Exponential distribution

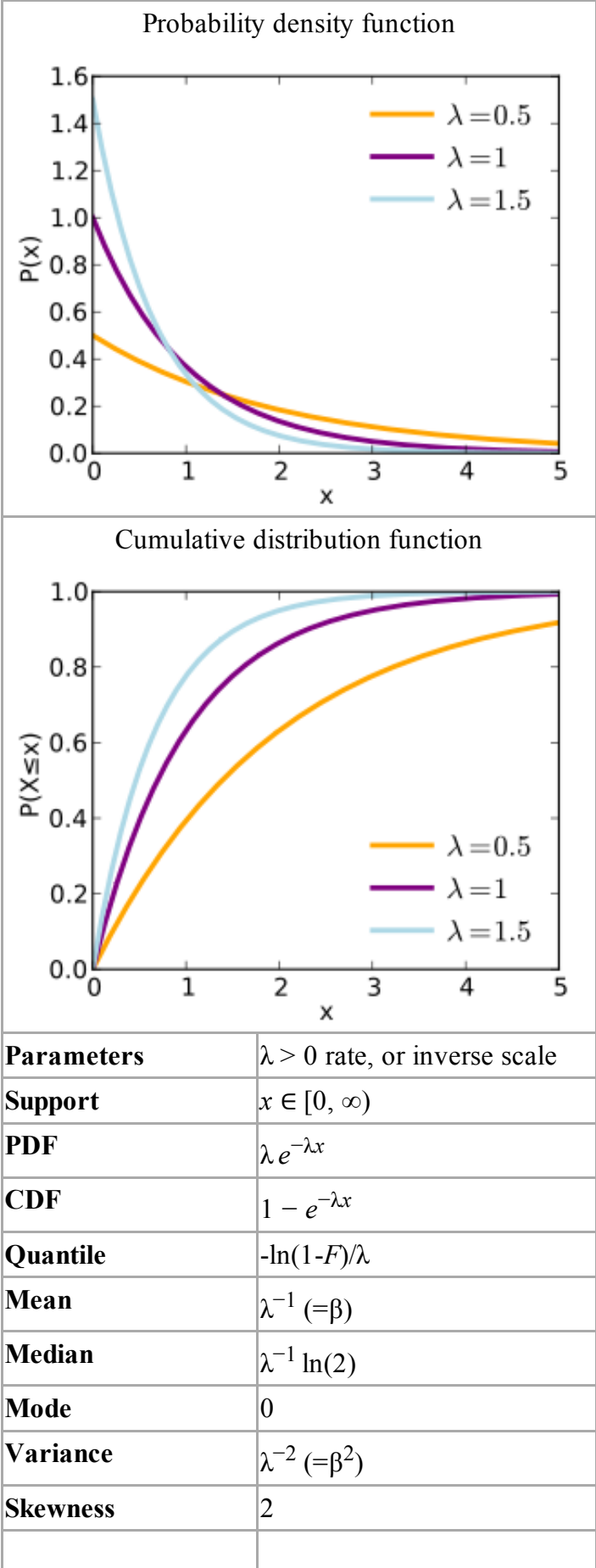From Wikipedia, the free encyclopedia

In probability theory and statistics, the **exponential distribution** (a.k.a. **negative exponential distribution**) is the probability distribution that describes the time between events in a Poisson process, i.e. a process in which events occur continuously and independently at a constant average rate. It is a particular case of the gamma distribution. It is the continuous analogue of the geometric distribution, and it has the key property of being memoryless. In addition to being used for the analysis of Poisson processes, it is found in various other contexts.

The exponential distribution is not the same as the class of exponential families of distributions, which is a large class of probability distributions that includes the exponential distribution as one of its members, but also includes the normal distribution, binomial distribution, gamma distribution, Poisson, and many others.

Contents

Exponential

Probability density function

Cumulative distribution function

Parameters	$\lambda > 0$ rate, or inverse scale
Support	$x \in [0, \infty)$
PDF	$\lambda e^{-\lambda x}$
CDF	$1 - e^{-\lambda x}$
Quantile	$-\ln(1-F)/\lambda$
Mean	λ^{-1} $(=\beta)$
Median	$\lambda^{-1} \ln(2)$
Mode	0
Variance	λ^{-2} $(=\beta^2)$
Skewness	2

Ex. kurtosis	6
Entropy	$\log(e/\lambda)$
MGF	$\dfrac{\lambda}{\lambda - t}$, for $t < \lambda$
CF	$\dfrac{\lambda}{\lambda - it}$
Fisher information	λ^{-2}

Characterization

Probability density function

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

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

Alternatively, this can be defined using the right-continuous Heaviside step function, $H(x)$ where $H(0)=1$:

$$f(x; \lambda) = \lambda e^{-\lambda x} H(x)$$

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.

