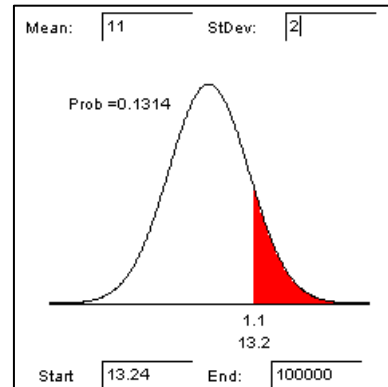


## Working with Normal Distributions

Example: Suppose  $x$  is a normally distributed random variable with  $\mu = 11$  and  $\sigma = 2$ . Find  $P(x > 13.24)$ .

- a. Use Gary McCelland's Applet.



- b. Convert to  $z$ -score and use Table III.

$$z = \frac{x - \mu}{\sigma} = \frac{13.24 - 11}{2} = 1.12.$$

$$\begin{aligned} \text{It follows that } P(x > 13.24) &= P(z > 1.12) \\ &= 0.5 - P(0 < z < 1.12) \\ &\approx 0.5 - 0.3686 \\ &\approx 0.1314. \end{aligned}$$

- c. Use a TI graphing calculator.

Nonstandard Normal Probabilities (See pp.205-206)

Press 2<sup>nd</sup> VARS for DIST and select Normalcdf(.

Enter Normalcdf(13.24, 21, 11, 2)

Press ENTER.

```
normalcdf(13.24,
100,11,2)
.1313569269
```

Standard Normal Probabilities

Press 2<sup>nd</sup> VARS for DIST and select Normalcdf(.

Enter Normalcdf(1.12, 12, 0, 1)

Press ENTER.

```
normalcdf(1.12,1
2,0,1)
.1313569269
```

Standard Normal Probabilities with Graph (See pp. 203-204)

Press WINDOW

Set Xmin = -5; Xmax = 5; Xscl = 1;

Ymin = -.2; Ymax = .5; Yscl = 0; Xres = 1

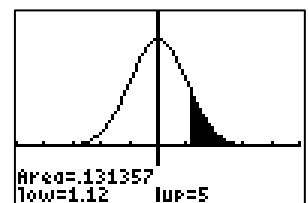
Press 2<sup>nd</sup> VARS

Arrow right to DRAW

Press ENTER and select 1:ShadeNorm(.

Enter 1.12, 5)

Press ENTER



## Assessing Normality

In the weeks ahead we will learn how to make inferences about a population on the basis of information derived from a sample. Some of the techniques we will use assume the population of interest is approximately normally distributed. So, it will be important for us to determine whether it is likely that our sample data comes from a normal population before we can apply those techniques.

Is Our Data from an Approximately Normal Distribution?

1. Examine a histogram or stem-and-leaf display for the data, and note the shape of the graph.
2. Compute the intervals  $\bar{x} \pm s$ ,  $\bar{x} \pm 2s$ ,  $\bar{x} \pm 3s$ , and determine the percent of measurements falling in each interval.
3. Calculate the ratio  $IQR/s$ . If the data are approximately normal then  $IQR/s \approx 1.3$ .
4. Construct a **normal probability plot** for the data. If the data are approximately normal, the points will fall (approximately) on a straight line.

A **normal probability plot** for a data set is a scatterplot with the ranked data values on one axis and their corresponding expected z-scores on the other axis. (We will use statistical packages to generate these plots.)

**Example:** Recall the EPA mileage ratings on 100 cars presented earlier in this course. We shall determine if the EPA mileage ratings are from an approximate normal distribution.

