**Joint, Marginal and Conditional Probabilities**

Probabilities may be either marginal, joint or conditional.  Understanding their differences and how to manipulate among them is key to success in understanding the foundations of statistics.

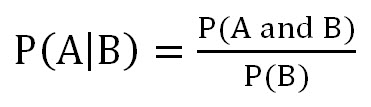
**Marginal probability**: the probability of an event occurring (p(A)), it may be thought of as an unconditional probability.  It is not conditioned on another event.  Example:  the probability that a card drawn is red (p(red) = 0.5).  Another example:  the probability that a card drawn is a 4  (p(four)=1/13).

**Joint probability:**  p(A and B).  The probability of event A and event B occurring.  It is the probability of the intersection of two or more events.  The probability of the intersection of A and B may be written p(A ∩ B). Example:  the probability that a card is a four and red =p(four and red) = 2/52=1/26.  (There are two red fours in a deck of 52, the 4 of hearts and the 4 of diamonds).

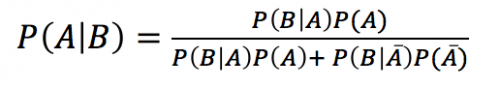
**Conditional probability:**  p(A|B) is the probability of event A occurring, given that event B occurs. Example:  given that you drew a red card, what’s the probability that it’s a four (p(four|red))=2/26=1/13.  So out of the 26 red cards (given a red card), there are two fours so 2/26=1/13.

# Manipulate among Joint, Conditional and Marginal Probabilities

### The equation below is a means to manipulate among joint, conditional and marginal probabilities.  As you can see, the conditional probability of A given B is equal to the joint probability of A and B divided by the marginal of B.  Let’s use our card example to illustrate.  We know that the conditional probability of a four, given a red card equals 2/26 or 1/13.  This should be equivalent to the joint probability of a red and four (2/52 or 1/26) divided by the marginal P(red) = 1/2.  And low and behold, it works!  As 1/13 = 1/26 divided by 1/2.  For the diagnostic exam, you should be able to manipulate among joint, marginal and conditional probabilities.

[](http://sites.nicholas.duke.edu/statsreview/files/2013/06/jointmargcond.jpg)

**Bayes’ Theorem** an equation that allows us to manipulate conditional probabilities. For two events, A and B, Bayes’ theorem lets us to go from p(B|A) to p(A|B) if we know themarginal probabilities of the outcomes of A and the probability of B, given the outcomes of A. Here is the equation for Bayes’ theorem for two events with two possible outcome (A and not A).

[](http://sites.nicholas.duke.edu/statsreview/files/2014/08/Screenshot-2014-08-24-22.37.48-e1408934339835.png)

### Let’s assume we know that 1% of women over the age of 40 have breast cancer.

### [p(cancer)=0.01]

### Let’s assume that 90% of women who have breast cancer will test positive for breast cancer in a mammogram.

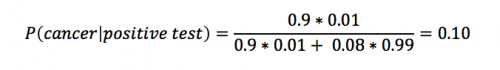
### [p(positive test|cancer)=0.9]

### Eight percent ofwomen that do NOT have cancer will also test positive.

### [p(positive test|no cancer)=0.08]

### What is the probability that a woman has cancer if she tests positive [p(cancer|positive test)]?

### We will call p(cancer) = P(A), and the P(positive test) = P(B). We want to know P(A|B)–the probability of having cancer if you have a positive test.

[](http://sites.nicholas.duke.edu/statsreview/files/2014/08/Screenshot-2014-08-24-22.46.21-e1408934861227.png)

### Using Bayes’ theorem, we calculate that the likelihood that a woman has breast cancer, given a positive test equals approximately 0.10. This makes intuitive sense as (1) this result is greater than 1% (the percent of breast cancer in the general public).

(SOURCE: Elizabeth A. Albright, PhD <http://sites.nicholas.duke.edu/statsreview/probability/jmc/> )

Probability

|  |  |  |
| --- | --- | --- |
| Problem: | A spinner has 4 equal sectors colored yellow, blue, green and red. What are the chances of landing on blue after spinning the spinner? What are the chances of landing on red? | spinner  Top of Form  Bottom of Form |
| Solution: | The chances of landing on blue are 1 in 4, or one fourth. |
|  | The chances of landing on red are 1 in 4, or one fourth. |

|  |
| --- |
| This problem asked us to find some probabilities involving a spinner. Let's look at some definitions and examples from the problem above. |

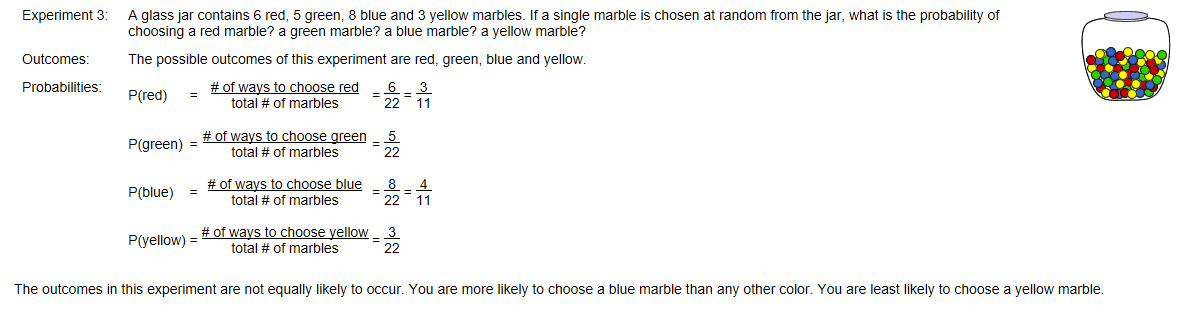
|  |  |
| --- | --- |
| **Definition** | **Example** |
| An **experiment** is a situation involving chance or probability that leads to results called outcomes. | In the problem above, the experiment is spinning the spinner. |
| An **outcome** is the result of a single trial of an experiment. | The possible outcomes are landing on yellow, blue, green or red. |
| An **event** is one or more outcomes of an experiment. | One event of this experiment is landing on blue. |
| **Probability** is the measure of how likely an event is. | The probability of landing on blue is one fourth. |

|  |
| --- |
| In order to measure probabilities, mathematicians have devised the following formula for finding the probability of an event. |

|  |
| --- |
| **Probability Of An Event** |
| |  |  | | --- | --- | | P(A) = | The Number Of Ways Event A Can Occur | | The total number Of Possible Outcomes | |

|  |
| --- |
| **The probability of event A is the number of ways event A can occur divided by the total number of possible outcomes.** Let's take a look at a slight modification of the problem from the top of the page. |

|  |  |  |  |
| --- | --- | --- | --- |
| Experiment 1: | A spinner has 4 equal sectors colored yellow, blue, green and red. After spinning the spinner, what is the probability of landing on each color? |  | spinner |
| Outcomes: | The possible outcomes of this experiment are yellow, blue, green, and red. |
| Probabilities: | |  |  |  |  |  | | --- | --- | --- | --- | --- | | P(yellow) | = | # of ways to land on yellow | = | 1 | | total # of colors | 4 | |  | | | | | | P(blue) | = | # of ways to land on blue | = | 1 | | total # of colors | 4 | |  | | | | | | P(green) | = | # of ways to land on green | = | 1 | | total # of colors | 4 | |  | | | | | | P(red) | = | # of ways to land on red | = | 1 | | total # of colors | 4 | |
|  |  |  |