Cumulative distribution function

The cumulative distribution function is given by

$$F(x; \lambda) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0, \\ 0 & x < 0. \end{cases}$$

Alternatively, this can be defined using the Heaviside step function, $H(x)$.

$$F(x; \lambda) = (1 - e^{-\lambda x})H(x)$$

Alternative parameterization

A commonly used alternative parametrization is to define the probability density function (pdf) of an exponential distribution as

$$f(x; \beta) = \begin{cases} \frac{1}{\beta}e^{-\frac{x}{\beta}} & x \geq 0, \\ 0 & x < 0. \end{cases}$$

where $\beta > 0$ is mean, standard deviation, and scale parameter of the distribution, the reciprocal of the *rate parameter*, λ, defined above. In this specification, β is a *survival parameter* in the sense that if a random variable X is the duration of time that a given biological or mechanical system manages to survive and $X \sim \text{Exp}(\beta)$ then $E[X] = \beta$. That is to say, the expected duration of survival of the system is β units of time. The parametrization involving the "rate" parameter arises in the context of events arriving at a rate λ, when the time between events (which might be modeled using an exponential distribution) has a mean of $\beta = \lambda^{-1}$.

The alternative specification is sometimes more convenient than the one given above, and some authors will use it as a standard definition. This alternative specification is not used here. Unfortunately this gives rise to a notational ambiguity. In general, the reader must check which of these two specifications is being used if an author writes "$X \sim \text{Exp}(\lambda)$", since either the notation in the previous (using λ) or the notation in this section (here, using β to avoid confusion) could be intended. An example of this notational switch: reference[1] uses λ for β.

Properties

Mean, variance, moments and median

The mean or expected value of an exponentially distributed random variable X with rate parameter λ is given by

$$E[X] = \frac{1}{\lambda} = \beta, \text{ see above.}$$

In light of the examples given above, this makes sense: if you receive phone calls at an average rate of 2 per hour, then you can expect to wait half an hour for every call.

The variance of X is given by

$$\text{Var}[X] = \frac{1}{\lambda^2},$$

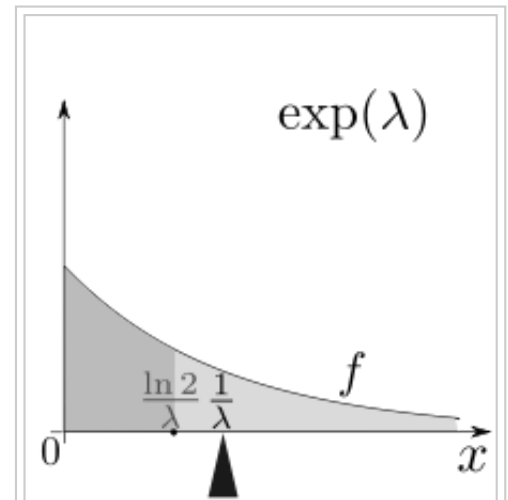so the standard deviation is equal to the mean.

The moments of X, for $n = 1, 2, \ldots$, are given by
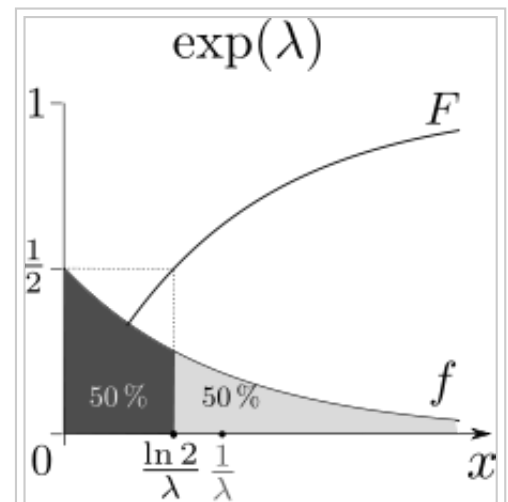
$$E[X^n] = \frac{n!}{\lambda^n}.$$

The median of X is given by

$$m[X] = \frac{\ln(2)}{\lambda} < E[X],$$

where ln refers to the natural logarithm. Thus the absolute difference between the mean and median is

The mean is the probability mass centre, that is the first moment.

