Basics of Probability

Probability

The study of the occurrence of random events or phenomena.

It does not deal with guarantees, but with the likelihood of an occurrence of an event.

Experiment:

- Any observation or measurement of a random phenomenon.

Outcomes:

- The possible results of an experiment.

Sample Space:

- The set of all possible outcomes of an experiment.

Event:

- A particular collection of possible outcomes from a sample space.

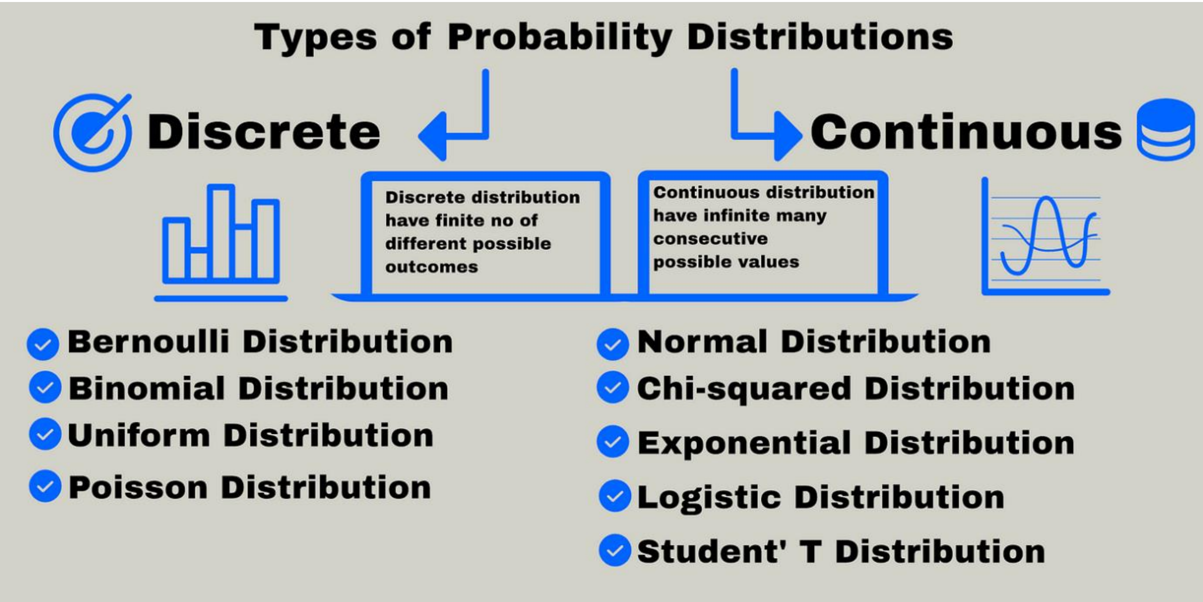
Basic Probability Formulas

Let A and B are two events. The probability formulas are listed below:

All Probability Formulas List in Maths	
Probability Range	$0 \leq P(A) \leq 1$
Rule of Addition	$P(A \cup B) = P(A) + P(B) - P(A \cap B)$
Rule of Complementary Events	$P(A') + P(A) = 1$
Disjoint Events	$P(A \cap B) = 0$
Independent Events	$P(A \cap B) = P(A) \cdot P(B)$
Conditional Probability	$P(A \mid B) = P(A \cap B) / P(B)$
Bayes Formula	$P(A \mid B) = P(B \mid A) \cdot P(A) / P(B)$

What Are Probability Distributions?

- A probability distribution is a statistical function that describes all the possible values and probabilities for a random variable within a given range.
- This range will be bound by the minimum and maximum possible values, but where the possible value would be plotted on the probability distribution will be determined by a number of factors.
- The mean (average), standard deviation, skewness, and kurtosis of the distribution are among these factors.



Types of Probability Distribution

The probability distribution is divided into two parts:

1. **Discrete Probability Distributions**
2. **Continuous Probability Distributions**

Discrete Probability Distribution

A discrete distribution describes the probability of occurrence of each value of a discrete random variable. The number of spoiled apples out of 6 in your refrigerator can be an example of a discrete probability distribution.

