Statistics How To

Statistics for the rest of us!

Skewed Distribution: Definition, Examples

Share on

Find an article

Search

Skewed Distribution / Asymmetric Distribution: Contents:

- What is a Skewed Distribution?
- Skewed Left
- Skewed Right
- Log Transformations and Statistical Tests
- Skew normal distribution

What is a Skewed Distribution?

Feel like "cheating" at Statistics? Check out the grade-increasing book that's recommended reading at top universities!

Watch the video or read the article below:

Skewed Distribution: left skewed vs right skewed

If one tail is longer than another, the distribution is skewed. These distributions are sometimes called asymmetric or asymmetrical distributions as they don't show any kind of symmetry. Symmetry means that one half of the distribution is a mirror image of the other half. For example, the normal distribution is a symmetric distribution with no skew. The tails are exactly the same.

A normal curve.

A left-skewed distribution has a long left tail. Left-skewed distributions are also called *negatively-skewed* distributions. That's because there is a long tail in the negative direction on the number line. The mean is also to the left of the peak.

A right-skewed distribution has a long right tail. Right-skewed distributions are also called positive-skew distributions. That's because there is a long tail in the positive direction on the number line. The mean is also to the right of the peak.
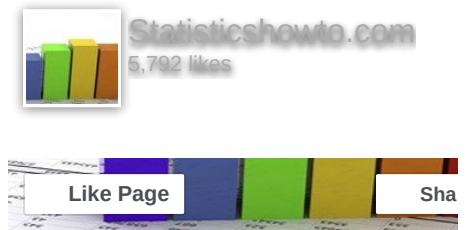
NEED HELP NOW with a homework problem? CLICK HERE!

Probability and Statistics Topic Indexes

Basic Statistics.
Bayesian Statistics and Probability
Descriptive Statistics: Charts, Graphs and Plots.
Probability.
Binomial Theorem.
Definitions for Common Statistics Terms.
Critical Values.
Hypothesis Testing.
Normal Distributions.
T-Distributions.
Central Limit Theorem.
Confidence Intervals.
Chebyshev's Theorem.
Sampling and Finding Sample Sizes.
Chi Square.
Online Tables (z-table, chi-square, t-dist etc.).
Regression Analysis / Linear Regression.
Non Normal Distributions.

The normal distribution is the most common distribution you'll come across. Next, you'll see a fair amount of negatively skewed distributions. For example, household income in the U.S. is negatively skewed with a very long left tail.

Income in the U.S. Image: NY Times.

Interestingly, you can take the **same data** and make it a right-skewed distribution. This positively-skewed graph plots number of household's income brackets:

Mean and Median in Skewed Distributions

In a normal distribution, the mean and the median are the same number while the mean and median in a skewed distribution become *different* numbers:

A left-skewed, negative distribution will have the mean to the **left** of the median.

A right-skewed distribution will have the mean to the **right** of the median.

Right skewed distribution: Mean is to the right
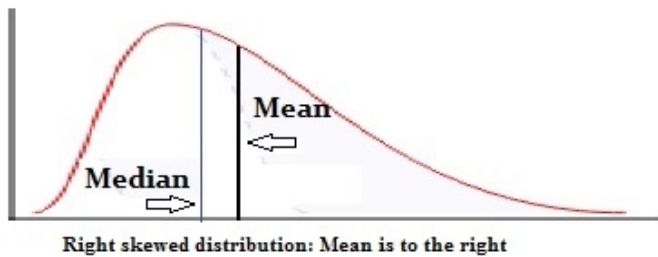
Effects on Statistics

The normal distribution is the easiest distribution to work with in order to gain an understanding about statistics. Real life distributions are usually skewed. Too much skewness, and many statistical techniques don't work. As a result, advanced mathematical techniques including logarithms and quantile regression techniques are used. Read more about quantile regression here.

Skewed Left (Negative Skew)

A left skewed distribution is sometimes called a negatively skewed distribution because it's long tail is on the negative direction on a number line.

A common misconception is that the *peak* of distribution is what defines "peakness." In other words, a peak that tends to the left is left skewed distribution. This is incorrect. There are two main things that make a distribution skewed left:

1. The mean is to the left of the peak. This is the main definition behind "skewness", which is technically a measure of the distribution of values around the mean.
2. The tail is longer on the left.
3. In most cases, the mean is to the left of the median. This isn't a reliable test for skewness though, as some distributions (i.e. many multimodal distributions) violate this rule. You should think of this as a "general idea" kind of rule, and not a set-in-stone one.

Left-Skewed (Negative Skewness)
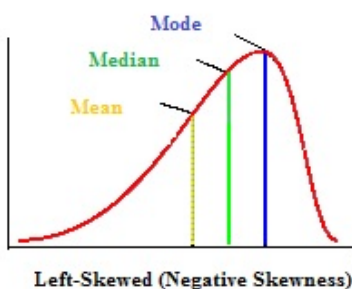
In a left skewed distribution, the mean is to the left of the peak.