The median is the preimage $F^{-1}(1/2)$.

$$|E[X] - m[X]| = \frac{1 - \ln(2)}{\lambda} < \frac{1}{\lambda} = \text{standard deviation},$$

in accordance with the median-mean inequality.

Memorylessness

An exponentially distributed random variable T obeys the relation

$$\Pr\left(T > s + t \mid T > s\right) = \Pr\left(T > t\right), \qquad \forall s, t \geq 0.$$

When T is interpreted as the waiting time for an event to occur relative to some initial time, this relation implies that, if T is conditioned on a failure to observe the event over some initial period of time s, the distribution of the remaining waiting time is the same as the original unconditional distribution. For example, if an event has not occurred after 30 seconds, the conditional probability that occurrence will take at least 10 more seconds is equal to the unconditional probability of observing the event more than 10 seconds relative to the initial time.

The exponential distribution and the geometric distribution are the only memoryless probability distributions.

The exponential distribution is consequently also necessarily the only continuous probability distribution that has a constant Failure rate.

Quantiles

The quantile function (inverse cumulative distribution function) for Exp(λ) is

$$F^{-1}(p; \lambda) = \frac{-\ln(1-p)}{\lambda}, \qquad 0 \le p < 1$$

The quartiles are therefore:

- first quartile: ln(4/3)/λ
- median: ln(2)/λ
- third quartile: ln(4)/λ

And as a consequence the interquartile range is ln(3)/λ.

Tukey criteria for anomalies.[2]

Kullback–Leibler divergence

The directed Kullback–Leibler divergence of e^{λ} ('approximating' distribution) from e^{λ_0} ('true' distribution) is given by

$$\Delta(\lambda_0 || \lambda) = \log(\lambda_0) - \log(\lambda) + \frac{\lambda}{\lambda_0} - 1.$$

Maximum entropy distribution

Among all continuous probability distributions with support $[0, \infty)$ and mean μ, the exponential distribution with $\lambda = 1/\mu$ has the largest differential entropy. In other words, it is the maximum entropy probability distribution for a random variate X which is greater than or equal to zero and for which E[X] is fixed.[3]

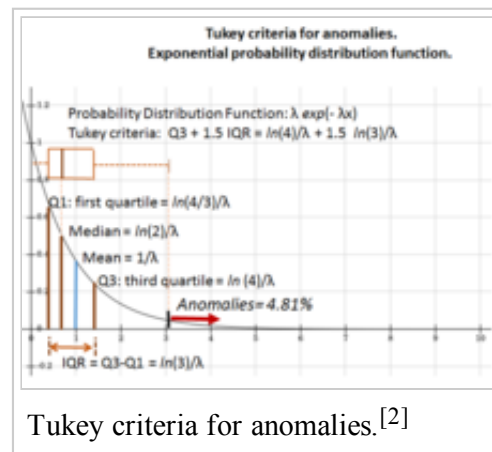## Distribution of the minimum of exponential random variables

Let $X_1, ..., X_n$ be independent exponentially distributed random variables with rate parameters $\lambda_1, ..., \lambda_n$. Then

$$\min \{X_1, \ldots, X_n\}$$

is also exponentially distributed, with parameter

$$\lambda = \lambda_1 + \cdots + \lambda_n.$$

This can be seen by considering the complementary cumulative distribution function:

$$\Pr\left(\min\{X_1,\ldots,X_n\} > x\right) = \Pr\left(X_1 > x,\ldots,X_n > x\right)$$
$$= \prod_{i=1}^{n}\Pr(X_i > x)$$
$$= \prod_{i=1}^{n}\exp(-x\lambda_i) = \exp\left(-x\sum_{i=1}^{n}\lambda_i\right).$$

The index of the variable which achieves the minimum is distributed according to the law

$$\Pr\left(X_k = \min\{X_1,\ldots,X_n\}\right) = \frac{\lambda_k}{\lambda_1 + \cdots + \lambda_n}.$$

Note that

$$\max\{X_1,\ldots,X_n\}$$

is not exponentially distributed.