Each possible value of the discrete random variable can be associated with a non-zero probability in a discrete probability distribution.

Let's discuss some significant probability distribution functions.

Binomial Distribution

The binomial distribution is a discrete distribution with a finite number of possibilities. When observing a series of what are known as Bernoulli trials, the binomial distribution emerges. A Bernoulli trial is a scientific experiment with only two outcomes: success or failure.

Consider a random experiment in which you toss a biased coin six times with a 0.4 chance of getting head. If 'getting a head' is considered a ‘success’, the binomial distribution will show the probability of r successes for each value of r.

The binomial random variable represents the number of successes (r) in n consecutive independent Bernoulli trials.

Binomial Distribution Formula

$$P(x) = \binom{n}{x} p^x q^{n-x} = \frac{n!}{(n-x)!\,x!} p^x q^{n-x}$$

where

n = the number of trials (or the number being sampled)

x = the number of successes desired

p = probability of getting a success in one trial

$q = 1 - p$ = the probability of getting a failure in one trial

Bernoulli's Distribution

The Bernoulli distribution is a variant of the Binomial distribution in which only one experiment is conducted, resulting in a single observation. As a result, the Bernoulli distribution describes events that have exactly two outcomes.

Definition

A random variable X is said to follow a Bernoulli distribution if its probability mass function is given by

$$P(X = x) = \begin{cases} p^x q^{1-x} & x = 0,1 \\ 0 & otherwise \end{cases}$$

Poisson Distribution

A Poisson distribution is a probability distribution used in statistics to show how many times an event is likely to happen over a given period of time. To put it another way, it's a count distribution. Poisson distributions are frequently used to comprehend independent events at a constant rate over a given time interval.

Poisson Distribution Formula

$$P(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

Where:

$x =$ number of successes

$\lambda =$ rate of success

$e =$ Euler's number (≈ 2.71828))

Continuous Probability Distributions

A continuous distribution describes the probabilities of a continuous random variable's possible values. A continuous random variable has an infinite and uncountable set of possible values (known as the range). The mapping of time can be considered as an example of the continuous probability distribution. It can be from 1 second to 1 billion seconds, and so on.

The area under the curve of a continuous random variable's PDF is used to calculate its probability. As a result, only value ranges can have a non-zero probability. A continuous random variable's probability of equaling some value is always zero.

Now, look at some varieties of the continuous probability distribution.

Normal Distribution

Normal Distribution is one of the most basic continuous distribution types. Gaussian distribution is another name for it. Around its mean value, this probability distribution is symmetrical. It also demonstrates that data close to the mean occurs more frequently than data far from it. Here, the mean is 0, and the variance is a finite value.

Normal Distribution Formula

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

$\mu =$ mean of x

$\sigma =$ standard deviation of x

$\pi \approx$ 3.14159 ...

$e \approx$ 2.71828 ...

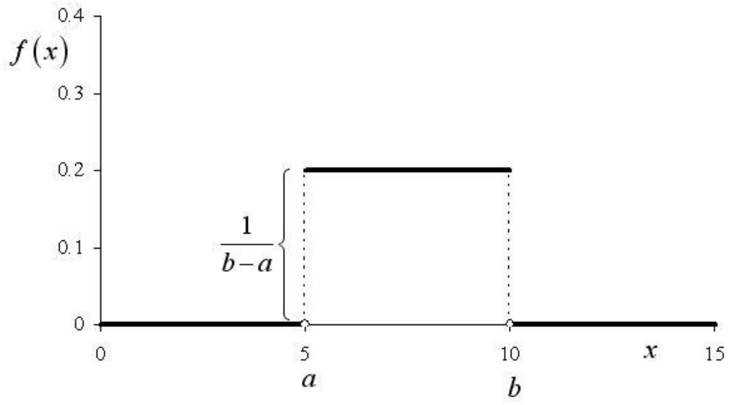
Continuous Uniform Distribution

In continuous uniform distribution, all outcomes are equally possible. Each variable has the same chance of being hit as a result. Random variables are spaced evenly in this symmetric probabilistic distribution, with a 1/ (b-a) probability.

