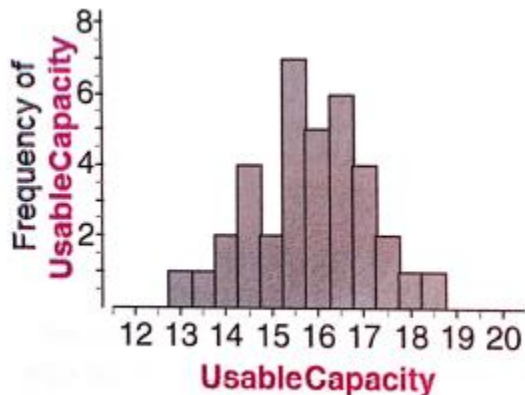
Step 1: *Plot the data* – make a dotplot, stemplot, or histogram. See if the graph is approximately symmetric and bell-shaped. Is the mean close to the median?

Step 2: *Check whether the data follow the 68-95-99.7 rule*. Find the mean and standard deviation. Then count the number of observations within one, two, and three standard deviations from the mean and compute these to percents.

Example. The measurements listed below describe the usable capacity (in cubic feet) of 36 side-by-side refrigerators. Are the data close to Normal?

12.9 13.7 14.1 14.2 14.5 14.5 14.6 14.7 15.1 15.2 15.3 15.3 15.3 15.3 15.5 15.6 15.6 15.8
16.0 16.0 16.2 16.2 16.3 16.4 16.5 16.6 16.6 16.6 16.8 17.0 17.0 17.2 17.4 17.4 17.9 18.4

The mean and standard deviation of these data are 15.825 and 1.217 cubic feet. The histogram is shown below.



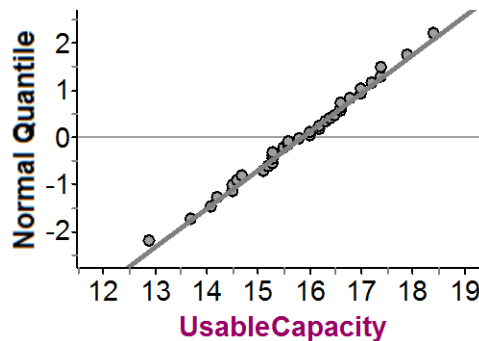
Normal Probability Plot

If a graph of the data is clearly skewed, has multiple peaks, or isn't bell-shaped, that's evidence that the distribution is *not* Normal. However, just because a plot of the data *looks* Normal, we can't say that the distribution *is* Normal. The 68–95–99.7 rule can give additional evidence in favor of or against Normality. A **Normal probability plot** also provides a good assessment of whether a data set follows a Normal distribution.

Normal Probability Plot: a graph that pairs an observation with its expected z-score to decide whether a distribution is approximately Normal

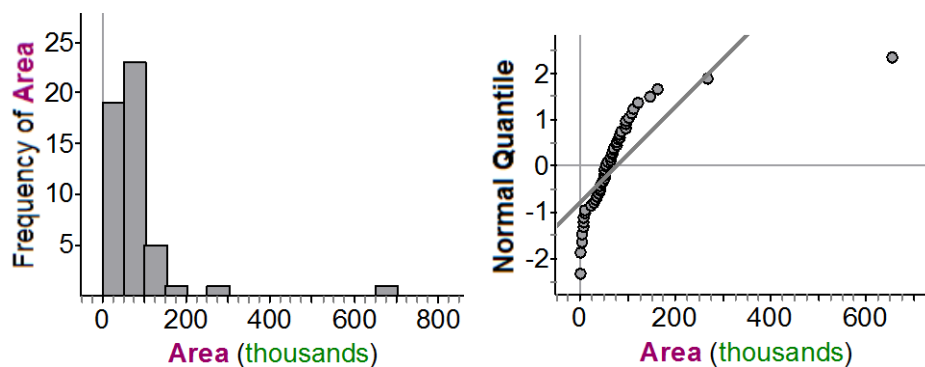
If the points on a Normal Probability Plot lie close to a straight line, the plot indicates that the data are normal. Systematic deviations from a straight line indicate a non-Normal distribution. Outliers appear as points that are far from the overall pattern of the plot.

Example: Here is a Normal probability plot (also called a Normal quantile plot) of the refrigerator data from the previous page. It is quite linear, supporting our earlier decision that the distribution is close to Normal.



Example: State land areas

Problem: The histogram and Normal probability plot below display the land areas for the 50 states. Is this distribution approximately Normal?



Example: NBA free throw percentage

This is an example of a distribution that is skewed to the left. Notice that the lowest free throw percentages are too the left of what we would expect and the highest free throw percentages are not as far to the right as we would expect.