SOURCE: <http://www.mathgoodies.com/lessons/vol6/intro_probability.html>

A **probability distribution** is a table or an equation that links each possible value that a [random variable](http://stattrek.com/Help/Glossary.aspx?Target=Random%20variable) can assume with its probability of occurrence.

## Discrete Probability Distributions

The probability distribution of a [discrete](http://stattrek.com/Help/Glossary.aspx?Target=Discrete%20variable) random variable can always be represented by a table. For example, suppose you flip a coin two times. This simple exercise can have four possible outcomes: HH, HT, TH, and TT. Now, let the variable X represent the number of heads that result from the coin flips. The variable X can take on the values 0, 1, or 2; and X is a discrete random variable.

The table below shows the probabilities associated with each possible value of X. The probability of getting 0 heads is 0.25; 1 head, 0.50; and 2 heads, 0.25. It is an example of a probability distribution for a discrete random variable.

|  |  |
| --- | --- |
| **Number of heads, x** | **Probability, P(x)** |
| 0 | 0.25 |
| 1 | 0.50 |
| 2 | 0.25 |

**Note:** Given a probability distribution, you can find [cumulative probabilities](http://stattrek.com/Help/Glossary.aspx?Target=Cumulative%20probability). For example, the probability of getting 1 or fewer heads [ P(X < 1) ] is P(X = 0) + P(X = 1), which is equal to 0.25 + 0.50 or 0.75.

## Continuous Probability Distributions

The probability distribution of a [continuous](http://stattrek.com/Help/Glossary.aspx?Target=Continuous%20variable) random variable is represented by an equation, called the **probability density function** (pdf). All probability density functions satisfy the following conditions:

* The random variable Y is a function of X; that is, y = f(x).
* The value of y is greater than or equal to zero for all values of x.
* The total area under the curve of the function is equal to one.

The charts below show two continuous probability distributions. The chart on the left shows a probability density function described by the equation y = 1 over the range of 0 to 1 and y = 0 elsewhere. The chart on the right shows a probability density function described by the equation y = 1 - 0.5x over the range of 0 to 2 and y = 0 elsewhere. The area under the curve is equal to 1 for both charts.

|  |  |
| --- | --- |
| http://stattrek.com/Images/Sp26.jpg | http://stattrek.com/Images/Sp27.jpg |
| y = 1 | y = 1 - 0.5x |

The probability that a continuous random variable falls in the interval between a and b is equal to the area under the pdf curve between a and b. For example, in the first chart above, the shaded area shows the probability that the random variable X will fall between 0.6 and 1.0. That probability is 0.40. And in the second chart, the shaded area shows the probability of falling between 1.0 and 2.0. That probability is 0.25.

**Note:** With a continuous distribution, there are an infinite number of values between any two data points. As a result, the probability that a continuous random variable will assume a particular value is always zero. For example, in both of the above charts, the probability that variable X will equal exactly 0.4 is zero.

# Mean and Variance of Random Variables

Just like [variables](http://stattrek.com/Help/Glossary.aspx?Target=Variable) from a data set, [random variables](http://stattrek.com/Help/Glossary.aspx?Target=Random%20variable) are described by measures of central tendency (like the mean) and measures of variability (like variance). This lesson shows how to compute these measures for [discrete](http://stattrek.com/Help/Glossary.aspx?Target=Discrete%20variable) random variables.

## **Mean of a Discrete Random Variable**

The mean of the discrete random variable X is also called the **expected value** of X. Notationally, the expected value of X is denoted by E(X). Use the following formula to compute the mean of a discrete random variable.

E(X) = μx = Σ [ xi \* P(xi) ]

where xi is the value of the random variable for outcome i, μx is the mean of random variable X, and P(xi) is the probability that the random variable will be outcome i.

**Example 1**

In a recent little league softball game, each player went to bat 4 times. The number of hits made by each player is described by the following probability distribution.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Number of hits, x** | 0 | 1 | 2 | 3 | 4 |
| **Probability, P(x)** | 0.10 | 0.20 | 0.30 | 0.25 | 0.15 |

What is the mean of the probability distribution?

(A) 1.00   
(B) 1.75   
(C) 2.00   
(D) 2.25   
(E) None of the above.

**Solution**

The correct answer is E. The mean of the probability distribution is 2.15, as defined by the following equation.

E(X) = Σ [ xi \* P(xi) ]   
E(X) = 0\*0.10 + 1\*0.20 + 2\*0.30 + 3\*0.25 + 4\*0.15 = 2.15

## Median of a Discrete Random Variable

The median of a discrete random variable is the "middle" value. It is the value of X for which P(X < x) is greater than or equal to 0.5 and P(X > x) is greater than or equal to 0.5.

Consider the problem presented above in Example 1. In Example 1, the median is 2; because P(X < 2) is equal to 0.60, and P(X > 2) is equal to 0.70. The computations are shown below.

P(X < 2) = P(x=0) + P(x=1) + P(x=2) = 0.10 + 0.20 + 0.30 = 0.60   
  
P(X > 2) = P(x=2) + P(x=3) + P(x=4) = 0.30 + 0.25 + 0.15 = 0.70

## Variability of a Discrete Random Variable

The equation for computing the variance of a discrete random variable is shown below.

σ2 = Σ { [ xi - E(x) ]2 \* P(xi) }

where xi is the value of the random variable for outcome i, P(xi) is the probability that the random variable will be outcome i, E(x) is the expected value of the discrete random variable x.

**Example 2**

The number of adults living in homes on a randomly selected city block is described by the following probability distribution.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Number of adults, x** | 1 | 2 | 3 | 4 |
| **Probability, P(x)** | 0.25 | 0.50 | 0.15 | 0.10 |

What is the standard deviation of the probability distribution?

(A) 0.50   
(B) 0.62   
(C) 0.79   
(D) 0.89   
(E) 2.10

**Solution**

The correct answer is D. The solution has three parts. First, find the expected value; then, find the variance; then, find the standard deviation. Computations are shown below, beginning with the expected value.