<u>Continuous Distributions</u>

<u>The Uniform distribution from a to b</u>

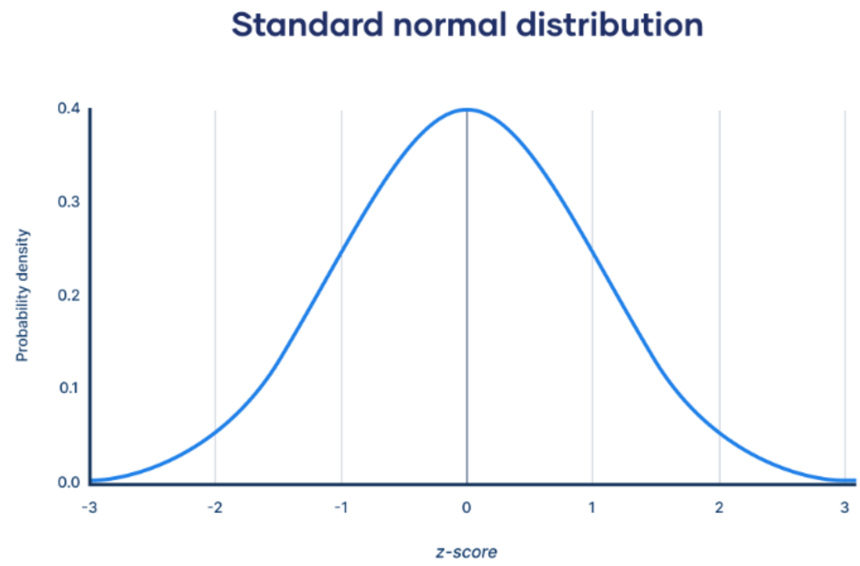
$$f(x) = \begin{cases} \dfrac{1}{b-a} & a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$



Standard normal distribution

The standard normal distribution, also called the z-distribution, is a special normal distribution where the mean is 0 and the standard deviation is 1.

Any normal distribution can be standardized by converting its values into z scores. Z scores tell you how many standard deviations from the mean each value lies.



Converting a normal distribution into a *z*-distribution allows you to calculate the probability of certain values occurring and to compare different data sets.

Note: All normal distributions, like the standard normal distribution, are unimodal and symmetrically distributed with a bell-shaped curve. However, a normal distribution can take on any value as its mean and standard deviation. In the standard normal distribution, the mean and standard deviation are always fixed.

Z-score formula	Explanation
$z = \frac{x - \mu}{\sigma}$	• x = individual value • μ = mean • σ = standard deviation

Chi-squared distribution

Chi-square (X2) distributions are a family of continuous probability distributions. They're widely used in hypothesis tests, including the chi-square goodness of fit test and the chi-square test of independence.

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

Where:

χ^2 = Chi Square obtained
$\sum$ = the sum of
O = observed score
E = expected score

The Chi-Squared Distribution

Definition - A random variable X is said to have the Chi-Squared distribution with parameter v, called degrees of freedom, if the probability density function of X is:

$$f(x) = \begin{cases} \dfrac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}}, & \text{for } x > 0 \\ 0, & \text{elsewhere} \end{cases}$$

where v is a positive integer.

Where T- Gamma Function

Gamma Distribution:

The Gamma distribution is a particular case of the normal distribution, which describes many life events including predicted rainfall, the reliability of mechanical tools and machines, or any applications that only have positive results.

The Gamma distribution

Let the continuous random variable X have density function:

$$f(x) = \begin{cases} \dfrac{\lambda^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Then X is said to have a **Gamma distribution** with parameters α and λ.

where α is the shape parameter, , β is the scale parameter, and Γ is the gamma function.

Exponential distribution

The exponential distribution is a continuous distribution that is commonly used to measure the expected time for an event to occur.