Parameter estimation

Suppose random variable X is exponentially distributed with rate parameter λ, and $x_1,\ldots,x_n$ are n independent samples from X, with sample mean $\bar{x}$. Among the estimators of λ, the maximum likelihood estimator (MLE) is $\frac{1}{\bar{x}}$ and the uniformly minimum variance unbiased estimator (UMVUE) is $\left(\frac{n-1}{n}\right)\left(\frac{1}{\bar{x}}\right)$. Of the estimators of the form $\left(\frac{k}{n}\right)\left(\frac{1}{\bar{x}}\right)$, the one that minimizes the expected mean squared error is $\left(\frac{n-2}{n}\right)\left(\frac{1}{\bar{x}}\right)$.[4]

Maximum likelihood

The likelihood function for λ, given an independent and identically distributed sample $x = (x_1, ..., x_n)$ drawn from the variable, is:

$$L(\lambda) = \prod_{i=1}^{n}\lambda\exp(-\lambda x_i) = \lambda^n\exp\left(-\lambda\sum_{i=1}^{n}x_i\right) = \lambda^n\exp\left(-\lambda n\bar{x}\right),$$

where:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n}x_i$$

is the sample mean.

The derivative of the likelihood function's logarithm is:

$$\frac{d}{d\lambda}\ln(L(\lambda)) = \frac{d}{d\lambda}\left(n\ln(\lambda) - \lambda n\bar{x}\right) = \frac{n}{\lambda} - n\bar{x} \begin{cases} > 0, & 0 < \lambda < \frac{1}{\bar{x}}, \\ = 0, & \lambda = \frac{1}{\bar{x}}, \\ < 0, & \lambda > \frac{1}{\bar{x}}. \end{cases}$$

Consequently the maximum likelihood estimate for the rate parameter is:

$$\widehat{\lambda} = \frac{1}{\bar{x}}.$$

Although this is *not* an unbiased estimator of λ, $\bar{x}$ *is* an unbiased[5] MLE[6] estimator of $1/\lambda = \beta$, where β is the scale parameter defined in the 'Alternative parameterization' section above and the distribution mean.

Confidence intervals

The $100(1-\alpha)\%$ confidence interval for the rate parameter of an exponential distribution is given by:[7]

$$\frac{2n}{\widehat{\lambda}\chi^2_{1-\frac{\alpha}{2},2n}} < \frac{1}{\lambda} < \frac{2n}{\widehat{\lambda}\chi^2_{\frac{\alpha}{2},2n}}$$

which is also equal to:

$$\frac{2n\bar{x}}{\chi^2_{1-\frac{\alpha}{2},2n}} < \frac{1}{\lambda} < \frac{2n\bar{x}}{\chi^2_{\frac{\alpha}{2},2n}}$$

where $\chi^2_{p,v}$ is the $100(p)$ percentile of the chi squared distribution with v degrees of freedom, n is the number of observations of inter-arrival times in the sample, and x-bar is the sample average. A simple approximation to the exact interval endpoints can be derived using a normal approximation to the $\chi^2_{p,v}$ distribution. This approximation gives the following values for a 95% confidence interval:

$$\lambda_{low} = \widehat{\lambda}\left(1 - \frac{1.96}{\sqrt{n}}\right)$$
$$\lambda_{upp} = \widehat{\lambda}\left(1 + \frac{1.96}{\sqrt{n}}\right)$$

This approximation may be acceptable for samples containing at least 15 to 20 elements.[8]

Bayesian inference

The conjugate prior for the exponential distribution is the gamma distribution (of which the exponential distribution is a special case). The following parameterization of the gamma probability density function is useful:

$$\mathrm{Gamma}(\lambda; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}\exp(-\lambda\beta).$$

The posterior distribution p can then be expressed in terms of the likelihood function defined above and a gamma prior:

$$p(\lambda) \propto L(\lambda) \times \mathrm{Gamma}(\lambda; \alpha, \beta)$$
$$= \lambda^n \exp\left(-\lambda n\bar{x}\right) \times \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} \exp(-\lambda\beta)$$
$$\propto \lambda^{(\alpha+n)-1} \exp(-\lambda\left(\beta + n\bar{x}\right)).$$