E(X) = Σ [ xi \* P(xi) ]   
E(X) = 1\*0.25 + 2\*0.50 + 3\*0.15 + 4\*0.10 = 2.10

Now that we know the expected value, we find the variance.

σ2 = Σ { [ xi - E(x) ]2 \* P(xi) }   
σ2 = (1 - 2.1)2 \* 0.25 + (2 - 2.1)2 \* 0.50 + (3 - 2.1)2 \* 0.15 + (4 - 2.1)2 \* 0.10   
σ2 = (1.21 \* 0.25) + (0.01 \* 0.50) + (0.81) \* 0.15) + (3.61 \* 0.10) = 0.3025 + 0.0050 + 0.1215 + 0.3610 = 0.79

And finally, the standard deviation is equal to the square root of the variance; sqrt(0.79) or 0.889.

# Independent Random Variables

When a study involves pairs of [random variables](http://stattrek.com/Help/Glossary.aspx?Target=random_variable), it is often useful to know whether or not the random variables are independent. This lesson explains how to assess the independence of random variables.

## Independence of Random Variables

If two random variables, X and Y, are **independent**, they satisfy the following conditions.

* P(x|y) = P(x), for all values of X and Y.
* P(x ∩ y) = P(x) \* P(y), for all values of X and Y.

The above conditions are equivalent. If either one is met, the other condition also met; and X and Y are independent. If either condition is not met, X and Y are **dependent**.

**Note:** If X and Y are independent, then the [correlation](http://stattrek.com/Help/Glossary.aspx?Target=Correlation) between X and Y is equal to zero.

## Joint Probability Distributions

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | X | | |
| 0 | 1 | 2 |
| Y | 3 | 0.1 | 0.2 | 0.2 |
| 4 | 0.1 | 0.2 | 0.2 |

The table on the right shows the joint probability distribution between two discrete random variables - X and Y.

In a joint probability distribution table, numbers in the cells of the table represent the probability that particular values of X and Y occur together. From this table, you can see that the probability that X=0 and Y=3 is 0.1; the probability that X=1 and Y=3 is 0.2; and so on.

You can use tables like this to figure out whether two discrete random variables are independent or dependent. Problem 1 below shows how.

## Test Your Understanding

**Problem 1**

The table on the left shows the joint probability distribution between two random variables - X and Y; and the table on the right shows the joint probability distribution between two random variables - A and B.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | | X | | | | 0 | 1 | 2 | | Y | 3 | 0.1 | 0.2 | 0.2 | | 4 | 0.1 | 0.2 | 0.2 | | |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | | A | | | | 0 | 1 | 2 | | B | 3 | 0.1 | 0.2 | 0.2 | | 4 | 0.2 | 0.2 | 0.1 | |

Which of the following statements are true?

I. X and Y are independent random variables.   
II. A and B are independent random variables.

(A) I only   
(B) II only   
(C) I and II   
(D) Neither statement is true.   
(E) It is not possible to answer this question, based on the information given.

**Solution**

The correct answer is A. The solution requires several computations to test the independence of random variables. Those computations are shown below. X and Y are independent if P(x|y) = P(x), for all values of X and Y. From the probability distribution table, we know the following:

P(x=0) = 0.2;      P(x=0 | y=3) = 0.2;      P(x=0 | y = 4) = 0.2   
P(x=1) = 0.4;      P(x=1 | y=3) = 0.4;      P(x=1 | y = 4) = 0.4   
P(x=2) = 0.4;      P(x=2 | y=3) = 0.4;      P(x=2 | y = 4) = 0.4

Thus, P(x|y) = P(x), for all values of X and Y, which means that X and Y are independent. We repeat the same analysis to test the independence of A and B.

P(a=0) = 0.3;      P(a=0 | b=3) = 0.2;      P(a=0 | b = 4) = 0.4   
P(a=1) = 0.4;      P(a=1 | b=3) = 0.4;      P(a=1 | b = 4) = 0.4   
P(a=2) = 0.3;      P(a=2 | b=3) = 0.4;      P(a=2 | b = 4) = 0.2

Thus, P(a|b) is not equal to P(a), for all values of A and B. For example, P(a=0) = 0.3; but P(a=0 | b=3) = 0.2. This means that A and B are not independent.

# Simple Random Sampling

Simple random sampling is the most widely-used probability sampling method, probably because it is easy to implement and easy to analyze. To understand simple random sampling, you need to first understand a few key definitions.

* The total [set](http://stattrek.com/Help/Glossary.aspx?Target=Set) of observations that can be made is called the **population**.
* A **sample** is a set of observations drawn from a population.
* A **parameter** is a measurable characteristic of a population, such as a [mean](http://stattrek.com/Help/Glossary.aspx?Target=Mean) or [standard deviation](http://stattrek.com/Help/Glossary.aspx?Target=Standard_deviation).
* A **statistic** is a measurable characteristic of a sample, such as a mean or standard deviation.
* A **sampling method** is a procedure for selecting sample elements from a population.
* A **random number** is a number determined totally by chance, with no predictable relationship to any other number.
* A **random number table** is a list of numbers, composed of the digits 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9. Numbers in the list are arranged so that each digit has no predictable relationship to the digits that preceded it or to the digits that followed it. In short, the digits are arranged randomly.

## **Simple Random Sampling** refers to a sampling method that has the following properties.

* The population consists of *N* objects.
* The sample consists of *n* objects.
* All possible samples of *n* objects are equally likely to occur.

An important benefit of simple random sampling is that it allows researchers to use statistical methods to analyze sample results. For example, given a simple random sample, researchers can use statistical methods to define a [confidence interval](http://stattrek.com/statistics/dictionary.aspx?definition=confidence_interval) around a sample mean. Statistical analysis is not appropriate when non-random sampling methods are used.

There are many ways to obtain a simple random sample. One way would be the lottery method. Each of the *N* population members is assigned a unique number. The numbers are placed in a bowl and thoroughly mixed. Then, a blind-folded researcher selects *n* numbers. Population members having the selected numbers are included in the sample.

## **Sampling With Replacement and Without Replacement**

Suppose we use the lottery method described above to select a simple random sample. After we pick a number from the bowl, we can put the number aside or we can put it back into the bowl. If we put the number back in the bowl, it may be selected more than once; if we put it aside, it can selected only one time.

When a population element can be selected more than one time, we are **sampling with replacement**. When a population element can be selected only one time, we are **sampling without replacement**.

# Measures of Central Tendency

Researchers are often interested in defining a value that best describes some attribute of the population. Often this attribute is a measure of central tendency or a proportion.

## Measures of Central Tendency

Several different measures of central tendency are defined below.