Probability density function [edit]

The probability density function (pdf) of an exponential distribution is

$$f(x;\lambda) = \begin{cases} \lambda e^{-(\lambda x)} & x \geq 0, \\ 0 & x < 0. \end{cases}$$

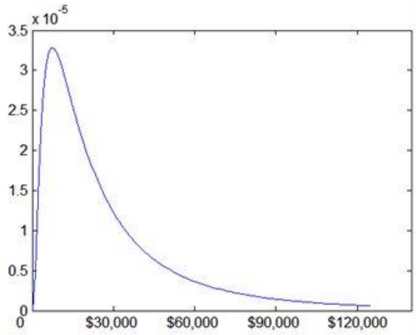
Here $\lambda > 0$ is the parameter of the distribution, often called the *rate parameter*. The distribution is supported on the interval $[0, \infty)$. If a random variable X has this distribution, we write $X \sim \text{Exp}(\lambda)$.

The exponential distribution exhibits infinite divisibility.

Lognormal distribution

Lognormal distribution

- If ln(X) has normal distribution X has lognormal distribution. That is, if X is normally distributed exp(X) is lognormally distributed.
- Notation: $\ln N(\mu, \sigma^2)$
- PDF $f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right]$
- Mean and variance



$$\mu_X = \exp\left(\mu + \sigma^2 / 2\right), \qquad \sigma_X^2 = Var(X) = \left(e^{\sigma^2} - 1\right) e^{2\mu + \sigma^2}$$

Weibull distribution

One of the most widely used distributions in reliability engineering is "Weibull Distribution". It is a kind of versatile distribution that can take the values from the other distributions using the parameter called the shape parameter.

The Weibull Distribution is a continuous probability distribution used to analyse life data, model failure times and access product reliability. It can also fit a huge range of data from many other fields like economics, hydrology, biology, engineering sciences. It is an extreme value of probability distribution which is frequently used to model the reliability, survival, wind speeds and other data.

The only reason to use Weibull distribution is because of its flexibility. Because it can simulate various distributions like normal and exponential distributions. Weibull's distribution reliability is measured with the help of parameters. The two versions of Weibull probability density function(pdf) are

1. Two parameter pdf
2. Three parameter pdf

Weibull Distribution Formulas

The formula general Weibull Distribution for **three-parameter pdf** is given as

$$f(x) = \frac{\gamma}{\alpha}\left(\frac{(x-\mu)}{\alpha}\right)^{\gamma-1} exp^{(-(\frac{(x-\mu)}{\alpha})^{\gamma})} \quad x \geq \mu; \gamma, \alpha > 0$$

Where,

- γ is the **shape parameter**, also called as the Weibull slope or the threshold parameter.
- α is the **scale parameter**, also called the characteristic life parameter.
- μ is the **location parameter**, also called the waiting time parameter or sometimes the shift parameter.

The standard Weibull distribution is derived, when μ=0 and α =1, the formula is reduced and it becomes

$$f(x) = \gamma x^{\gamma-1} exp^{(-x)^{\gamma}}, x \geq 0; y > 0$$

Two-Parameter Weibull Distribution

The formula is practically similar to the three parameters Weibull, except that μ isn't included:

$$f(x) = \frac{\gamma}{\alpha}\left(\frac{(x)}{\alpha}\right)^{\gamma-1} exp^{(-(\frac{(x)}{\alpha})^{\gamma})} \quad x \geq 0$$

