

Not all Data is Normal

- Modality
- Symmetry
- Peakedness

That is not Normal!

Modality

Unimodal one-peak	Bimodal two-peaks	Multimodal two or more peaks
-----------------------------	-----------------------------	--

Skewness

$$\text{Skewness} = \frac{(\text{Mean} - \text{Median})}{\text{Standard Deviation}}$$

Skewness

$$\text{Skewness} = \frac{(\text{Mean} - \text{Median})}{\text{Standard Deviation}}$$

Skewness = 0
Mean = Median = Mode

Skewness

$$\text{Skewness} = \frac{(\text{Mean} - \text{Median})}{\text{Standard Deviation}}$$

Positively Skewed
Mode < Median < Mean

Skewness = 0
Mean = Median = Mode

This is a "Tail"

Skewness

$$\text{Skewness} = \frac{(\text{Mean} - \text{Median})}{\text{Standard Deviation}}$$

Positively Skewed
Mode < Median < Mean

Skewness = 0
Mean = Median = Mode

Negatively Skewed
Mean < Median < Mode

Kurtosis

Leptokurtic
Peaks sharply with fat tails
Less variability $K > 0$

Mesokurtic
Normal Distribution
 $K = 0$

Platykurtic
Flattened
Highly dispersed
 $K < 0$

EXAMPLE I

On a standardized test, the test scores are normally distributed with a mean of 60 and a standard deviation of 6.

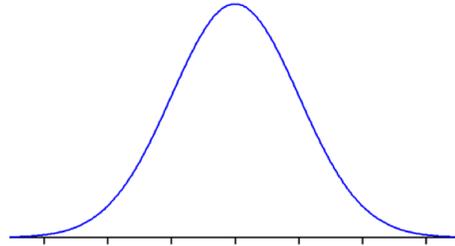
- a. Of the data, 84% of the scores are at or below what score?
- b. Of the data, 16% of the scores are at or below what score?
- c. What is the z-score of a score of 48?
- d. If 2,000 students took the test, how many would be expected to score at or below 48?

$$\text{Zscore} = \frac{X - \bar{X}}{\sigma}$$

Practice Exercises: Normal Distributions

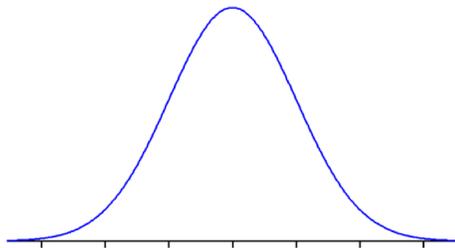
1. The distribution of blood cholesterol levels in 14-year-old boys is roughly normal; the mean is 165 milligrams of cholesterol per deciliter of blood and the standard deviation is 30.

- a) Label the horizontal axis appropriately on the Normal curve at right.
- b) What proportion of 14-year-old boys have level between 120 and 200? Shade the appropriate area in the figure, and then determine the proportion.



- c) What proportion have level *over* 20? (The table can help with this one, but you'll need to use general principles to answer.)

- d) [5] What level separates the boys who have level in the highest 70% from those in the lowest 30%?



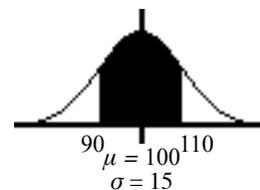
Using Your TI-83/84 Calculator: Normal Probability Distributions Elementary Statistics

Always start by drawing a sketch of the normal distribution that you are working with. Shade in the area (probability) that you are given or trying to find, and label the mean, standard deviation, lower bound, and upper bound that you are given or trying to find. Don't worry about making your drawing to scale; the purpose of the sketch is to get you thinking clearly about the problem you are trying to solve. For illustration purposes, let's consider the distribution of adult scores on the Weschler IQ test. These IQ scores are normally distributed with $\mu = 100$ and $\sigma = 15$.

Using the normalcdf command ← 2nd VARS 2:normalcdf(

The normalcdf command is used for finding an area under the normal density curve. This area corresponds to the probability of randomly selecting a value between the specified lower and upper bounds. You can also interpret this area as the percentage of all values that fall between the two specified boundaries.