* The **mode** is the most frequently appearing value in the population or sample. Suppose we draw a sample of five women and measure their weights. They weigh 100 pounds, 100 pounds, 130 pounds, 140 pounds, and 150 pounds. Since more women weigh 100 pounds than any other weight, the mode would equal 100 pounds.
* To find the **median**, we arrange the observations in order from smallest to largest value. If there is an odd number of observations, the median is the middle value. If there is an even number of observations, the median is the average of the two middle values. Thus, in the sample of five women, the median value would be 130 pounds; since 130 pounds is the middle weight.
* The **mean** of a sample or a population is computed by adding all of the observations and dividing by the number of observations. Returning to the example of the five women, the mean weight would equal (100 + 100 + 130 + 140 + 150)/5 = 620/5 = 124 pounds.

# How to Measure Variability in a Data Set

In this lesson, we discuss three measures that are used to quantify the amount of variation in a data set - the range, the variance, and the standard deviation.

For example, consider a population of elements {5, 5 ,5, 5}. Here, each of the values in the data set are equal, so there is no variation. The set {3, 5, 5, 7}, on the other hand, has some variation since some some elements in the data set have different values.

## **Notation**

The following notation is helpful, when we talk about variability.

* σ2: The variance of the population.
* σ: The standard deviation of the population.
* s2: The variance of the sample.
* s: The standard deviation of the sample.
* μ: The population [mean](http://stattrek.com/Help/Glossary.aspx?Target=Mean).
* x: The sample mean.
* N: Number of observations in the population.
* n: Number of observations in the sample.
* P: The proportion of elements in the population that has a particular attribute.
* p: The proportion of elements in the sample that has a particular attribute.
* Q: The proportion of elements in the population that does not have a specified attribute. Note that Q = 1 - P.
* q: The proportion of elements in the sample that does not have a specified attribute. Note that q = 1 - p.

Note that capital letters refer to population [parameters](http://stattrek.com/Help/Glossary.aspx?Target=Parameter), and lower-case letters refer to sample [statistics](http://stattrek.com/Help/Glossary.aspx?Target=Statistic).

## **The Range**

The **range** is the simplest measure of variation. It is difference between the biggest and smallest random variable.

Range = Maximum value - Minimum value

Therefore, the range of the four random variables (3, 5, 5, 7} would be 7 minus 3 or 4.

## Variance of the Mean

It is important to distinguish between the variance of a population mean and the variance of a sample mean. They have different notation, and they are computed differently. The variance of a population mean is denoted by σ2; and the variance of a sample mean, by *s*2.

The **variance** of a population mean is the average squared deviation from the population mean, as defined by the following formula:

σ2 = Σ ( Xi - μ )2 / N

where σ2 is the population variance, μ is the population mean, Xi is the *i*th element from the population, and N is the number of elements in the population.

The variance of a sample mean is defined by slightly different formula:

*s*2 = Σ ( xi - x )2 / ( n - 1 )

where *s*2 is the sample variance, x is the sample mean, xi is the *i*th element from the sample, and n is the number of elements in the sample. If you are working with a simple random sample, the sample variance can be considered an unbiased estimate of the true population variance. Therefore, if you want to estimate the unknown population variance, based on known data from a simple random sample, use this formula.

**Example 1**  
A population consists of four observations: {1, 3, 5, 7}. What is the variance?

*Solution:* First, we need to compute the population mean.

μ = ( 1 + 3 + 5 + 7 ) / 4 = 4

Then we plug all of the known values in to formula for the variance of a population, as shown below:

σ2 = Σ ( Xi - μ )2 / N

σ2 = [ ( 1 - 4 )2 + ( 3 - 4 )2 + ( 5 - 4 )2 + ( 7 - 4 )2 ] / 4

σ2 = [ ( -3 )2 + ( -1 )2 + ( 1 )2 + ( 3 )2 ] / 4

σ2 = [ 9 + 1 + 1 + 9 ] / 4 = 20 / 4 = 5

**Example 2**  
A simple random sample consists of four observations: {1, 3, 5, 7}. What is the best estimate of the population variance?

*Solution:* This problem is handled exactly like the previous problem, except that we use the formula for calculating sample variance, rather than the formula for calculating population variance.

*s*2 = Σ ( xi - x )2 / ( n - 1 )

*s*2 = [ ( 1 - 4 )2 + ( 3 - 4 )2 + ( 5 - 4 )2 + ( 7 - 4 )2 ] / ( 4 - 1 )

*s*2 = [ ( -3 )2 + ( -1 )2 + ( 1 )2 + ( 3 )2 ] / 3

*s*2 = [ 9 + 1 + 1 + 9 ] / 3 = 20 / 3 = 6.667

# Sampling Distributions

Suppose that we draw all possible samples of size *n* from a given population. Suppose further that we compute a [statistic](http://stattrek.com/Help/Glossary.aspx?Target=Statistic) (e.g., a mean, proportion, standard deviation) for each sample. The [probability distribution](http://stattrek.com/Help/Glossary.aspx?Target=Probability_distribution) of this statistic is called a **sampling distribution**. And the standard deviation of this statistic is called the **standard error**.

## Variability of a Sampling Distribution

The variability of a sampling distribution is measured by its [variance](http://stattrek.com/Help/Glossary.aspx?Target=Variance) or its [standard deviation](http://stattrek.com/Help/Glossary.aspx?Target=Standard%20deviation). The variability of a sampling distribution depends on three factors:

* N: The number of observations in the population.
* n: The number of observations in the sample.
* The way that the random sample is chosen.

If the population size is much larger than the sample size, then the sampling distribution has roughly the same standard error, whether we sample [with](http://stattrek.com/Help/Glossary.aspx?Target=Sampling_with_replacement) or [without replacement](http://stattrek.com/Help/Glossary.aspx?Target=Sampling_without_replacement). On the other hand, if the sample represents a significant fraction (say, 1/20) of the population size, the standard error will be meaningfully smaller, when we sample without replacement.

## **Sampling Distribution of the Mean**

Suppose we draw all possible samples of size *n* from a population of size *N*. Suppose further that we compute a mean score for each sample. In this way, we create a sampling distribution of the mean.

We know the following about the sampling distribution of the mean. The mean of the sampling distribution (μx) is equal to the mean of the population (μ). And the standard error of the sampling distribution (σx) is determined by the standard deviation of the population (σ), the population size (N), and the sample size (n). These relationships are shown in the equations below:

μx = μ      and      σx = [ σ / sqrt(n) ] \* sqrt[ (N - n ) / (N - 1) ]

In the standard error formula, the factor sqrt[ (N - n ) / (N - 1) ] is called the finite population correction or fpc. When the population size is very large relative to the sample size, the fpc is approximately equal to one; and the standard error formula can be approximated by:

σx = σ / sqrt(n).

