Probability Distribution Summary

Dr. Mohammed Hawa
October 2007

There are many probability distributions that have been studied over the years. Each of which was developed based on observations of a particular physical phenomenon. The table below shows some of these distributions:

	Probability distributions	
	Univariate	**Multivariate**
Discrete:	Benford • Bernoulli • binomial • Boltzmann • categorical • compound Poisson • discrete phase-type • degenerate • Gauss-Kuzmin • geometric • hypergeometric • logarithmic • negative binomial • parabolic fractal • Poisson • Rademacher • Skellam • uniform • Yule-Simon • zeta • Zipf • Zipf-Mandelbrot	Ewens • multinomial • multivariate Polya
Continuous:	Beta • Beta prime • Cauchy • chi-square • Dirac delta function • Coxian • Erlang • exponential • exponential power • F • fading • Fermi-Dirac • Fisher's z • Fisher-Tippett • Gamma • generalized extreme value • generalized hyperbolic • generalized inverse Gaussian • Half-Logistic • Hotelling's T-square • hyperbolic secant • hyper-exponential • hypoexponential • inverse chi-square (scaled inverse chi-square) • inverse Gaussian • inverse gamma (scaled inverse gamma) • Kumaraswamy • Landau • Laplace • Lévy • Lévy skew alpha-stable • logistic • log-normal • Maxwell-Boltzmann • Maxwell speed • Nakagami • normal (Gaussian) • normal-gamma • normal inverse Gaussian • Pareto • Pearson • phase-type • polar • raised cosine • Rayleigh • relativistic Breit-Wigner • Rice • shifted Gompertz • Student's t • triangular • truncated normal • type-1 Gumbel • type-2 Gumbel • uniform • Variance-Gamma • Voigt • von Mises • Weibull • Wigner semicircle • Wilks' lambda	Dirichlet • Generalized Dirichlet distribution . inverse-Wishart • Kent • matrix normal • multivariate normal • multivariate Student • von Mises-Fisher • Wigner quasi • Wishart

In this article we will summarize the characteristics of some popular probability distributions, which you are likely to encounter later in your study of communication engineering. We will begin with *discrete* distributions then move to *continuous* distributions.

A - Discrete Random Variables:

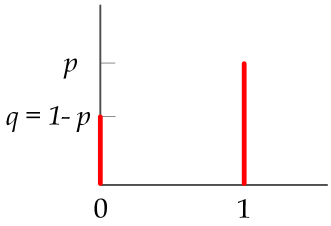
• Bernoulli distribution

The **Bernoulli random variable** takes value 1 with success probability p and value 0 with failure probability $q = 1 - p$.

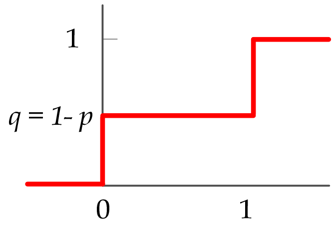
Applications:

- Coin toss.
- Success/Failure experiments.

Probability mass function



Cumulative distribution function



Parameters	$p > 0$ success probability (p is real)
Support	$k = \{0, 1\}$
Probability mass function **(pmf)**	$\begin{matrix} q & \text{for } k = 0 \\ p & \text{for } k = 1 \end{matrix}$
Cumulative distribution function **(CDF)**	$\begin{matrix} 0 & \text{for } k < 0 \\ q & \text{for } 0 \leq k < 1 \\ 1 & \text{for } k \geq 1 \end{matrix}$
Mean	p
Median	N/A
Mode	$\begin{matrix} 0 & \text{if } q > p \\ 0, 1 & \text{if } q = p \\ 1 & \text{if } q < p \end{matrix}$
Variance	pq
Skewness	$\frac{q - p}{\sqrt{pq}}$
Excess kurtosis	$\frac{6p^2 - 6p + 1}{p(1 - p)}$
Probability-generating function **(PGF)**	$1 - p + pz$
Moment-generating function **(MGF)**	$1 - p + pe^t$
Characteristic function	$1 - p + pe^{it}$

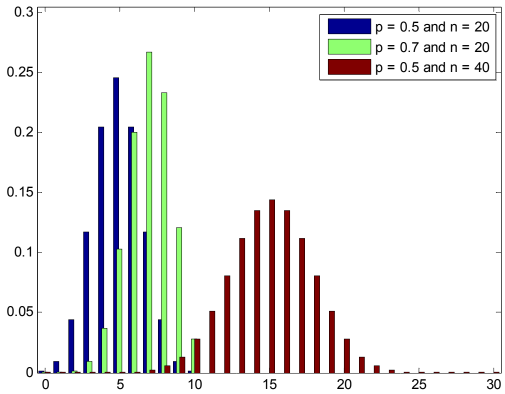
• Binomial distribution

The **binomial random variable** is the sum of n independent, identically distributed Bernoulli random variables, each of which yields success with probability p. In fact, when $n = 1$, the binomial random variable *is* a Bernoulli random variable.

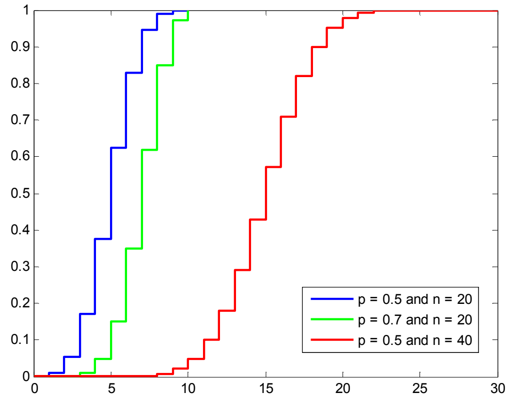
Applications:

- Obtaining k heads in n tossings of a coin.
- Receiving k bits correctly in n transmitted bits.
- Batch arrivals of k packets from n inputs at an ATM switch.