Now the posterior density p has been specified up to a missing normalizing constant. Since it has the form of a gamma pdf, this can easily be filled in, and one obtains:

$$p(\lambda) = \mathrm{Gamma}(\lambda; \alpha + n, \beta + n\bar{x}).$$

Here the hyperparameter α can be interpreted as the number of prior observations, and β as the sum of the prior observations. The posterior mean here is:

$$\frac{\alpha + n}{\beta + n\bar{x}}.$$

Generating exponential variates

A conceptually very simple method for generating exponential variates is based on inverse transform sampling: Given a random variate U drawn from the uniform distribution on the unit interval (0, 1), the variate

$$T = F^{-1}(U)$$

has an exponential distribution, where F^{-1} is the quantile function, defined by

$$F^{-1}(p) = \frac{-\ln(1-p)}{\lambda}.$$

Moreover, if U is uniform on (0, 1), then so is $1 - U$. This means one can generate exponential variates as follows:

$$T = \frac{-\ln(U)}{\lambda}.$$

Other methods for generating exponential variates are discussed by Knuth[9] and Devroye.[10]

The ziggurat algorithm is a fast method for generating exponential variates.

A fast method for generating a set of ready-ordered exponential variates without using a sorting routine is also available.[10]

Related distributions

- Exponential distribution is closed under scaling by a positive factor. If $X \sim \mathrm{Exp}(\lambda)$ then $kX \sim \mathrm{Exp}(\lambda/k)$.

- If $X \sim \text{Exp}(1/2)$ then $X \sim \chi_2^2$, i.e. X has a chi-squared distribution with 2 degrees of freedom. Hence:

$$\exp(\lambda) \sim \frac{1}{2\lambda}\chi_2^2 \Rightarrow \sum_{i=1}^{n} \exp(\lambda) \sim \frac{1}{2\lambda}\chi_{2n}^2$$

- If $X \sim \text{Exp}(\lambda)$ and $Y \sim \text{Exp}(\nu)$ then $\min(X, Y) \sim \text{Exp}(\lambda + \nu)$.
- If $X_i \sim \text{Exp}(\lambda)$ then $\min\{X_1, ..., X_n\} \sim \text{Exp}(n\lambda)$.
- The Benktander Weibull distribution reduces to a truncated exponential distribution. If $X \sim \text{Exp}(\lambda)$ then $1+X \sim \text{BenktanderWeibull}(\lambda, 1)$.
- The exponential distribution is a limit of a scaled beta distribution:

$$\lim_{n \to \infty} n\text{Beta}(1, n) = \text{Exp}(1).$$

- If $X_i \sim \text{Exp}(\lambda)$ then the sum $X_1 + ... + X_k = \sum_i X_i \sim \text{Erlang}(k, \lambda)$ which is just a Gamma(k, λ^{-1}) (in (k, θ) parametrization) or Gamma(k, λ) (in (α, β) parametrization) with an integer shape parameter k.
- If $X \sim \text{Exp}(1)$ then $\mu - \sigma \log(X) \sim \text{GEV}(\mu, \sigma, 0)$.
- If $X \sim \text{Exp}(\lambda)$ then $X \sim \text{Gamma}(1, \lambda^{-1})$ (in (k, θ) parametrization) or Gamma$(1, \lambda)$ (in (α, β) parametrization).
- If $X \sim \text{Exp}(\lambda)$ and $Y \sim \text{Exp}(\nu)$ then $\lambda X - \nu Y \sim \text{Laplace}(0, 1)$.
- If $X, Y \sim \text{Exp}(\lambda)$ then $X - Y \sim \text{Laplace}(0, \lambda^{-1})$.
- If $X \sim \text{Laplace}(\mu, \beta^{-1})$ then $|X - \mu| \sim \text{Exp}(\beta)$.
- If $X \sim \text{Exp}(1)$ then (logistic distribution):

$$\mu - \beta \log\left(\frac{e^{-X}}{1-e^{-X}}\right) \sim \text{Logistic}(\mu, \beta)$$

