Following handout is the material taken from following website which is a good source of learning statistics online. <http://stattrek.com/sampling/sampling-distribution.aspx?Tutorial=AP>. The purpose is to encourage you people to discover good website which provide a great platform of self learning

**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.

Probability Distribution Prerequisites

To understand probability distributions, it is important to understand variables. random variables, and some notation.

* 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 |

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. We call one of these 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.

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* ) ].