Probability mass function



Cumulative distribution function



Parameters	$n \geq 0$ $0 \leq p \leq 1$	number of trials (integer) success probability (real)
Support	$k \in \{0, \ldots, n\}$	
Probability mass function **(pmf)**	$\binom{n}{k} p^k (1-p)^{n-k}$	
Cumulative distribution function **(CDF)**	$\sum_{j=0}^{\lfloor x \rfloor} \binom{n}{j} p^j (1-p)^{n-j}$	
Mean	np	
Median	one of $\{\lfloor np \rfloor - 1, \lfloor np \rfloor, \lfloor np \rfloor + 1\}$	
Mode	$\lfloor (n+1)p \rfloor$	
Variance	$np(1-p)$	
Skewness	$\frac{1-2p}{\sqrt{np(1-p)}}$	
Excess kurtosis	$\frac{1-6p(1-p)}{np(1-p)}$	
Probability-generating function **(PGF)**	$(1-p+pz)^n$	
Moment-generating function **(MGF)**	$(1-p+pe^t)^n$	
Characteristic function	$(1-p+pe^{it})^n$	

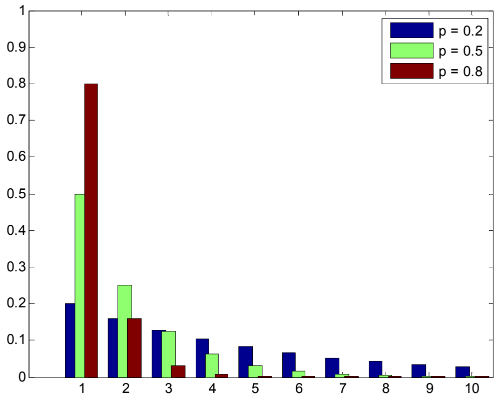
• Geometric distribution

This is the distribution of time slots between the successes of consecutive *independent* Bernoulli trails. Like its continuous analogue (the exponential distribution), the geometric distribution is *memoryless*, which is an important property in certain applications.

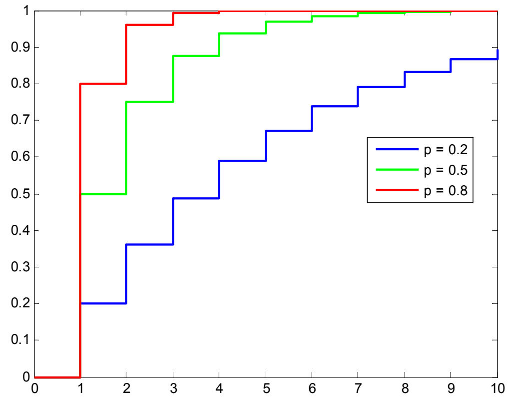
Applications:

- To represent the number of dice rolls needed until you roll a six.
- Queueing theory and discrete Markov chains.

Probability mass function



Cumulative distribution function



Parameters	$0 < p \leq 1$ success probability (real)
Support	$k \in \{1, 2, 3, \dots\}$
Probability mass function **(pmf)**	$(1-p)^{k-1}\,p$
Cumulative distribution function **(CDF)**	$1-(1-p)^{\lfloor k \rfloor}$
Mean	$\frac{1}{p}$
Median	$\left\lceil \frac{-\ln(2)}{\ln(1-p)} \right\rceil$
Mode	1
Variance	$\frac{1-p}{p^2}$
Skewness	$\frac{2-p}{\sqrt{1-p}}$
Excess kurtosis	$6+\frac{p^2}{1-p}$
Probability-generating function **(PGF)**	$\frac{pz}{1-(1-p)z}$
Moment-generating function **(MGF)**	$\frac{p\,e^t}{1-(1-p)e^t}$
Characteristic function	$\frac{pe^{it}}{1-(1-p)e^{it}}$

• Poisson distribution

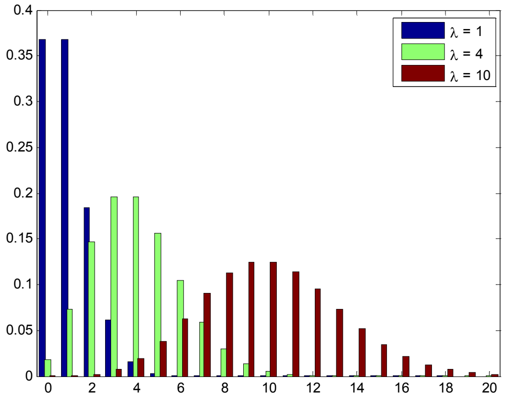
The **Poisson distribution** describes the probability that a number k of events occur in a fixed period of time assuming these events occur at *random* with a rate λ.

Applications:

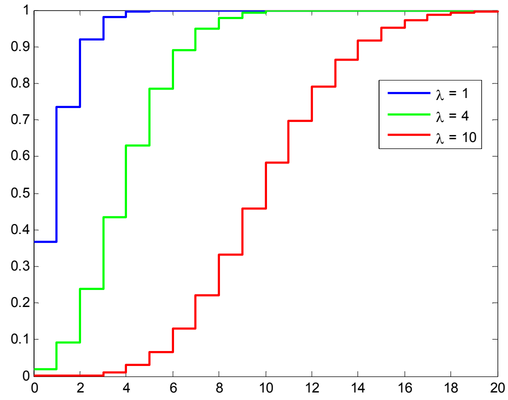
- The number of phone calls at a call center per minute.
- The number of times a web server is accessed per minute.

- The number of spelling mistakes one makes while typing a single page.
- The number of mutations in a given stretch of DNA after radiation.
- The number of unstable nuclei that decayed within a given period of time in a piece of radioactive substance.
- The number of stars in a given volume of space.