- If $X, Y \sim \text{Exp}(1)$ then (logistic distribution):

$$\mu - \beta \log\left(\frac{X}{Y}\right) \sim \text{Logistic}(\mu, \beta)$$

- If $X \sim \text{Exp}(\lambda)$ then $ke^X \sim \text{Pareto}(k, \lambda)$.
- If $X \sim \text{Pareto}(1, \lambda)$ then $\log(X) \sim \text{Exp}(\lambda)$.
- Exponential distribution is a special case of type 3 Pearson distribution.
- If $X \sim \text{Exp}(\lambda)$ then $\frac{e^{-X}}{k} \sim \text{PowerLaw}(k, \lambda)$ (power law)
- If $X \sim \text{Rayleigh}(\lambda^{-1/2})$ then $X^2/2 \sim \text{Exp}(\lambda)$.
- If $X \sim \text{Exp}(\lambda)$ then $X \sim \text{Weibull}(\frac{1}{\lambda}, 1)$ (Weibull distribution)
- If $X_i \sim U(0, 1)$ then

$$\lim_{n \to \infty} n \min(X_1, \ldots, X_n) \sim \text{Exp}(1)$$

- If $Y|X \sim \text{Poisson}(X)$ where $X \sim \text{Exp}(\lambda^{-1})$ then $Y \sim \text{Geometric}(\frac{1}{1+\lambda})$ (geometric distribution)
- If $X \sim \text{Exp}(1)$ and $Y \sim \Gamma(\alpha, \frac{\beta}{\alpha})$ then $\sqrt{XY} \sim K(\alpha, \beta)$ (K-distribution)
- The Hoyt distribution can be obtained from Exponential distribution and Arcsine distribution
- If $X \sim \text{Exp}(\lambda)$ and $Y \sim \text{Erlang}(n, \lambda)$ then:

$$\frac{X}{Y} \sim \text{Pareto}(1, n)$$

- If $X \sim \text{Exp}(\lambda)$ and $Y \sim \Gamma(n, \frac{1}{\lambda})$ then $\frac{X}{Y} \sim \text{Pareto}(1, n)$
- If $X \sim \text{SkewLogistic}(\theta)$, then $\log(1 + e^{-X}) \sim \text{Exp}(\theta)$.
- If $X \sim \text{Exp}(\lambda)$ and $Y = \mu - \beta \log(X\lambda)$ then $Y \sim \text{Gumbel}(\mu, \beta)$.

- Let $X \sim \text{Exp}(\lambda_X)$ and $Y \sim \text{Exp}(\lambda_Y)$ be independent. Then $Z = \frac{\lambda_X X}{\lambda_Y Y}$ has probability density function $f_Z(z) = \frac{1}{(z+1)^2}$. This can be used to obtain a confidence interval for $\frac{\lambda_X}{\lambda_Y}$.

Other related distributions:

- Hyper-exponential distribution – the distribution whose density is a weighted sum of exponential densities.
- Hypoexponential distribution – the distribution of a general sum of exponential random variables.
- exGaussian distribution – the sum of an exponential distribution and a normal distribution.

Applications of exponential distribution

Occurrence of events

The exponential distribution occurs naturally when describing the lengths of the inter-arrival times in a homogeneous Poisson process.

The exponential distribution may be viewed as a continuous counterpart of the geometric distribution, which describes the number of Bernoulli trials necessary for a *discrete* process to change state. In contrast, the exponential distribution describes the time for a continuous process to change state.

In real-world scenarios, the assumption of a constant rate (or probability per unit time) is rarely satisfied. For example, the rate of incoming phone calls differs according to the time of day. But if we focus on a time interval during which the rate is roughly constant, such as from 2 to 4 p.m. during work days, the exponential distribution can be used as a good approximate model for the time until the next phone call arrives. Similar caveats apply to the following examples which yield approximately exponentially distributed variables:

- The time until a radioactive particle decays, or the time between clicks of a geiger counter
- The time it takes before your next telephone call
- The time until default (on payment to company debt holders) in reduced form credit risk modeling

Exponential variables can also be used to model situations where certain events occur with a constant probability per unit length, such as the distance between mutations on a DNA strand, or between roadkills on a given road.

In queuing theory, the service times of agents in a system (e.g. how long it takes for a bank teller etc. to serve a customer) are often modeled as exponentially distributed variables. (The arrival of customers for instance is also modeled by the Poisson distribution if the arrivals are independent and distributed

identically.) The length of a process that can be thought of as a sequence of several independent tasks follows the Erlang distribution (which is the distribution of the sum of several independent exponentially distributed variables).