You often see this "approximate" formula in introductory statistics texts. As a general rule, it is safe to use the approximate formula when the sample size is no bigger than 1/20 of the population size.

## **Sampling Distribution of the Proportion**

In a population of size *N*, suppose that the probability of the occurrence of an event (dubbed a "success") is P; and the probability of the event's non-occurrence (dubbed a "failure") is Q. From this population, suppose that we draw all possible samples of size *n*. And finally, within each sample, suppose that we determine the proportion of successes *p* and failures *q*. In this way, we create a sampling distribution of the proportion.

We find that the mean of the sampling distribution of the proportion (μp) is equal to the probability of success in the population (P). And the standard error of the sampling distribution (σp) is determined by the standard deviation of the population (σ), the population size, and the sample size. These relationships are shown in the equations below:

μp = P

σp = [ σ / sqrt(n) ] \* sqrt[ (N - n ) / (N - 1) ]

σp = sqrt[ PQ/n ] \* sqrt[ (N - n ) / (N - 1) ]

where σ = sqrt[ PQ ].

Like the formula for the standard error of the mean, the formula for the standard error of the proportion uses the finite population correction, sqrt[ (N - n ) / (N - 1) ]. When the population size is very large relative to the sample size, the fpc is approximately equal to one; and the standard error formula can be approximated by:

σp = sqrt[ PQ/n ]

You often see this "approximate" formula in introductory statistics texts. As a general rule, it is safe to use the approximate formula when the sample size is no bigger than 1/20 of the population size.

## **Central Limit Theorem**

The **central limit theorem** states that the sampling distribution of the mean of any [independent](http://stattrek.com/help/glossary.aspx?target=independent), [random variable](http://stattrek.com/help/glossary.aspx?target=random_variable) will be normal or nearly normal, if the sample size is large enough.

How large is "large enough"? The answer depends on two factors.

* Requirements for accuracy. The more closely the sampling distribution needs to resemble a normal distribution, the more sample points will be required.
* The shape of the underlying population. The more closely the original population resembles a normal distribution, the fewer sample points will be required.

In practice, some statisticians say that a sample size of 30 is large enough when the population distribution is roughly bell-shaped. Others recommend a sample size of at least 40. But if the original population is distinctly not normal (e.g., is badly skewed, has multiple peaks, and/or has outliers), researchers like the sample size to be even larger.

## **T-Distribution vs. Normal Distribution**

The t distribution and the normal distribution can both be used with statistics that have a bell-shaped distribution. This suggests that we might use either the t-distribution or the normal distribution to analyze sampling distributions. Which should we choose?

Guidelines exist to help you make that choice. Some focus on the population standard deviation.

* If the population standard deviation is known, use the normal distribution
* If the population standard deviation is unknown, use the t-distribution.

Other guidelines focus on sample size.

* If the sample size is large, use the normal distribution. (See the discussion above in the section on the Central Limit Theorem to understand what is meant by a "large" sample.)
* If the sample size is small, use the t-distribution.

In practice, researchers employ a mix of the above guidelines. On this site, we use the normal distribution when the population standard deviation is known and the sample size is large. We might use either distribution when standard deviation is unknown and the sample size is very large. We use the t-distribution when the sample size is small, unless the underlying distribution is not normal. The t distribution should not be used with small samples from populations that are not approximately normal.

**Example 1**  
Assume that a school district has 10,000 6th graders. In this district, the average weight of a 6th grader is 80 pounds, with a standard deviation of 20 pounds. Suppose you draw a random sample of 50 students. What is the probability that the average weight of a sampled student will be less than 75 pounds?

*Solution:* To solve this problem, we need to define the sampling distribution of the mean. Because our sample size is greater than 30, the Central Limit Theorem tells us that the sampling distribution will approximate a normal distribution.

To define our normal distribution, we need to know both the mean of the sampling distribution and the standard deviation. Finding the mean of the sampling distribution is easy, since it is equal to the mean of the population. Thus, the mean of the sampling distribution is equal to 80.

The standard deviation of the sampling distribution can be computed using the following formula.

σx = [ σ / sqrt(n) ] \* sqrt[ (N - n ) / (N - 1) ]   
σx = [ 20 / sqrt(50) ] \* sqrt[ (10,000 - 50 ) / (10,000 - 1) ] = (20/7.071) \* (0.995) = 2.81

Let's review what we know and what we want to know. We know that the sampling distribution of the mean is normally distributed with a mean of 80 and a standard deviation of 2.82. We want to know the probability that a sample mean is less than or equal to 75 pounds.

Because we know the population standard deviation and the sample size is large, we'll use the normal distribution to find probability. To solve the problem, we plug these inputs into the Normal Probability Calculator: mean = 80, standard deviation = 2.81, and normal random variable = 75. The Calculator tells us that the probability that the average weight of a sampled student is less than 75 pounds is equal to 0.038.

Note: Since the population size is more than 20 times greater than the sample size, we could have used the "approximate" formula σx = [ σ / sqrt(n) ] to compute the standard error. Had we done that, we would have found a standard error equal to [ 20 / sqrt(50) ] or 2.83.

**Example 2**  
Find the probability that of the next 120 births, no more than 40% will be boys. Assume equal probabilities for the births of boys and girls. Assume also that the number of births in the population (N) is very large, essentially infinite.

*Solution:* The Central Limit Theorem tells us that the proportion of boys in 120 births will be approximately normally distributed.

The mean of the sampling distribution will be equal to the mean of the population distribution. In the population, half of the births result in boys; and half, in girls. Therefore, the probability of boy births in the population is 0.50. Thus, the mean proportion in the sampling distribution should also be 0.50.

The standard deviation of the sampling distribution (i.e., the standard error) can be computed using the following formula.

σp = sqrt[ PQ/n ] \* sqrt[ (N - n ) / (N - 1) ]

Here, the finite population correction is equal to 1.0, since the population size (N) was assumed to be infinite. Therefore, standard error formula reduces to:

σp = sqrt[ PQ/n ]   
σp = sqrt[ (0.5)(0.5)/120 ] = sqrt[0.25/120 ] = 0.04564

Let's review what we know and what we want to know. We know that the sampling distribution of the proportion is normally distributed with a mean of 0.50 and a standard deviation of 0.04564. We want to know the probability that no more than 40% of the sampled births are boys.

Because we know the population standard deviation and the sample size is large, we'll use the normal distribution to find probability. To solve the problem, we plug these inputs into the Normal Probability Calculator: mean = .5, standard deviation = 0.04564, and the normal random variable = .4. The Calculator tells us that the probability that no more than 40% of the sampled births are boys is equal to 0.014.