Probability mass function



Cumulative distribution function

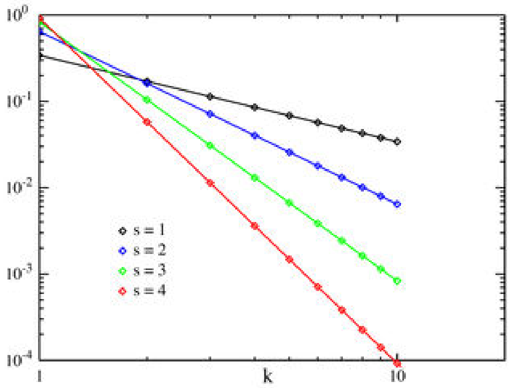


Parameters	$\lambda \in (0, \infty)$	rate (real)
Support	$k \in \{0, 1, 2, \ldots\}$	
Probability mass function **(pmf)**	$\frac{e^{-\lambda}\lambda^k}{k!}$	
Cumulative distribution function **(CDF)**	$\frac{\Gamma(\lfloor k+1 \rfloor, \lambda)}{\lfloor k \rfloor!}$ for $k \geq 0$	where $\Gamma(x, y)$ is the Incomplete gamma function
Mean	λ	
Median	usually about $\lfloor \lambda + 1/3 - 0.02/\lambda \rfloor$	
Mode	$\lfloor \lambda \rfloor$ (and $\lambda - 1$ if λ is an integer)	
Variance	λ	
Skewness	$\lambda^{-1/2}$	
Excess kurtosis	λ^{-1}	
Probability-generating function **(PGF)**	$e^{\lambda(z-1)}$	
Moment-generating function **(MGF)**	$e^{\lambda(e^t-1)}$	
Characteristic function	$e^{\lambda(e^{it}-1)}$	

• Zipf distribution

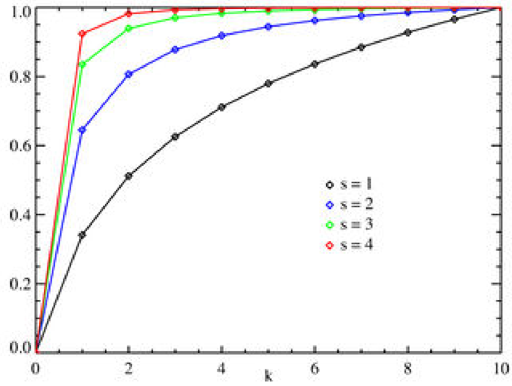
The term Zipf's law refers to *frequency distributions* of **rank data**. Originally, Zipf's law stated that, in natural language utterances, the frequency of any word is roughly inversely proportional to its rank in the frequency table. So, the most frequent word will occur approximately twice as often as the second most frequent word, which occurs twice as often as the fourth most frequent word, etc.

Probability mass function



N=10 (log-log scale). Note that the function is only defined at integer values of k. The connecting lines do not indicate continuity.

Cumulative distribution function



N=10. Note that the function is only defined at integer values of k. The connecting lines do not indicate continuity.

Parameters	$s > 0$ $N \in \{1, 2, 3 \ldots\}$	exponent (real) number of elements (integer)
Support	$k \in \{1, 2, \ldots, N\}$	(rank)
Probability mass function **(pmf)**	$\frac{1/k^s}{H_{N,s}}$	where $H_{N,s}$ is the Nth generalized harmonic number $H_{N,s} = \sum_{k=1}^{N} \frac{1}{k^s}$
Cumulative distribution function **(CDF)**	$\frac{H_{k,s}}{H_{N,s}}$	
Mean	$\frac{H_{N,s-1}}{H_{N,s}}$	
Moment-generating function **(MGF)**	$\frac{1}{H_{N,s}} \sum_{n=1}^{N} \frac{e^{nt}}{n^s}$	
Characteristic function	$\frac{1}{H_{N,s}} \sum_{n=1}^{N} \frac{e^{int}}{n^s}$	

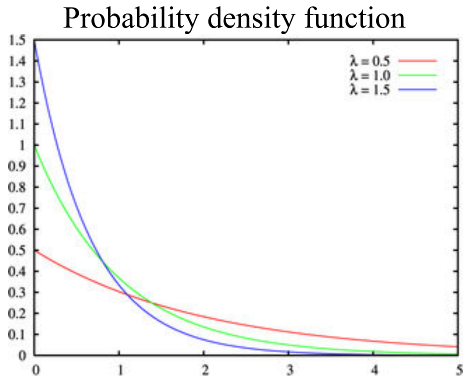
B - Continuous Random Variables:

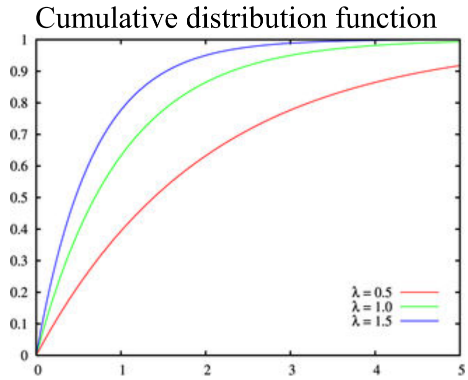
• Exponential distribution

The exponential distribution represents the probability distribution of the time intervals between successive Poisson arrivals. The exponential distribution is the continuous counterpart of the geometric distribution, and is the only continuous distribution that is memoryless.

Applications:

- Similar to that of Poisson distribution (e.g., the time it takes before your next telephone call, the distance between mutations on a DNA strand, etc).





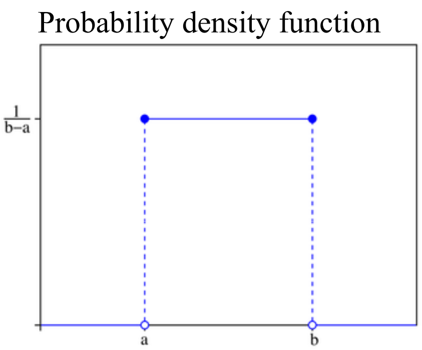
Parameters	$\lambda > 0$ rate or inverse scale (real)
Support	$x \in [0; +\infty)$
Probability density function **(pdf)**	$\lambda e^{-\lambda x}$
Cumulative distribution function **(CDF)**	$1 - e^{-\lambda x}$
Mean	λ^{-1}
Median	$\ln(2)/\lambda$
Mode	0
Variance	λ^{-2}
Skewness	2
Excess kurtosis	6
Moment-generating function **(MGF)**	$\left(1 - \frac{t}{\lambda}\right)^{-1}$
Characteristic function	$\left(1 - \frac{it}{\lambda}\right)^{-1}$

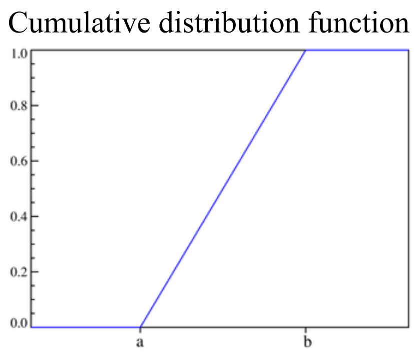
• Uniform distribution

The uniform distribution is often abbreviated $U(a, b)$. In this distribution, all time intervals of the same length are equally probable.