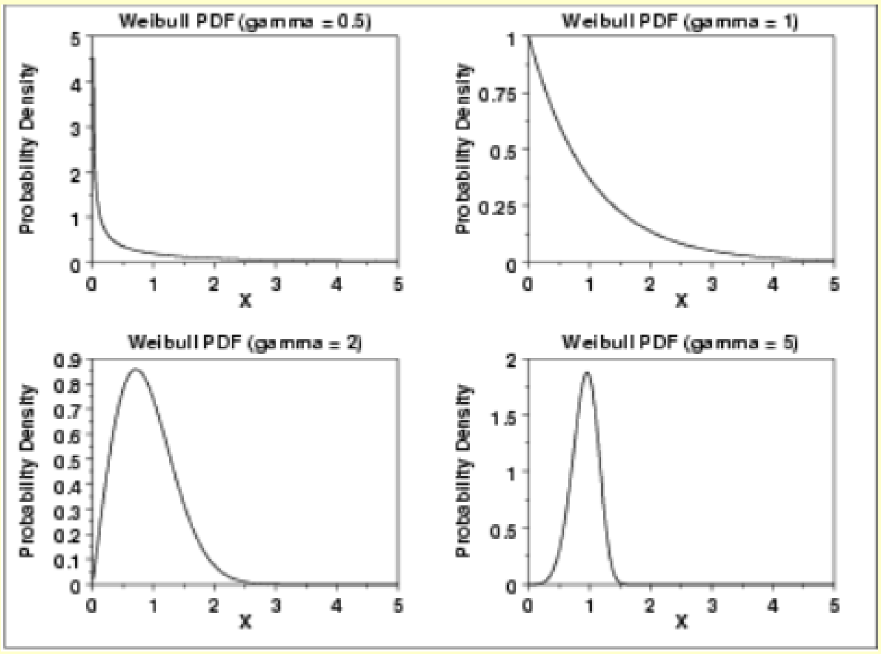
The failure rate is determined by the value of the shape parameter **γ.**

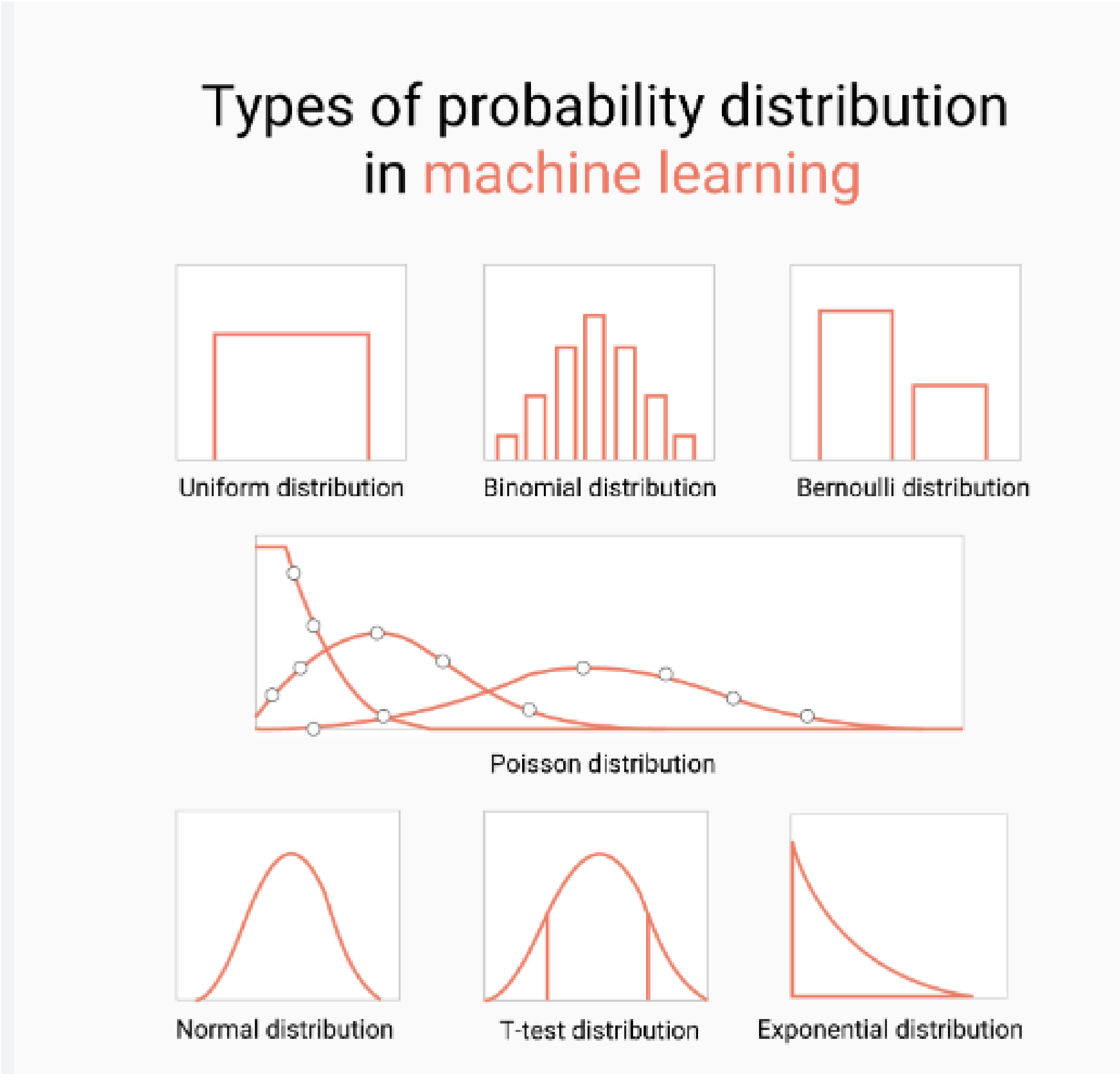
- If γ < 1, then the **failure rate decreases with time**
- If γ = 1, then the **failure rate is constant**