Reliability theory and reliability engineering also make extensive use of the exponential distribution. Because of the *memoryless* property of this distribution, it is well-suited to model the constant hazard rate portion of the bathtub curve used in reliability theory. It is also very convenient because it is so easy to add failure rates in a reliability model. The exponential distribution is however not appropriate to model the overall lifetime of organisms or technical devices, because the "failure rates" here are not constant: more failures occur for very young and for very old systems.

Fitted cumulative exponential distribution to annually maximum 1-day rainfalls using CumFreq[11]

In physics, if you observe a gas at a fixed temperature and pressure in a uniform gravitational field, the heights of the various molecules also follow an approximate exponential distribution, known as the Barometric formula. This is a consequence of the entropy property mentioned below.

In hydrology, the exponential distribution is used to analyze extreme values of such variables as monthly and annual maximum values of daily rainfall and river discharge volumes.[12]

> The blue picture illustrates an example of fitting the exponential distribution to ranked annually maximum one-day rainfalls showing also the 90% confidence belt based on the binomial distribution. The rainfall data are represented by plotting positions as part of the cumulative frequency analysis.
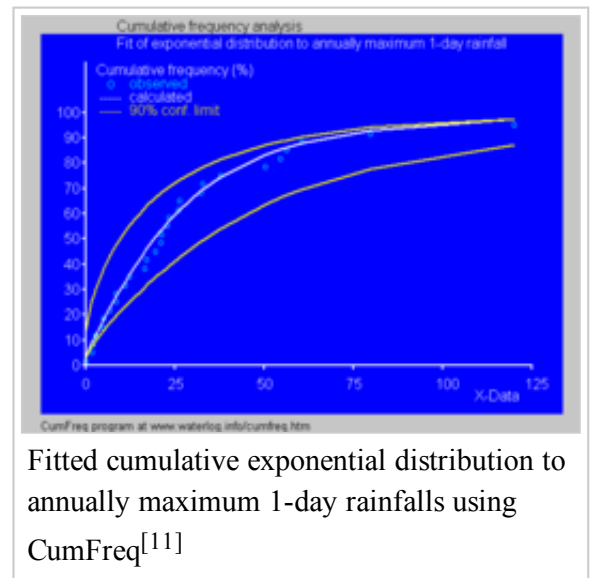
Prediction

Having observed a sample of *n* data points from an unknown exponential distribution a common task is to use these samples to make predictions about future data from the same source. A common predictive distribution over future samples is the so-called plug-in distribution, formed by plugging a suitable estimate for the rate parameter λ into the exponential density function. A common choice of estimate is the one provided by the principle of maximum likelihood, and using this yields the predictive density over a future sample x_{n+1}, conditioned on the observed samples $x = (x_1, ..., x_n)$ given by

$$p_{\mathrm{ML}}(x_{n+1} \mid x_1, \ldots, x_n) = \left(\frac{1}{\overline{x}}\right) \exp\left(-\frac{x_{n+1}}{\overline{x}}\right)$$

The Bayesian approach provides a predictive distribution which takes into account the uncertainty of the estimated parameter, although this may depend crucially on the choice of prior.

A predictive distribution free of the issues of choosing priors that arise under the subjective Bayesian approach is

$$p_{\mathrm{CNML}}(x_{n+1} \mid x_1, \ldots, x_n) = \frac{n^{n+1} \left(\overline{x}\right)^n}{\left(n\overline{x} + x_{n+1}\right)^{n+1}},$$

which can be considered as

- (1) a frequentist confidence distribution, obtained from the distribution of the pivotal quantity $x_{n+1}/\overline{x}$;[13]
- (2) a profile predictive likelihood, obtained by eliminating the parameter λ from the joint likelihood of x_{n+1} and λ by maximization;[14]
- (3) an objective Bayesian predictive posterior distribution, obtained using the non-informative Jeffreys prior $1/\lambda$;
- (4) the Conditional Normalized Maximum Likelihood (CNML) predictive distribution, from information theoretic considerations.[15]