Applications:

- To test a statistic for the simple null hypothesis.
- In simulation software packages, in which uniformly-distributed pseudo-random numbers are generated first and then converted to other types of distributions.





Parameters	$a, b \in (-\infty, \infty)$	lower/upper limits
Support	$a \le x \le b$	
Probability density function **(pdf)**	$\frac{1}{b-a}$ for $a \le x \le b$ 0 for $x < a$ or $x > b$	
Cumulative distribution function **(CDF)**	0 for $x < a$ $\frac{x-a}{b-a}$ for $a \le x < b$ 1 for $x \ge b$	
Mean	$\dfrac{a+b}{2}$	
Median	$\dfrac{a+b}{2}$	
Mode	any value in $[a, b]$	
Variance	$\dfrac{(b-a)^2}{12}$	
Skewness	0	
Excess kurtosis	$-\dfrac{6}{5}$	
Moment-generating function **(MGF)**	$\dfrac{e^{tb}-e^{ta}}{t(b-a)}$	The raw moments are: $m_k = \dfrac{1}{k+1}\sum_{i=0}^{k} a^i b^{k-i}.$
Characteristic function	$\dfrac{e^{itb}-e^{ita}}{it(b-a)}$	

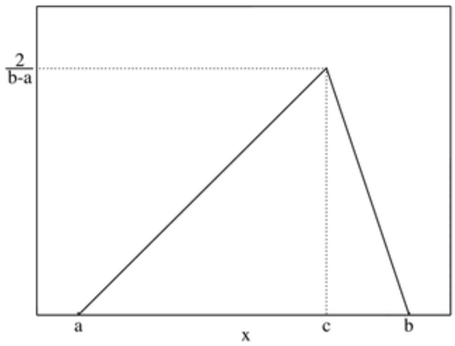
• Triangular distribution

The Triangular distribution is used as a subjective description of a population for which there is only limited sample data.

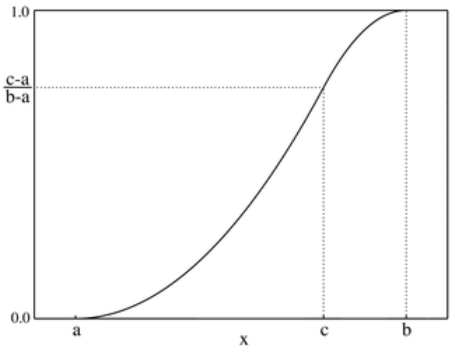
Applications:

- Used in business decision making and project management to describe the time to completion of a task.

Probability density function



Cumulative distribution function



Parameters	$a:\ a \in (-\infty, \infty)$	lower limit
	$b:\ b > a$	upper limit
	$c:\ a \le c \le b$	mode

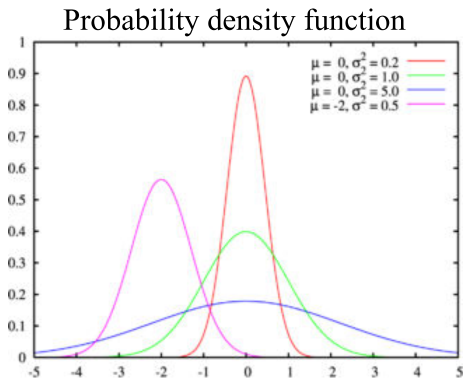
Support	$a \le x \le b$
Probability density function **(pdf)**	$\begin{cases} \frac{2(x-a)}{(b-a)(c-a)} & \text{for } a \le x \le c \\ \frac{2(b-x)}{(b-a)(b-c)} & \text{for } c \le x \le b \end{cases}$
Cumulative distribution function **(CDF)**	$\begin{cases} \frac{(x-a)^2}{(b-a)(c-a)} & \text{for } a \le x \le c \\ 1-\frac{(b-x)^2}{(b-a)(b-c)} & \text{for } c \le x \le b \end{cases}$
Mean	$\frac{a+b+c}{3}$
Median	$\begin{cases} a+\frac{\sqrt{(b-a)(c-a)}}{\sqrt{2}} & \text{for } c \ge \frac{b-a}{2} \\ b-\frac{\sqrt{(b-a)(b-c)}}{\sqrt{2}} & \text{for } c \le \frac{b-a}{2} \end{cases}$
Mode	c
Variance	$\frac{a^2+b^2+c^2-ab-ac-bc}{18}$
Skewness	$\frac{\sqrt{2}(a+b-2c)(2a-b-c)(a-2b+c)}{5(a^2+b^2+c^2-ab-ac-bc)^{\frac{3}{2}}}$
Excess kurtosis	$-\frac{3}{5}$
Moment-generating function **(MGF)**	$2\frac{(b-c)e^{at}-(b-a)e^{ct}+(c-a)e^{bt}}{(b-a)(c-a)(b-c)t^2}$
Characteristic function	$-2\frac{(b-c)e^{iat}-(b-a)e^{ict}+(c-a)e^{ibt}}{(b-a)(c-a)(b-c)t^2}$

• Normal distribution

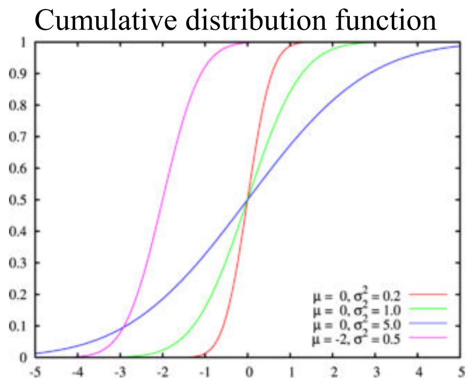
The **normal distribution** is also called the **Gaussian distribution**, and is denoted by $N(\mu, \sigma^2)$. The **standard normal distribution** is the normal distribution with a *mean* of zero and a *variance* of one. It is often called the **bell curve** because the graph of its probability density resembles a bell.