If γ > 1, the **failure rate increases with time**

Plot Weibull Distribution

The following is the plot of the Weibull probability density function.

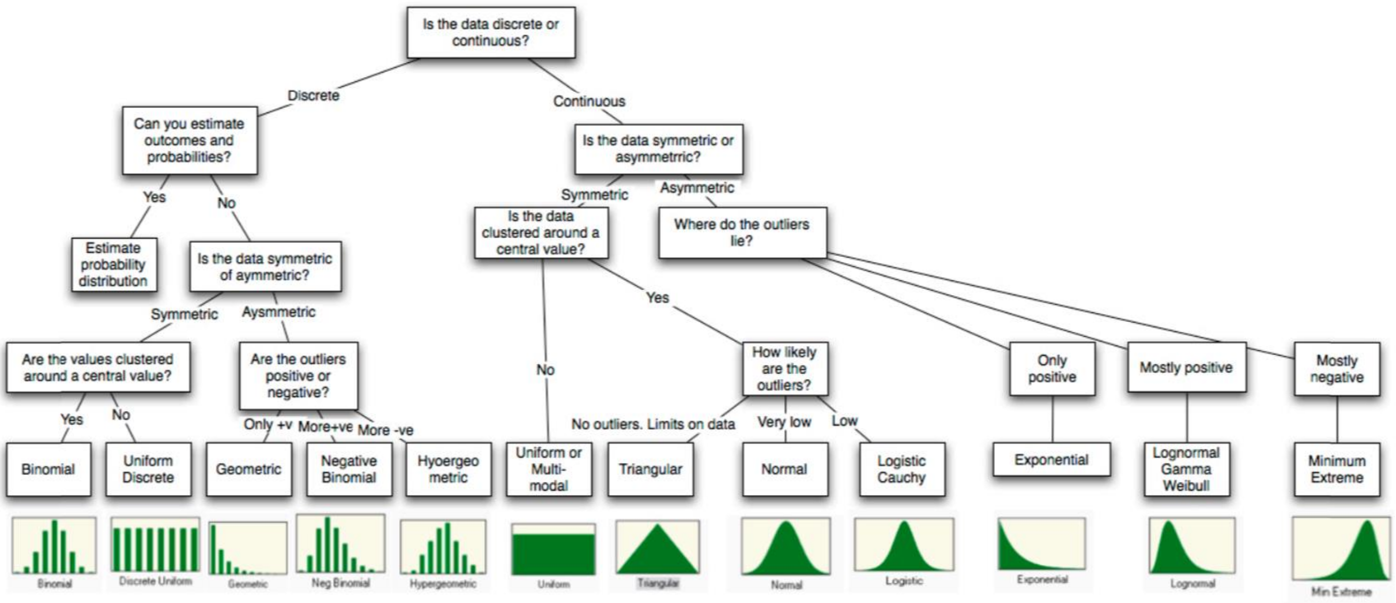



Types of probability distribution
in machine learning
Uniform distribution
Binomial distribution
Bernoulli distribution
Poisson distribution
Normal distribution
T-test distribution
Exponential distribution