Note: This problem can also be treated as a [binomial experiment](http://stattrek.com/Help/Glossary.aspx?Target=Binomial%20experiment). Elsewhere, we showed [how to analyze a binomial experiment](http://stattrek.com/Lesson2/Binomial.aspx?Tutorial=AP). The binomial experiment is actually the more exact analysis. It produces a probability of 0.018 (versus a probability of 0.14 that we found using the normal distribution). Without a computer, the binomial approach is computationally demanding. Therefore, many statistics texts emphasize the approach presented above, which uses the normal distribution to approximate the binomial.

# What is a Probability Distribution?

A probability distribution is a table or an equation that links each outcome of a statistical experiment with its probability of occurrence.

* A **variable** is a symbol (*A*, *B*, *x*, *y*, etc.) that can take on any of a specified set of values.
* When the value of a variable is the outcome of a [statistical experiment](http://stattrek.com/Help/Glossary.aspx?Target=Statistical_experiment), that variable is a **random variable**.

Generally, statisticians use a capital letter to represent a random variable and a lower-case letter, to represent one of its values. For example,

* X represents the random variable X.
* P(X) represents the probability of X.
* P(X = x) refers to the probability that the random variable X is equal to a particular value, denoted by x. As an example, P(X = 1) refers to the probability that the random variable X is equal to 1.

## **Probability Distributions**

An example will make clear the relationship between random variables and probability distributions. Suppose you flip a coin two times. This simple statistical experiment can have four possible outcomes: HH, HT, TH, and TT. Now, let the variable X represent the number of Heads that result from this experiment. The variable X can take on the values 0, 1, or 2. In this example, X is a random variable; because its value is determined by the outcome of a statistical experiment.

A **probability distribution** is a table or an equation that links each outcome of a statistical experiment with its probability of occurrence. Consider the coin flip experiment described above. The table below, which associates each outcome with its probability, is an example of a probability distribution.

|  |  |
| --- | --- |
| **Number of heads** | **Probability** |
| 0 | 0.25 |
| 1 | 0.50 |
| 2 | 0.25 |

The above table represents the probability distribution of the random variable X.

## Cumulative Probability Distributions

A **cumulative probability** refers to the probability that the value of a random variable falls within a specified range.

Let us return to the coin flip experiment. If we flip a coin two times, we might ask: What is the probability that the coin flips would result in one or fewer heads? The answer would be a cumulative probability. It would be the probability that the coin flip experiment results in zero heads plus the probability that the experiment results in one head.

P(X < 1) = P(X = 0) + P(X = 1) = 0.25 + 0.50 = 0.75

Like a probability distribution, a cumulative probability distribution can be represented by a table or an equation. In the table below, the cumulative probability refers to the probability than the random variable X is less than or equal to x.

|  |  |  |
| --- | --- | --- |
| **Number of heads: x** | **Probability: P(X = x)** | **Cumulative Probability: P(X < x)** |
| 0 | 0.25 | 0.25 |
| 1 | 0.50 | 0.75 |
| 2 | 0.25 | 1.00 |

## **Uniform Probability Distribution**

The simplest probability distribution occurs when all of the values of a random variable occur with equal probability. This probability distribution is called the **uniform distribution**.

**Uniform Distribution.** Suppose the random variable X can assume k different values. Suppose also that the P(X = xk) is constant. Then,

P(X = xk) = 1/k

**Example 1**  
Suppose a die is tossed. What is the probability that the die will land on 5 ?

*Solution:* When a die is tossed, there are 6 possible outcomes represented by: S = { 1, 2, 3, 4, 5, 6 }. Each possible outcome is a random variable (X), and each outcome is equally likely to occur. Thus, we have a uniform distribution. Therefore, the P(X = 5) = 1/6.

**Example 2**  
Suppose we repeat the dice tossing experiment described in Example 1. This time, we ask what is the probability that the die will land on a number that is smaller than 5 ?

*Solution:* When a die is tossed, there are 6 possible outcomes represented by: S = { 1, 2, 3, 4, 5, 6 }. Each possible outcome is equally likely to occur. Thus, we have a uniform distribution.

This problem involves a cumulative probability. The probability that the die will land on a number smaller than 5 is :

P( X < 5 ) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) = 1/6 + 1/6 + 1/6 + 1/6 = 2/3

# Probability Distributions: Discrete vs. Continuous

All probability distributions can be classified as **discrete probability distributions** or as **continuous probability distributions**, depending on whether they define probabilities associated with discrete variables or continuous variables.

## Discrete vs. Continuous Variables

If a [variable](http://stattrek.com/Help/Glossary.aspx?Target=Variable) can take on any value between two specified values, it is called a **continuous variable**; otherwise, it is called a **discrete variable**.

Some examples will clarify the difference between discrete and continuous variables.

* Suppose the fire department mandates that all fire fighters must weigh between 150 and 250 pounds. The weight of a fire fighter would be an example of a continuous variable; since a fire fighter's weight could take on any value between 150 and 250 pounds.
* Suppose we flip a coin and count the number of heads. The number of heads could be any integer value between 0 and plus infinity. However, it could not be any number between 0 and plus infinity. We could not, for example, get 2.5 heads. Therefore, the number of heads must be a discrete variable.

Just like variables, [probability distributions](http://stattrek.com/Help/Glossary.aspx?Target=Probability_distribution) can be classified as discrete or continuous.

## **Discrete Probability Distributions**

If a [random variable](http://stattrek.com/Help/Glossary.aspx?Target=Random_variable) is a discrete variable, its [probability distribution](http://stattrek.com/Help/Glossary.aspx?Target=Probability_distribution) is called a **discrete probability distribution**.

An example will make this clear. Suppose you flip a coin two times. This simple [statistical experiment](http://stattrek.com/Help/Glossary.aspx?Target=Statistical_experiment) can have four possible outcomes: HH, HT, TH, and TT. Now, let the random variable X represent the number of Heads that result from this experiment. The random variable X can only take on the values 0, 1, or 2, so it is a discrete random variable.

The probability distribution for this statistical experiment appears below.

|  |  |
| --- | --- |
| **Number of heads** | **Probability** |
| 0 | 0.25 |
| 1 | 0.50 |
| 2 | 0.25 |

The above table represents a *discrete* probability distribution because it relates each value of a discrete random variable with its probability of occurrence. In subsequent lessons, we will cover the following discrete probability distributions.

* [Binomial probability distribution](http://stattrek.com/Lesson2/Binomial.aspx)
* [Hypergeometric probability distribution](http://stattrek.com/Lesson2/Hypergeometric.aspx)
* [Multinomial probability distribution](http://stattrek.com/Lesson2/Multinomial.aspx)
* [Negative binomial distribution](http://stattrek.com/online-calculator/negative-binomial.aspx)
* [Poisson probability distribution](http://stattrek.com/Lesson2/Poisson.aspx)

**Note:** With a discrete probability distribution, each possible value of the discrete random variable can be associated with a non-zero probability. Thus, a discrete probability distribution can always be presented in tabular form.