- Let's find the percentage of adults who score between 90 and 110 on the Weschler IQ test. Begin by sketching the distribution and labeling the relevant information. We are ultimately trying to find the area under the normal density curve that is bounded by 90 and 110, so shade in that area on your sketch.



- Press **2nd**[VARS] for the [DISTR] menu. Scroll down to 2:normalcdf(and then press **ENTER**.
- Enter the two x values (or z scores) that form the boundaries of the area that you are trying to find, the population mean, and the population standard deviation using the following syntax: lowerbound, upperbound, μ , σ and then press **ENTER**. For this example, type 90, 110, 100, 15) and then press **ENTER**.

```

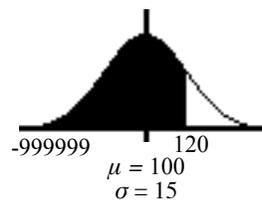
DISTR DRAW
1:normalPdf(
2:normalcdf(
3:invNorm(
4:invT(
5:tPdf(
6:tCdf(
7:χ²Pdf(
    
```

```

normalcdf(90,110
,100,15)
.495015066
    
```

- Your calculator will return the area under the normal curve bounded by 90 and 110. Thus, we find that 49.5% of adults score between 90 and 110 on the Weschler IQ test. Remember to round percentages to three significant figures.

- What is the probability that a randomly selected adult scores less than 120 on the Weschler IQ test? Problems like this are a bit trickier because your calculator requires both a lower bound and an upper bound. Ultimately, you need to specify an approximation of $-\infty$ for the lower bound. In most cases, -999999 is a good choice for the lower bound. Hence, the syntax for problems of this sort is normalcdf(-999999, upperbound, μ , σ).

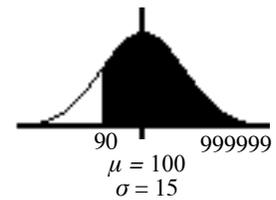


- Start by drawing a sketch. Then, press **2nd**[VARS] for the [DISTR] menu. Scroll down to 2:normalcdf(and press **ENTER**. Type **(-)**999999, 120, 100, 15) and then press **ENTER** again. We find that the probability of randomly selecting an adult whose IQ is below 120 is 0.909. In symbols, $P(x \leq 120) = 0.909$. Remember to round probability values to 3 significant figures.

```

normalcdf(-99999
9,120,100,15)
.9087887181
    
```

- What percentage of adults score at least 90 on the Weschler IQ test? Because your calculator requires both a lower bound and an upper bound, you will need to specify an approximation of ∞ for your upper bound. Using 999999 is good enough in most cases. Hence, the syntax for problems of this sort is `normalcdf(lowerbound, 999999, μ , σ)`.
- Start by drawing a sketch. Then, press `2nd|VARS` for the `[DISTR]` menu. Scroll down to `2:normalcdf(` and press `[ENTER]`. Type `90, 999999, 100, 15)` and then press `[ENTER]` again.
- We find that 74.8% of adults score at least 90 on the Weschler IQ test. Remember to round to 3 significant figures.



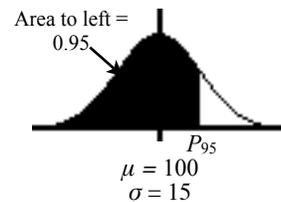
```
normalcdf(90, 999
999, 100, 15)
.747507533
```

Using the `invNorm` command `2nd VARS 3:invNorm(`

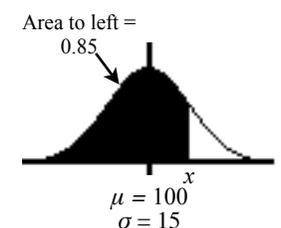
Use the `invNorm` command when you are given a probability or percentage and asked to find an x value or z score. This command is often used to find values corresponding to percentiles or quartiles. Your calculator requires that you enter the cumulative area to the left of the desired value; drawing a sketch is very useful for making sure you enter the correct area. Sometimes you will need to work with an area other than the one specified by the problem.