Applications:

- Many physical phenomena (like **noise**) can be approximated well by the normal distribution.
- Many statistical tests are based on the assumption of normal distribution.
- Normal distribution arises as the limiting distribution of several continuous and discrete families of distributions by the **central limit theorem**.
- Measurement errors are often assumed to be normally distributed.
- Financial variables such as stock values or commodity prices can be modeled by the normal distribution.
- Light intensity from a single source is usually assumed to be normally distributed.



The green line is the standard normal distribution



Colors match the image to the left

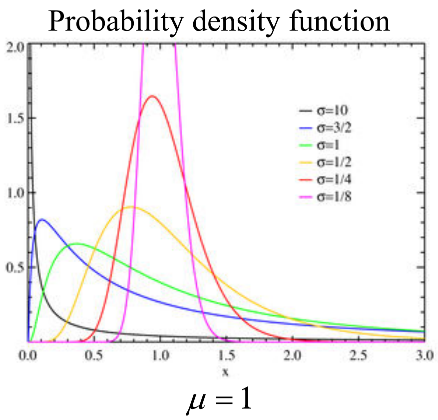
Parameters	μ $\sigma^2 > 0$	location / mean (real) squared scale / variance (real)
Support	$x \in \mathbb{R}$	
Probability density function **(pdf)**	$\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$	
Cumulative distribution function **(CDF)**	$\frac{1}{2}\left(1 + \operatorname{erf}\frac{x-\mu}{\sigma\sqrt{2}}\right)$	
Mean	μ	
Median	μ	
Mode	μ	
Variance	σ^2	
Skewness	0	
Excess kurtosis	0	
Moment-generating function **(MGF)**	$\exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)$	
Characteristic function	$\exp\left(\mu i t - \frac{\sigma^2 t^2}{2}\right)$	

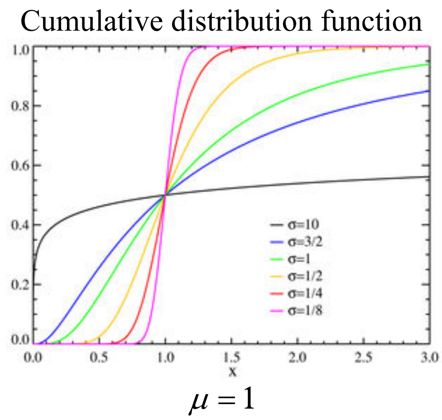
• Log-normal distribution

If Y is a random variable with a normal distribution, then $X = \exp(Y)$ has a log-normal distribution; likewise, if X is log-normally distributed, then $\ln(X)$ is normally distributed.

Applications:

- The long-term return rate on a stock investment can be modeled as log-normal.





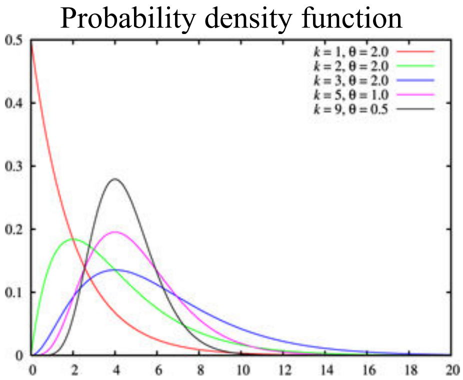
Parameters	$-\infty \le \mu \le \infty$ mean $\sigma \ge 0$ standard deviation
Support	$x \in [0; +\infty)$
Probability density function **(pdf)**	$\frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{[\ln(x)-\mu]^2}{2\sigma^2}\right)$
Cumulative distribution function **(CDF)**	$\frac{1}{2} + \frac{1}{2}\mathrm{erf}\left[\frac{\ln(x)-\mu}{\sigma\sqrt{2}}\right]$
Mean	$e^{\mu+\sigma^2/2}$
Median	e^{μ}
Mode	$e^{\mu-\sigma^2}$
Variance	$(e^{\sigma^2}-1)e^{2\mu+\sigma^2}$
Skewness	$(e^{\sigma^2}+2)\sqrt{e^{\sigma^2}-1}$
Excess kurtosis	$\frac{e^{6\sigma^2}-4e^{3\sigma^2}+6e^{\sigma^2}-3}{e^{4\mu+2\sigma^2}(e^{\sigma^2}-1)^4}$
Moment-generating function **(MGF)**	The MGF does not exist, but moments do, and are given by: $m_k = e^{k\mu+k^2\sigma^2/2}$

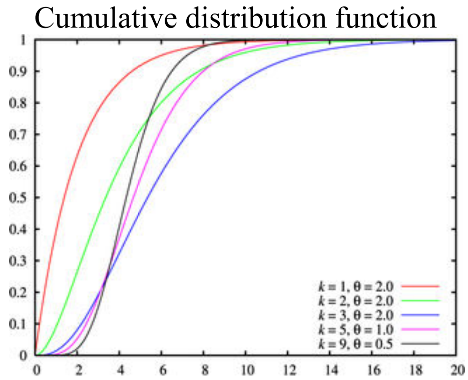
• Gamma (Erlang) distribution

A gamma distributed random variable is denoted by $\Gamma(k,\lambda)$. The Erlang distribution is a special case of the *gamma distribution* where the shape parameter k is an integer. In the Gamma distribution, this parameter is a real number.