Left Skewed and Numerical Values

Skewness can be shown with a list of numbers as well as on a graph. For example, take the numbers 1,2, and 3. They are evenly spaced, with 2 as the mean (1 + 2 + 3 / 3 = 6 / 3 = 2). If you add a number to the far left (think in terms of adding a value to the number line), the distribution becomes left skewed:
-10, 1, 2, 3.
Similarly, if you add a value to the far right, the set of numbers becomes right skewed:
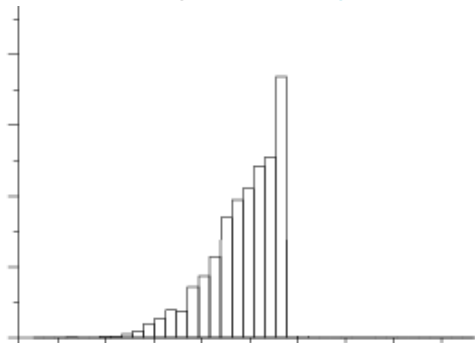1, 2, 3, 10.

Left Skewed Boxplot

If the bulk of observations are on the high end of the scale, a boxplot is left skewed. Consequently, the left whisker is longer than the right whisker.

A left skewed boxplot, showing a long left whisker. Image: SHU.EDU

Left Skewed Histogram

*Left skewed histograms* are Histograms with long tails on the left.

Skewed Right / Positive Skew

A right skewed distribution is sometimes called a positive skew distribution. That's because the tail is longer on the positive direction of the number line.

Right Skewed Histogram

A histogram is right skewed if the peak of the histogram veers to the left. Therefore, the histogram's tail has a positive skew to the right.
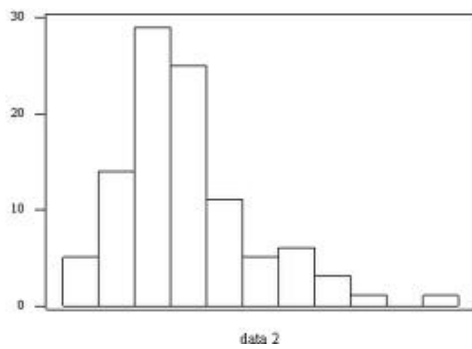
A skewed to the right histogram. Image: SUNY Oswego

Right Skewed Box Plot

If a box plot is skewed to the right, the box shifts to the left and the right whisker gets longer. As a result, the mean is greater than the median
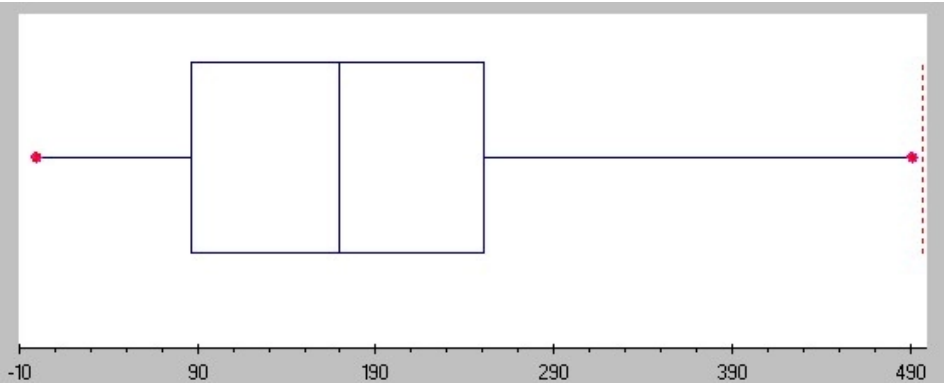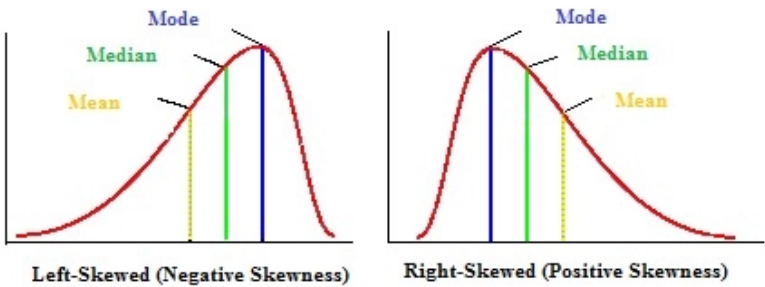

Image: Seton Hall University

Right Skewed Mean and Median

The rule of thumb is that in a right skewed distribution, the mean is usually to the right of the median.