- Press `2nd|VARS` for the `[DISTR]` menu. Scroll down to `3:invNorm(` and press `[ENTER]`.
- Let's start by finding the IQ score corresponding to the 95th percentile (P_{95}). Begin by drawing a sketch and labeling it.
- Enter the total area to the left of the desired x value (or z score) using the following syntax: `invNorm(area to left, μ , σ)` and then press `[ENTER]`. For this example, type `95, 100, 15)` and then press `[ENTER]`.
- Your calculator will return the x value (or z score). For this example, we find that the Weschler IQ score corresponding to the 95th percentile is 124.7. Round your answer to one more decimal place than what was provided for μ .
- Let's try a trickier example. This time, find the IQ score separating the top 15% of all Weschler IQ scores from the rest. Start by drawing a sketch. Given that the total area under the normal density curve is always 1, the area to the left of the IQ score that we are seeking can be found by subtracting the given area from 1. In this case, $1 - 0.15 = 0.85$. Don't forget to convert percentages to their equivalent decimal values.
- Press `2nd|VARS` for the `[DISTR]` menu. Scroll down to `3:invNorm(` and press `[ENTER]`.
- The area to the left of the IQ score we are seeking is 0.85, so type `0.85, 100, 15)` and then press `[ENTER]`. We find that an IQ score of 115.5 separates the top 15% of all adult Weschler IQ scores from the rest.

```
0:QUIT DRAW
1:normalpdf(
2:normalcdf(
3:invNorm(
4:invT(
5:tpdf(
6:tcdf(
7:Xdpdf(
```



```
invNorm(0.95, 100
, 15)
124.6728044
```



```
invNorm(0.85, 100
, 15)
115.5465007
```

Applying Skills

In 11–14, select the numeral that precedes the choice that best completes the statement or answers the question.

- 11.** The playing life of a Euclid mp3 player is normally distributed with a mean of 30,000 hours and a standard deviation of 500 hours. Matt's mp3 player lasted for 31,500 hours. His mp3 player lasted longer than what percent of other Euclid mp3 players?

(1) 68% (2) 95% (3) 99.7% (4) more than 99.8%
- 12.** The scores of a test are normally distributed. If the mean is 50 and the standard deviation is 8, then a student who scored 38 had a z-score of

(1) 1.5 (2) -1.5 (3) 12 (4) -12
- 13.** The heights of 10-year-old children are normally distributed with a mean of 138 centimeters with a standard deviation of 5 centimeters. The height of a 10-year-old child who is as tall as or taller than 95.6% of all 10-year-old children is

(1) between 138 and 140 cm. (2) between 140 and 145 cm.
(3) between 145 and 148 cm. (4) taller than 148 cm.
- 14.** The heights of 200 women are normally distributed. The mean height is 170 centimeters with a standard deviation of 10 centimeters. What is the best estimate of the number of women in this group who are between 160 and 170 centimeters tall?

(1) 20 (2) 34 (3) 68 (4) 136
- 15.** When coffee is packed by machine into 16-ounce cans, the amount can vary. The mean weight is 16.1 ounces and the standard deviation is 0.04 ounce. The weight of the coffee approximates a normal distribution.

 - a.** What percent of the cans of coffee can be expected to contain less than 16 ounces of coffee?
 - b.** What percent of the cans of coffee can be expected to contain between 16.0 and 16.2 ounces of coffee?

- 16.** The length of time that it takes Ken to drive to work represents a normal distribution with a mean of 25 minutes and a standard deviation of 4.5 minutes. If Ken allows 35 minutes to get to work, what percent of the time can he expect to be late?
- 17.** A librarian estimates that the average number of books checked out by a library patron is 4 with a standard deviation of 2 books. If the number of books checked out each day approximates a normal distribution, what percent of the library patrons checked out more than 7 books yesterday?
- 18.** The heights of a group of women are normally distributed with a mean of 170 centimeters and a standard deviation of 10 centimeters. What is the z-score of a member of the group who is 165 centimeters tall?
- 19.** The test grades for a standardized test are normally distributed with a mean of 50. A grade of 60 represents a z-score of 1.25. What is the standard deviation of the data?
- 20.** Nora scored 88 on a math test that had a mean of 80 and a standard deviation of 5. She also scored 80 on a science test that had a mean of 70 and a standard deviation of 3. On which test did Nora perform better compared with other students who took the tests?