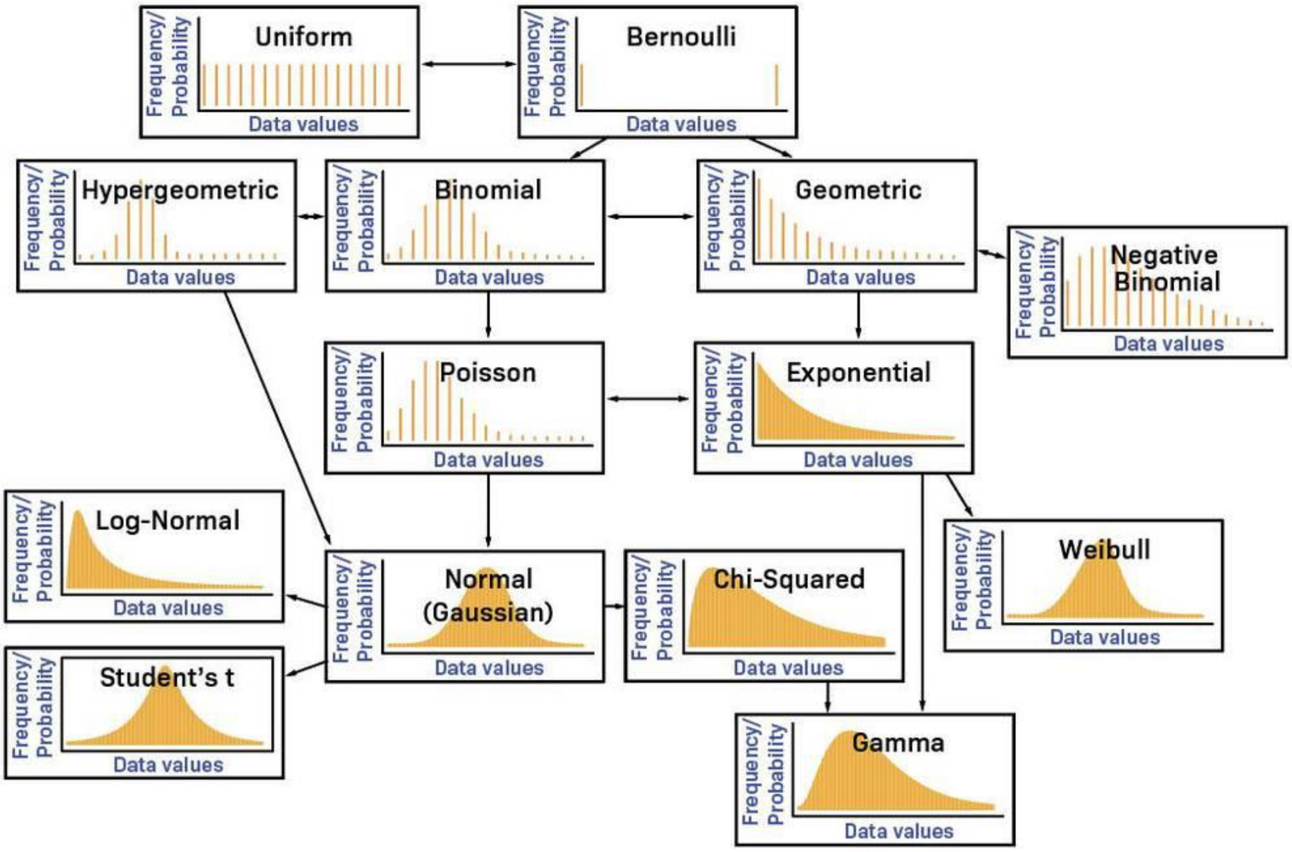
Note:

12 Common Distributions used in Statistics and Data Science

(And when to use them)



Note 2



Note 3:

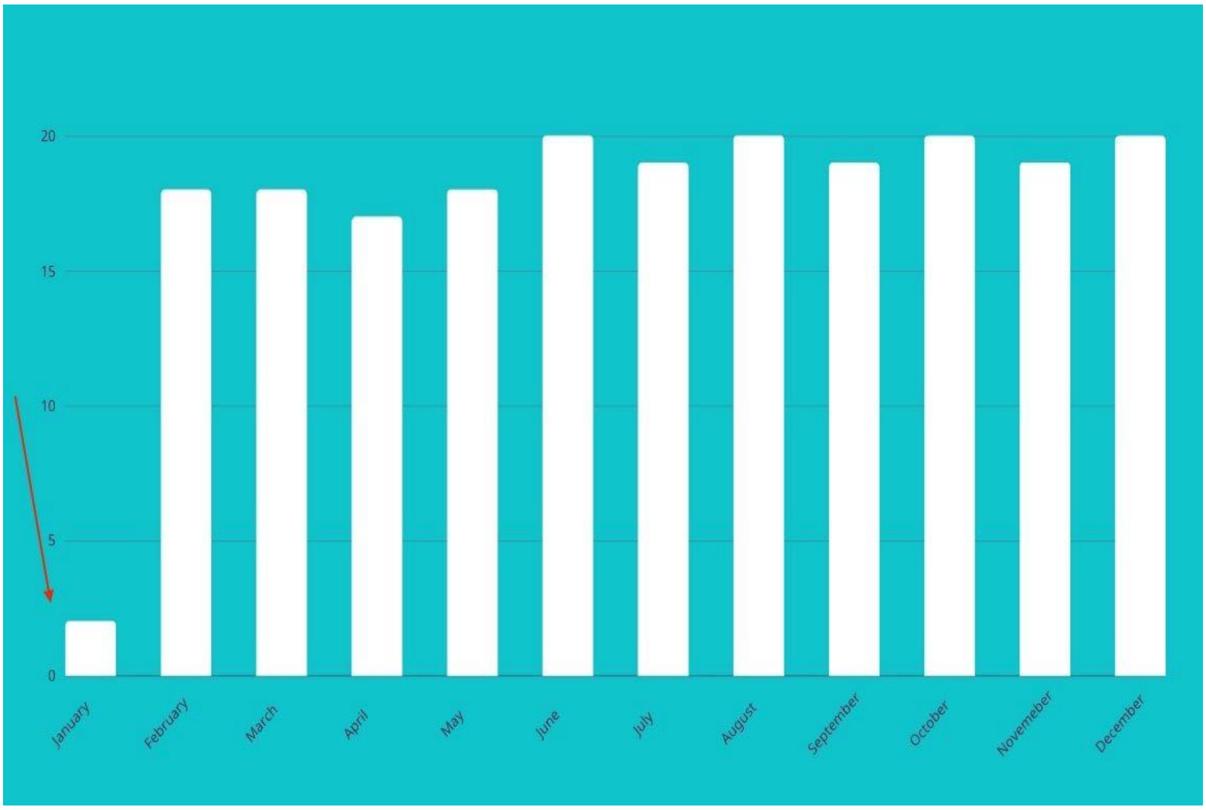
What is an Outlier in Statistics? A Definition

In simple terms, an outlier is an extremely high or extremely low data point relative to the nearest data point and the rest of the neighboring co-existing values in a data graph or dataset you're working with.

Outliers are extreme values that stand out greatly from the overall pattern of values in a dataset or graph.

Below, on the far left of the graph, there is an outlier.

The value in the month of January is significantly less than in the other months.



20
15
10
5
0
January
February
March
April
May
June
July
August
September
October
Novemeber
December