Applications:

- The number of telephone calls which might be made at the same time to a switching center.





Parameters	$k > 0$ $\lambda > 0$	shape (real) rate (real)
Support	$x \in [0; \infty)$	
Probability density function **(pdf)**	$\dfrac{\lambda^k x^{k-1} e^{-\lambda x}}{\Gamma(k)}$ where the gamma function $\Gamma(k) = \int_0^\infty y^{k-1} e^{-y} dy$ becomes $\Gamma(k) = (k-1)!$ if k is an integer.	
Cumulative distribution function **(CDF)**	$\dfrac{\int_0^{\lambda x} y^{k-1} e^{-y} dy}{\Gamma(k)}$	
Mean	k/λ	
Median	no simple closed form	
Mode	$(k-1)/\lambda$ for $k \geq 1$	
Variance	k/λ^2	
Skewness	$\dfrac{2}{\sqrt{k}}$	
Excess kurtosis	$\dfrac{6}{k}$	
Moment-generating function **(MGF)**	$(1 - t/\lambda)^{-k}$ for $t < \lambda$	
Characteristic function	$(1 - it/\lambda)^{-k}$	

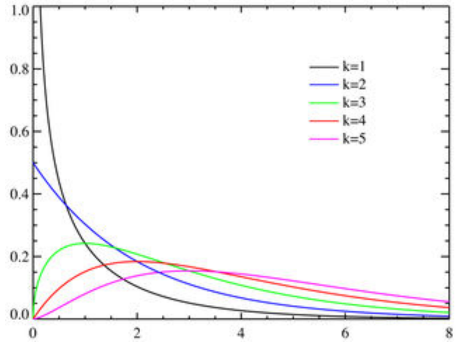
• chi-square distribution

The chi-square distribution ($\chi 2$ distribution) is a special case of the gamma distribution where $k = n/2$ and $\lambda = 1/2$.

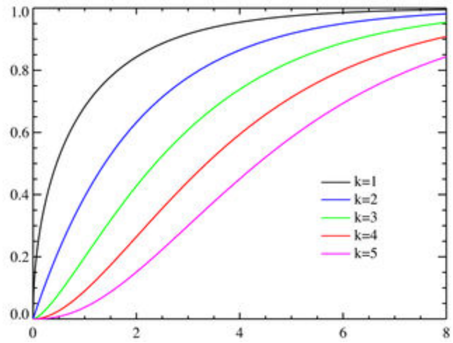
Applications:

- Used in the common chi-square tests for goodness of fit of observed data to a theoretical distribution.
- Used in statistical significance tests. One example is Friedman's analysis of variance by ranks.

Probability density function



Cumulative distribution function

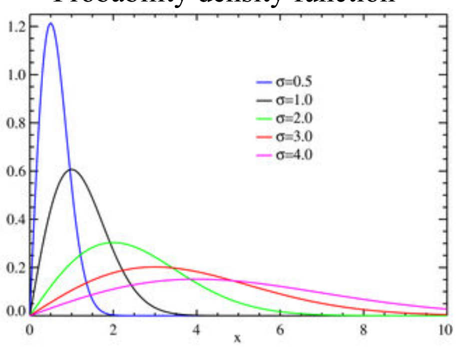


Parameters	$k > 0$ degrees of freedom
Support	$x \in [0; +\infty)$
Probability density function **(pdf)**	$\frac{(1/2)^{k/2}}{\Gamma(k/2)} x^{k/2-1} e^{-x/2}$
Cumulative distribution function **(CDF)**	$\frac{\int_0^{x/2} y^{\frac{k}{2}-1} e^{-y} dy}{\Gamma\left(\frac{k}{2}\right)}$
Mean	k
Median	approximately $k - 2/3$
Mode	$k - 2$ if $k \geq 2$
Variance	$2\,k$
Skewness	$\sqrt{8/k}$
Excess kurtosis	$12/k$
Moment-generating function **(MGF)**	$(1 - 2\,t)^{-k/2}$ for $2\,t < 1$
Characteristic function	$(1 - 2\,i\,t)^{-k/2}$

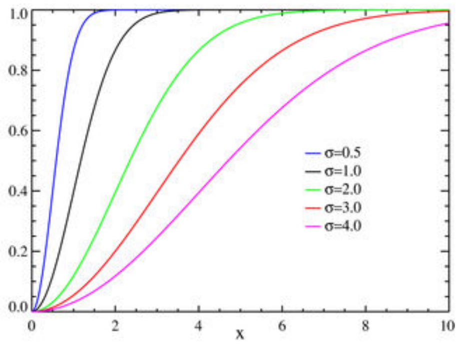
• Rayleigh distribution

The Rayleigh distribution has been used to model attenuation of wireless signals facing multi-path **fading**. The Chi, Rice and Weibull distributions are all generalizations of the Rayleigh distribution.

Probability density function



Cumulative distribution function



Parameters	$\sigma > 0$
Support	$x \in [0;\infty)$
Probability density function **(pdf)**	$\frac{x \exp\left(\frac{-x^2}{2\sigma^2}\right)}{\sigma^2}$
Cumulative distribution function **(CDF)**	$1 - \exp\left(\frac{-x^2}{2\sigma^2}\right)$
Mean	$\sigma\sqrt{\frac{\pi}{2}}$
Median	$\sigma\sqrt{\ln(4)}$
Mode	σ
Variance	$\frac{4-\pi}{2}\sigma^2$
Skewness	$\frac{2\sqrt{\pi}(\pi-3)}{(4-\pi)^{3/2}}$
Excess kurtosis	$-\frac{6\pi^2-24\pi+16}{(4-\pi)^2}$
Moment-generating function **(MGF)**	$1+\sigma t\, e^{\sigma^2t^2/2}\sqrt{\frac{\pi}{2}}\left(\text{erf}\left(\frac{\sigma t}{\sqrt{2}}\right)+1\right)$
Characteristic function	$1-\sigma t e^{-\sigma^2t^2/2}\sqrt{\frac{\pi}{2}}\left(\text{erfi}\left(\frac{\sigma t}{\sqrt{2}}\right)-i\right)$