The accuracy of a predictive distribution may be measured using the distance or divergence between the true exponential distribution with rate parameter, λ_0, and the predictive distribution based on the sample x. The Kullback–Leibler divergence is a commonly used, parameterisation free measure of the difference between two distributions. Letting $\Delta(\lambda_0\|p)$ denote the Kullback–Leibler divergence between an exponential with rate parameter λ_0 and a predictive distribution p it can be shown that

$$\mathrm{E}_{\lambda_0}\left[\Delta(\lambda_0\;\|\;p_{\mathrm{ML}})\right] = \psi(n) + \frac{1}{n-1} - \log(n)$$

$$\mathrm{E}_{\lambda_0}\left[\Delta(\lambda_0\;\|\;p_{\mathrm{CNML}})\right] = \psi(n) + \frac{1}{n} - \log(n)$$

where the expectation is taken with respect to the exponential distribution with rate parameter $\lambda_0 \in (0, \infty)$, and $\psi(\,\cdot\,)$ is the digamma function. It is clear that the CNML predictive distribution is strictly superior to the maximum likelihood plug-in distribution in terms of average Kullback–Leibler divergence for all sample sizes $n > 0$.

See also

- Dead time – an application of exponential distribution to particle detector analysis.
- Laplace distribution, or the "double exponential distribution".
- Relationships among probability distributions

References

1. David Olive, *Chapter 4. Truncated Distributions (http://lagrange.math.siu.edu/Olive/ch4.pdf)*, "Lemma 4.3", Southern Illinois University, February 18, 2010, p.107.
2. Template:Ref-missing
3. Park, Sung Y.; Bera, Anil K. (2009). "Maximum entropy autoregressive conditional heteroskedasticity model" (PDF). *Journal of Econometrics* (Elsevier): 219–230. Retrieved 2011-06-02.
4. Abdulaziz Elfessi and David M. Reineke, "A Bayesian Look at Classical Estimation: The Exponential Distribution (http://www.amstat.org/publications/jse/v9n1/elfessi.html)", *Journal of Statistics Education* Volume 9, Number 1 (2001).
5. Richard Arnold Johnson; Dean W. Wichern (2007). *Applied Multivariate Statistical Analysis*. Pearson Prentice Hall. ISBN 978-0-13-187715-3. Retrieved 10 August 2012.
6. *NIST/SEMATECH e-Handbook of Statistical Methods (http://www.itl.nist.gov/div898/handbook/eda/section3/eda3667.htm)*
7. Ross, Sheldon M. (2009). *Introduction to probability and statistics for engineers and scientists* (4th ed.). Associated Press. p. 267. ISBN 978-0-12-370483-2.
8. Guerriero, V. (2012). "Power Law Distribution: Method of Multi-scale Inferential Statistics". *Journal of*

Modern Mathematics Frontier (JMMF) **1**: 21–28.

9. Donald E. Knuth (1998). *The Art of Computer Programming*, volume 2: *Seminumerical Algorithms*, 3rd edn. Boston: Addison–Wesley. ISBN 0-201-89684-2. *See section 3.4.1, p. 133.*

10. Luc Devroye (1986). *Non-Uniform Random Variate Generation (http://luc.devroye.org/rnbookindex.html)*. New York: Springer-Verlag. ISBN 0-387-96305-7. *See chapter IX (http://luc.devroye.org/chapter_nine.pdf), section 2, pp. 392–401.*

11. "Cumfreq, a free computer program for cumulative frequency analysis".

12. Ritzema (ed.), H.P. (1994). *Frequency and Regression Analysis* (PDF). Chapter 6 in: Drainage Principles and Applications, Publication 16, International Institute for Land Reclamation and Improvement (ILRI), Wageningen, The Netherlands. pp. 175–224. ISBN 90-70754-33-9.

13. Lawless, J.F., Fredette, M.,"Frequentist predictions intervals and predictive distributions", Biometrika (2005), Vol 92, Issue 3, pp 529–542.

14. Bjornstad, J.F. (1990). "Predictive Likelihood: A Review". *Statist. Sci.* **5** (2): 242–254. doi:10.1214/ss/1177012175.

15. D. F. Schmidt and E. Makalic, "Universal Models for the Exponential Distribution (http://www.emakalic.org/blog/wp-content/uploads/2010/04/SchmidtMakalic09b.pdf)", *IEEE Transactions on Information Theory*, Volume 55, Number 7, pp. 3087–3090, 2009 doi:10.1109/TIT.2009.2018331 (https://dx.doi.org/10.1109%2FTIT.2009.2018331)

External links

- Hazewinkel, Michiel, ed. (2001), "Exponential distribution", *Encyclopedia of Mathematics*, Springer, ISBN 978-1-55608-010-4
- Online calculator of Exponential Distribution (http://www.elektro-energetika.cz/calculations/ex.php?language=english)