## Continuous Probability Distributions

If a [random variable](http://stattrek.com/Help/Glossary.aspx?Target=Random_variable) is a continuous variable, its [probability distribution](http://stattrek.com/Help/Glossary.aspx?Target=Probability_distribution) is called a **continuous probability distribution**.

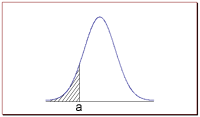
A continuous probability distribution differs from a discrete probability distribution in several ways.

* The probability that a continuous random variable will assume a particular value is zero.
* As a result, a continuous probability distribution cannot be expressed in tabular form.
* Instead, an equation or formula is used to describe a continuous probability distribution.

Most often, the equation used to describe a continuous probability distribution is called a **probability density function**. Sometimes, it is referred to as a **density function**, a **PDF**, or a **pdf**. For a continuous probability distribution, the density function has the following properties:

* Since the continuous random variable is defined over a continuous range of values (called the **domain** of the variable), the graph of the density function will also be continuous over that range.
* The area bounded by the curve of the density function and the x-axis is equal to 1, when computed over the domain of the variable.
* The probability that a random variable assumes a value between *a* and *b* is equal to the area under the density function bounded by *a* and *b*.

For example, consider the probability density function shown in the graph below. Suppose we wanted to know the probability that the random variable *X* was less than or equal to *a*. The probability that *X* is less than or equal to *a* is equal to the area under the curve bounded by *a* and minus infinity - as indicated by the shaded area.



**Note:** The shaded area in the graph represents the probability that the random variable *X* is less than or equal to *a*. This is a [cumulative probability](http://stattrek.com/Help/Glossary.aspx?Target=Cumulative_probability). However, the probability that *X* is *exactly* equal to *a* would be zero. A continuous random variable can take on an infinite number of values. The probability that it will equal a specific value (such as *a*) is always zero.

# Binomial Probability Distribution

To understand binomial distributions and binomial probability, it helps to understand binomial experiments and some associated notation; so we cover those topics first.

## Binomial Experiment

A **binomial experiment** is a [statistical experiment](http://stattrek.com/Help/Glossary.aspx?Target=Statistical_experiment) that has the following properties:

* The experiment consists of *n* repeated trials.
* Each trial can result in just two possible outcomes - a success and the other, a failure.
* The probability of success, denoted by *P*, is the same on every trial.
* The trials are [independent](http://stattrek.com/Help/Glossary.aspx?Target=Independent); that is, the outcome on one trial does not affect the outcome on other trials.

Consider the following statistical experiment. You flip a coin 2 times and count the number of times the coin lands on heads. This is a binomial experiment because:

* The experiment consists of repeated trials. We flip a coin 2 times.
* Each trial can result in just two possible outcomes - heads or tails.
* The probability of success is constant - 0.5 on every trial.
* The trials are independent; that is, getting heads on one trial does not affect whether we get heads on other trials.

## Notation

The following notation is helpful, when we talk about binomial probability.

* *x*: The number of successes that result from the binomial experiment.
* *n*: The number of trials in the binomial experiment.
* *P*: The probability of success on an individual trial.
* *Q*: The probability of failure on an individual trial. (This is equal to 1 - *P*.)
* *n!*: The [factorial](http://stattrek.com/statistics/dictionary.aspx?definition=factorial) of n (also known as n factorial).
* b(*x*; *n, P*): Binomial probability - the probability that an *n*-trial binomial experiment results in exactly *x* successes, when the probability of success on an individual trial is *P*.
* nCr: The number of [combinations](http://stattrek.com/Help/Glossary.aspx?Target=Combination) of *n* things, taken *r* at a time.

## Binomial Distribution

A **binomial random variable** is the number of successes *x* in *n* repeated trials of a binomial experiment. The [probability distribution](http://stattrek.com/Help/Glossary.aspx?Target=Probability_distribution) of a binomial random variable is called a **binomial distribution**.

Suppose we flip a coin two times and count the number of heads (successes). The binomial random variable is the number of heads, which can take on values of 0, 1, or 2. The binomial distribution is presented below.

|  |  |
| --- | --- |
| **Number of heads** | **Probability** |
| 0 | 0.25 |
| 1 | 0.50 |
| 2 | 0.25 |

The binomial distribution has the following properties:

* The mean of the distribution (μx) is equal to *n* \* *P* .
* The [variance](http://stattrek.com/Help/Glossary.aspx?Target=Variance) (σ2x) is *n* \* *P* \* ( 1 - *P* ).
* The [standard deviation](http://stattrek.com/Help/Glossary.aspx?Target=Standard%20deviation) (σx) is sqrt[ *n* \* *P* \* ( 1 - *P* ) ].

## Binomial Formula and Binomial Probability

The **binomial probability** refers to the probability that a binomial experiment results in exactly *x* successes. For example, in the above table, we see that the binomial probability of getting exactly one head in two coin flips is 0.50.

Given *x*, *n*, and *P*, we can compute the binomial probability based on the binomial formula:

**Binomial Formula.** Suppose a binomial experiment consists of *n* trials and results in *x* successes. If the probability of success on an individual trial is *P*, then the binomial probability is:

b(*x*; *n, P*) = nCx \* Px \* (1 - P)n - x   
or   
b(*x*; *n, P*) = { n! / [ x! (n - x)! ] } \* Px \* (1 - P)n - x

**Example 1**  
Suppose a die is tossed 5 times. What is the probability of getting exactly 2 fours?

*Solution:* This is a binomial experiment in which the number of trials is equal to 5, the number of successes is equal to 2, and the probability of success on a single trial is 1/6 or about 0.167. Therefore, the binomial probability is:

b(2; 5, 0.167) = 5C2 \* (0.167)2 \* (0.833)3   
b(2; 5, 0.167) = 0.161

# Poisson Distribution

A Poisson distribution is the probability distribution that results from a Poisson experiment.

## Attributes of a Poisson Experiment

A **Poisson experiment** is a [statistical experiment](http://stattrek.com/Help/Glossary.aspx?Target=Statistical_experiment) that has the following properties:

* The experiment results in outcomes that can be classified as successes or failures.
* The average number of successes (μ) that occurs in a specified region is known.
* The probability that a success will occur is proportional to the size of the region.
* The probability that a success will occur in an extremely small region is virtually zero.

Note that the specified region could take many forms. For instance, it could be a length, an area, a volume, a period of time, etc.