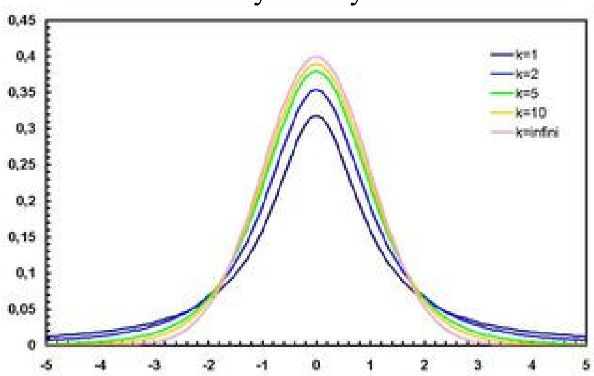
• Student's t distribution

The Student's t distribution arises in the problem of estimating the mean of a normally distributed population when the sample size is small.

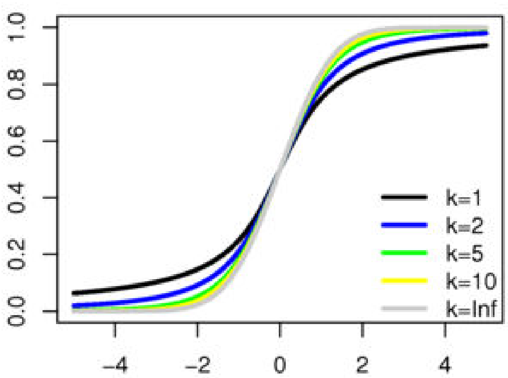
Applications:

- It is the basis of the popular Student's t-tests for the statistical significance of the difference between two sample means.

Probability density function



Cumulative distribution function



Parameters	$\nu > 0$ deg. of freedom (real)
Support	$x \in (-\infty; +\infty)$

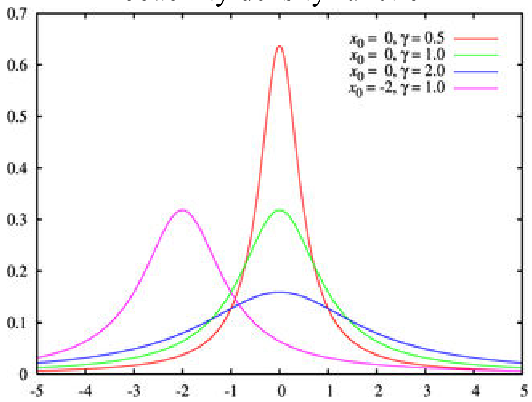
Probability density function **(pdf)**	$\frac{\Gamma((\nu+1)/2)}{\sqrt{\nu\pi}\,\Gamma(\nu/2)}(1+x^2/\nu)^{-(\nu+1)/2}$	
Cumulative distribution function **(CDF)**	$\frac{1}{2}+\frac{x\Gamma\left((\nu+1)/2\right)\,{}_2F_1\left(\frac{1}{2},(\nu+1)/2;\frac{3}{2};-\frac{x^2}{\nu}\right)}{\sqrt{\pi\nu}\,\Gamma(\nu/2)}$	where ${}_2F_1$ is the hypergeometric function
Mean	0 for $\nu > 1$, otherwise undefined	
Median	0	
Mode	0	
Variance	$\frac{\nu}{\nu-2}$ for $\nu > 2$, otherwise undefined	
Skewness	0 for $\nu > 3$	
Excess kurtosis	$\frac{6}{\nu-4}$ for $\nu > 4$	

• Cauchy distribution

Applications:

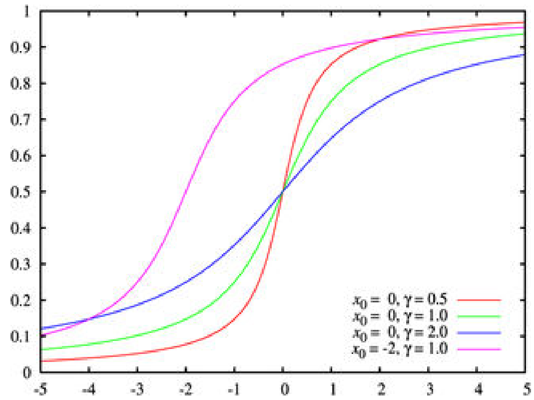
- In spectroscopy the Cauchy distribution is the description of the line shape of spectral lines which are broadened by collisions.
- In nuclear and particle physics, the energy profile of a resonance is described by the relativistic Cauchy distribution.

Probability density function



The green line is the standard Cauchy distribution

Cumulative distribution function



Colors match the pdf above

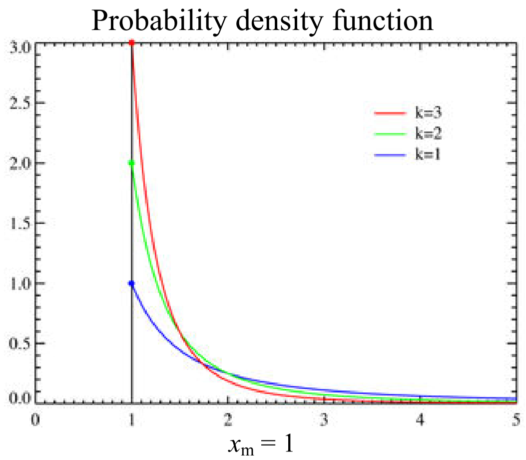
Parameters	x_0 $\gamma > 0$	location (real) scale (real)
Support	$x \in (-\infty; +\infty)$	
Probability density function **(pdf)**	$\frac{1}{\pi\gamma\left[1+\left(\frac{x-x_0}{\gamma}\right)^2\right]}$	
Cumulative distribution function **(CDF)**	$\frac{1}{\pi}\arctan\left(\frac{x-x_0}{\gamma}\right)+\frac{1}{2}$	