However, like most rules of thumb, there *are* exceptions. Most right skewed distributions you come across in elementary statistics will have the mean to the right of the median. The Journal of Statistics Education points out an exception to the rule:

In a data analysis course, a third moment formula calculates the skew (see: What is a Moment?). Consequently, some distributions can break the rule of thumb. The following distribution was made from a 2002 General Social Survey. Respondents stated how many

people older than 18 lived in their household. This is a right-skewed graph, but the mean is clearly to the left of the median.
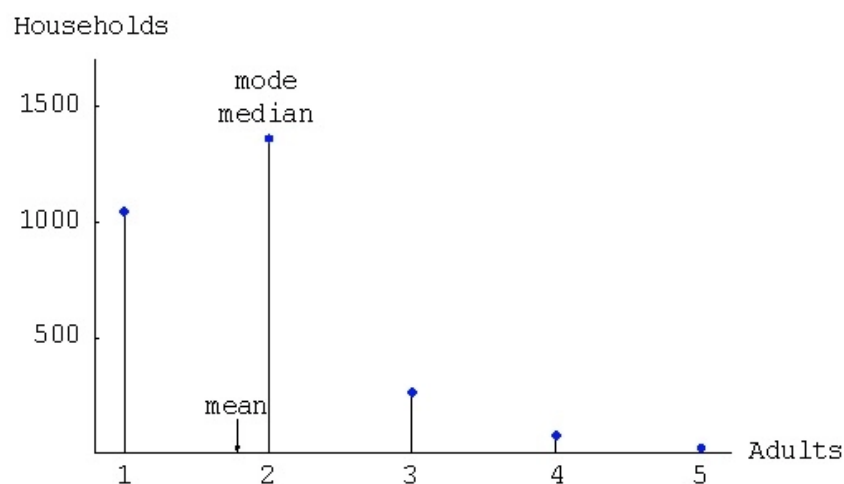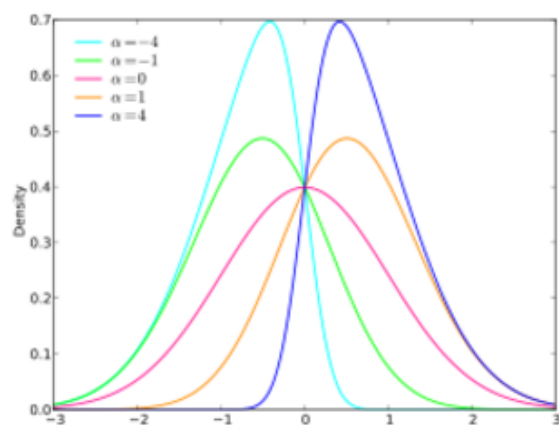

Image: Journal of Statistics Education

There are other exceptions which most involve theoretical mathematics and calculus. The important point to note is that although the mean is generally to the right of the median in a right skewed distribution, it isn't an absolute fact.

Skew Normal Distribution

The probability density function for the skew normal, showing various alphas. Image: skbkekas/Wikimedia Commons.

The **skew normal distribution** is a normal distribution with an extra shape parameter, α. The shape parameter skews the normal distribution to the left or right. As it is only the skew of the normal distribution that's being changed, the skew normal family has many of the same properties of the normal distribution:

- It's defined over the real number line.

- The square of a random variable is a chi-square variable (from a chi-square distribution) with one degree of freedom.
- The distribution is unimodal (one peak).
- The location parameter, μ(i.e. the mean), defines where the peak is and the scale parameter, σ(i.e. the standard deviation) determines the distribution's spread.

The skew normal has a number of interesting properties related to alpha:

- If the skew normal has a skew of zero, then it becomes the normal distribution.
- If the sign of alpha changes, the distribution will flip over the y-axis.
- As alpha increases (in absolute value), the skew also increases.
- As alpha tends towards infinity, the series converges to the folded normal density function.

Therefore, the normal distribution can be seen as a special case of the skew normal distribution.

This is a relatively new distribution, introduced by O'Hagan and Leonard in 1976 in a paper on Bayes' estimation. The work was a basic overview and it wasn't until the 1980s that an in-depth analysis of the distribution was published. It is mainly used in threshold autoregressive stochastic processes and in time series analysis, but can also be used to model various phenomena in a wide range of fields from the sciences to the stock market.

References

Abramowitz, M. and Stegun, I. A. (Eds.). Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 9th printing. New York: Dover, p. 928, 1972.
Kenney, J. F. and Keeping, E. S. "Skewness." §7.10 in Mathematics of Statistics, Pt. 1, 3rd ed. Princeton, NJ: Van Nostrand, pp. 100-101, 1962.
O'Hagan, A. and Leonard, T. (1976). Bayes estimation subject to uncertainty about parameter constraints. Biometrika, 63, 201-202.

Next: Finding Skewness.

CITE THIS AS:
Stephanie Glen. "Skewed Distribution: Definition, Examples" From
StatisticsHowTo.com: Elementary Statistics for the rest of us!
https://www.statisticshowto.com/probability-and-statistics/skewed-distribution/

Need help with a homework or test question? With Chegg Study, you can get step-by-step solutions to your questions from an expert in the field. Your first 30 minutes with a Chegg tutor is free!

Comments? Need to post a correction? Please post a comment on our *Facebook page*.