## Notation

The following notation is helpful, when we talk about the Poisson distribution.

* *e*: A constant equal to approximately 2.71828. (Actually, *e* is the base of the natural logarithm system.)
* μ: The mean number of successes that occur in a specified region.
* *x*: The actual number of successes that occur in a specified region.
* P(*x*; μ): The **Poisson probability** that exactly *x* successes occur in a Poisson experiment, when the mean number of successes is μ.

## **Poisson Distribution**

A **Poisson random variable** is the number of successes that result from a Poisson experiment. The [probability distribution](http://stattrek.com/Help/Glossary.aspx?Target=Probability_distribution) of a Poisson random variable is called a **Poisson distribution**.

Given the mean number of successes (μ) that occur in a specified region, we can compute the Poisson probability based on the following formula:

**Poisson Formula.** Suppose we conduct a Poisson experiment, in which the average number of successes within a given region is μ. Then, the Poisson probability is:

P(*x*; μ) = (e-μ) (μx) / x!

where *x* is the actual number of successes that result from the experiment, and *e* is approximately equal to 2.71828.

The Poisson distribution has the following properties:

* The mean of the distribution is equal to μ .
* The [variance](http://stattrek.com/Help/Glossary.aspx?Target=Variance) is also equal to μ .

**Example 1**  
The average number of homes sold by the Acme Realty company is 2 homes per day. What is the probability that exactly 3 homes will be sold tomorrow?

*Solution:* This is a Poisson experiment in which we know the following:

* μ = 2; since 2 homes are sold per day, on average.
* x = 3; since we want to find the likelihood that 3 homes will be sold tomorrow.
* e = 2.71828; since *e* is a constant equal to approximately 2.71828.

We plug these values into the Poisson formula as follows:

P(*x*; μ) = (e-μ) (μx) / x!   
P(3; 2) = (2.71828-2) (23) / 3!   
P(3; 2) = (0.13534) (8) / 6   
P(3; 2) = 0.180

Thus, the probability of selling 3 homes tomorrow is 0.180 .

# Normal Distribution

The **normal distribution** refers to a family of [continuous probability distributions](http://stattrek.com/Help/Glossary.aspx?Target=Continuous%20probability%20distribution) described by the normal equation.

**Normal equation.** The value of the random variable *Y* is:

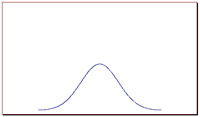
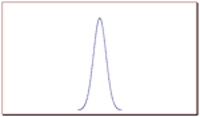
Y = { 1/[ σ \* sqrt(2π) ] } \* e-(x - μ)2/2σ2

where *X* is a normal random variable, μ is the mean, σ is the standard deviation, π is approximately 3.14159, and *e* is approximately 2.71828.

The random variable *X* in the normal equation is called the **normal random variable**. The normal equation is the [probability density function](http://stattrek.com/Help/Glossary.aspx?Target=Probability%20density%20function) for the normal distribution.

## **The Normal Curve**

The graph of the normal distribution depends on two factors - the mean and the standard deviation. The mean of the distribution determines the location of the center of the graph, and the standard deviation determines the height and width of the graph. When the standard deviation is large, the curve is short and wide; when the standard deviation is small, the curve is tall and narrow. All normal distributions look like a symmetric, bell-shaped curve, as shown below.

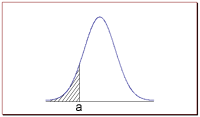
            

The curve on the left is shorter and wider than the curve on the right, because the curve on the left has a bigger standard deviation.

## Probability and the Normal Curve

The normal distribution is a continuous probability distribution. This has several implications for probability.

* The total area under the normal curve is equal to 1.
* The probability that a normal random variable *X* equals any particular value is 0.
* The probability that *X* is greater than *a* equals the area under the normal curve bounded by *a* and plus infinity (as indicated by the non-shaded area in the figure below).
* The probability that *X* is less than *a* equals the area under the normal curve bounded by *a* and minus infinity (as indicated by the shaded area in the figure below).



Additionally, every normal curve (regardless of its mean or standard deviation) conforms to the following "rule".

* About 68% of the area under the curve falls within 1 standard deviation of the mean.
* About 95% of the area under the curve falls within 2 standard deviations of the mean.
* About 99.7% of the area under the curve falls within 3 standard deviations of the mean.

Collectively, these points are known as the **empirical rule** or the **68-95-99.7 rule**. Clearly, given a normal distribution, most outcomes will be within 3 standard deviations of the mean.

To find the probability associated with a normal random variable, use a graphing calculator, an online normal distribution calculator, or a normal distribution table. In the examples below, we illustrate the use of Stat Trek's [Normal Distribution Calculator](http://stattrek.com/Tables/Normal.aspx), a free tool available on this site. In the next lesson, we demonstrate the use of normal distribution tables.

**Example 1**  
An average light bulb manufactured by the Acme Corporation lasts 300 days with a standard deviation of 50 days. Assuming that bulb life is normally distributed, what is the probability that an Acme light bulb will last at most 365 days?

*Solution:* Given a mean score of 300 days and a standard deviation of 50 days, we want to find the cumulative probability that bulb life is less than or equal to 365 days. Thus, we know the following:

* The value of the normal random variable is 365 days.
* The mean is equal to 300 days.
* The standard deviation is equal to 50 days.

We enter these values into the Normal Distribution Calculator and compute the cumulative probability. The answer is: P( X < 365) = 0.90. Hence, there is a 90% chance that a light bulb will burn out within 365 days.

**Example 2**  
Suppose scores on an IQ test are normally distributed. If the test has a mean of 100 and a standard deviation of 10, what is the probability that a person who takes the test will score between 90 and 110?

*Solution:* Here, we want to know the probability that the test score falls between 90 and 110. The "trick" to solving this problem is to realize the following:

P( 90 < *X* < 110 ) = P( X < 110 ) - P( X < 90 )

We use the Normal Distribution Calculator to compute both probabilities on the right side of the above equation.

* To compute P( X < 110 ), we enter the following inputs into the calculator: The value of the normal random variable is 110, the mean is 100, and the standard deviation is 10. We find that P( X < 110 ) is 0.84.
* To compute P( X < 90 ), we enter the following inputs into the calculator: The value of the normal random variable is 90, the mean is 100, and the standard deviation is 10. We find that P( X < 90 ) is 0.16.

We use these findings to compute our final answer as follows:

P( 90 < *X* < 110 ) = P( X < 110 ) - P( X < 90 )  
P( 90 < *X* < 110 ) = 0.84 - 0.16  
P( 90 < *X* < 110 ) = 0.68

Thus, about 68% of the test scores will fall between 90 and 110.

**Reference:** The above is from Statitstics Notes at <http://stattrek.com/>