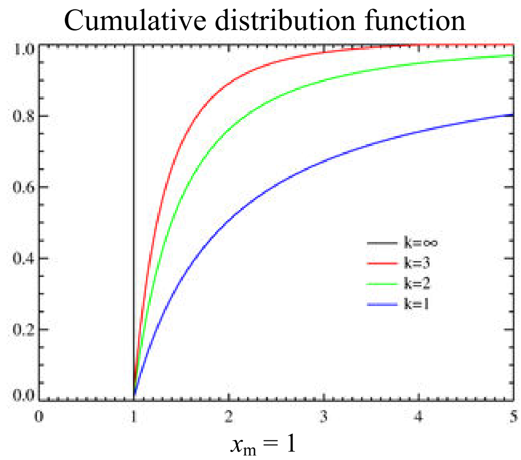
• Pareto distribution

The Pareto distribution is characterized by a curved *long tail* when plotted on a linear scale, thus is it used to describe long-range dependent phenomena.

Applications:

- Used to describe the allocation of wealth among individuals.
- Describes frequencies of words in longer texts (a few words are used often, lots of words are used infrequently).
- File size distribution of Internet traffic which uses the TCP protocol (many smaller files, few larger ones).
- Clusters of Bose-Einstein condensate near absolute zero.
- The values of oil reserves in oil fields (a few large fields, many small fields)
- The length distribution in jobs assigned to supercomputers (a few large ones, many small ones).
- The standardized price returns on individual stocks.
- Areas burnt in forest fires.

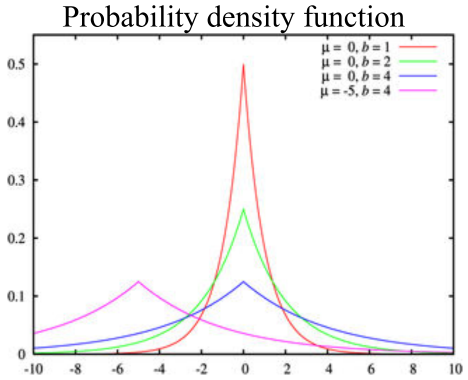


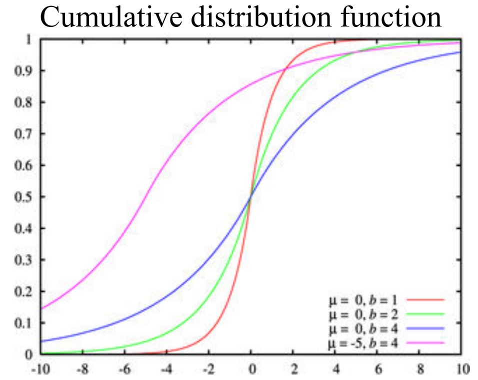


Parameters	$x_m > 0$	location (real)
	$k > 0$	shape (real)
Support	$x \in [x_m; +\infty)$	
Probability density function **(pdf)**	$\frac{k\, x_m^k}{x^{k+1}}$	
Cumulative distribution function **(CDF)**	$1 - \left(\frac{x_m}{x}\right)^k$	
Mean	$\frac{k\, x_m}{k-1}$ for $k > 1$	
Median	$x_m \sqrt[k]{2}$	
Mode	x_m	
Variance	$\frac{x_m^2 k}{(k-1)^2(k-2)}$ for $k > 2$	

Skewness	$\frac{2(1+k)}{k-3}\sqrt{\frac{k-2}{k}}$ for $k > 3$
Excess kurtosis	$\frac{6(k^3+k^2-6k-2)}{k(k-3)(k-4)}$ for $k > 4$

• Laplace distribution





Parameters	μ location (real) $b > 0$ scale (real)
Support	$x \in (-\infty; +\infty)$
Probability density function **(pdf)**	$\frac{1}{2\,b}\exp\left(-\frac{\vert x-\mu\vert}{b}\right)$ $= \frac{1}{2b}\begin{cases}\exp\left(-\frac{\mu-x}{b}\right) & \text{if } x < \mu\\ \exp\left(-\frac{x-\mu}{b}\right) & \text{if } x \geq \mu\end{cases}$
Cumulative distribution function **(CDF)**	$\begin{cases}\frac{1}{2}\exp\left(-\frac{\mu-x}{b}\right) & \text{if } x < \mu\\ 1-\frac{1}{2}\exp\left(-\frac{x-\mu}{b}\right) & \text{if } x \geq \mu\end{cases}$
Mean	μ
Median	μ
Mode	μ
Variance	$2\,b^2$
Skewness	0
Excess kurtosis	3
Moment-generating function **(MGF)**	$\frac{\exp(\mu\,t)}{1-b^2\,t^2}$ for $\vert t\vert < 1/b$
Characteristic function	$\frac{\exp(\mu\,i\,t)}{1+b^2\,t^2}$

Notes:
The expected value for a discrete random variable is:

$$\mu = E[X] = \sum_i p_i x_i$$

For a continuous random variable:

$$\mu = E[X] = \int_{-\infty}^{\infty} x f(x) dx$$

The **variance** of a random variable is a measure of its statistical dispersion, indicating how its possible values are spread around the expected value. For a random variable X, the variance is:

$$\operatorname{Var}(X) = \operatorname{E}((X - \mu)^2).$$

If the random variable is discrete the variance is equivalent to:

$$\operatorname{Var}[X] = \sum_{i=1}^{n} (x_i - \mu)^2 p_i$$

The **mode** is the most frequent value assumed by a random variable, or occurring in a sampling of a random variable. The mode is not necessarily unique, since the same maximum frequency may be attained at different values.

The **median** is the number separating the higher half of a sample or a probability distribution, from the lower half. The *median* of a finite list of numbers can be found by arranging all the observations from lowest value to highest value and picking the middle one. If there is an even number of observations, the median is not unique.

Skewness is a measure of the *asymmetry* of the probability density function.

Kurtosis (from the Greek word kurtos) is a measure of the *peakedness* of the probability density function.