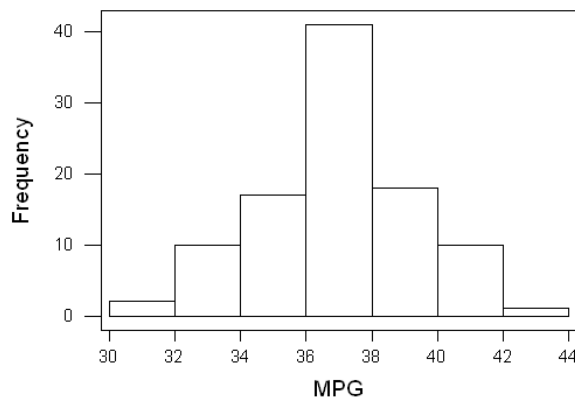
Sorted EPA Mileage Ratings on 100 Cars									
30.0	33.9	35.2	36.1	36.7	37.0	37.4	38.0	39.0	40.2
31.8	33.9	35.3	36.2	36.7	37.0	37.4	38.1	39.0	40.3
32.5	34.0	35.5	36.3	36.7	37.0	37.5	38.2	39.3	40.5
32.7	34.2	35.6	36.3	36.8	37.1	37.6	38.2	39.4	40.5
32.9	34.4	35.6	36.4	36.8	37.1	37.6	38.3	39.5	40.7
32.9	34.5	35.7	36.4	36.8	37.1	37.7	38.4	39.7	41.0
33.1	34.8	35.8	36.5	36.9	37.2	37.7	38.5	39.8	41.0
33.2	34.8	35.9	36.5	36.9	37.2	37.8	38.6	39.9	41.2
33.6	35.0	35.9	36.6	36.9	37.3	37.9	38.7	40.0	42.1
33.8	35.1	36.0	36.6	37.0	37.3	37.9	38.8	40.1	44.9

A Stem and Leaf Display

```

12  33 126899
18  34 024588
29  35 01235667899
49  36 01233445566777888999
(21) 37 000011122334456677899
30  38 0122345678
20  39 00345789
12  40 0123557
5   41 002
2   42 1
1   43
1   44
    
```

A Histogram for 100 EPA Mileage Ratings



An examination of the stem-and-leaf display and the histogram show the EPA ratings closely reflect a mound shaped, symmetric distribution centered around the mean of about 37 mpg.

We also found the following:

**Descriptive Statistics:**

Variable	N	Mean	Median	StDev	Min	Max	Q1	Q3
MPG	100	36.994	37.000	2.418	30.000	44.900	35.625	38.375

About 68% of the MPG values are within one StDev of the mean.

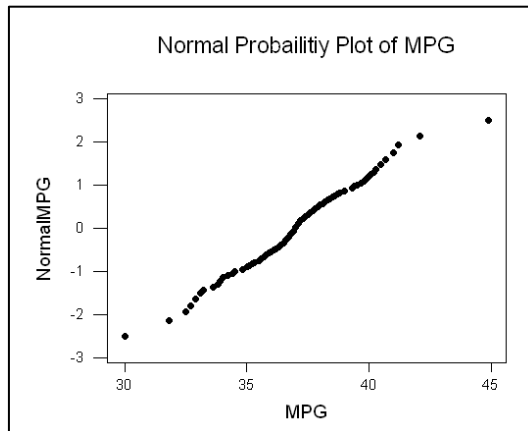
About 95% of the MPG values are within two StDevs of the mean.

About 99% of the MPG values are within three StDevs of the mean.

These percentages agree closely with those from a normal distribution.

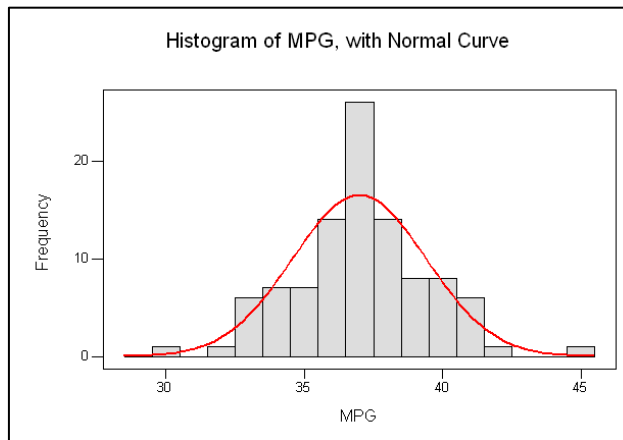
Examining the ratio  $IQR/s$ , we find  $IQR/s \approx 2.75/2.4 \approx 1.15$ . Since that ratio is close to 1.3, we have further confirmation that the data are approximately normal.

As a fourth test we examine a normal probability plot produced by MINITAB.



We note that the points in the plot fall reasonably close to a line. So, we have still further verification that the EPA data are approximately normally distributed.

MINITAB will also produce a histogram of the EPA data with a normal curve superimposed.



Exercise: Determine whether or not the following data is approximately normally distributed: 9.7, 93.1, 33.0, 21.1, 81.4, 51.1, 43.5, 10.6, 12.8, 7.8, 18.1, 12.7.