WIKIPEDIA

Rayleigh distribution

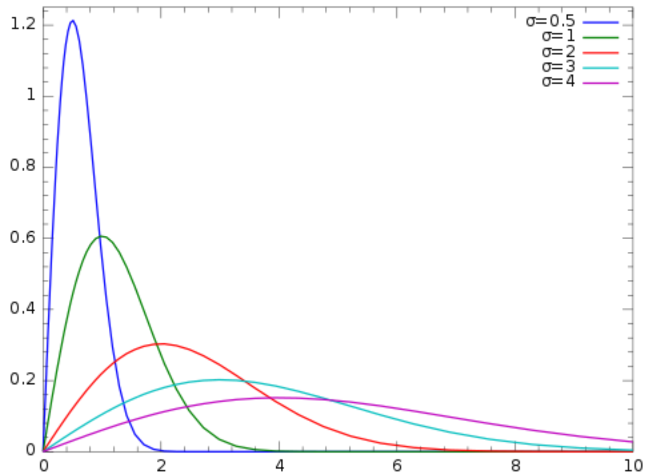
In probability theory and statistics, the **Rayleigh distribution** is a continuous probability distribution for nonnegative-valued random variables. It is essentially a chi distribution with two degrees of freedom.

A Rayleigh distribution is often observed when the overall magnitude of a vector is related to its directional components. One example where the Rayleigh distribution naturally arises is when wind velocity is analyzed in two dimensions. Assuming that each component is uncorrelated, normally distributed with equal variance, and zero mean, then the overall wind speed (vector magnitude) will be characterized by a Rayleigh distribution. A second example of the distribution arises in the case of random complex numbers whose real and imaginary components are independently and identically distributed Gaussian with equal variance and zero mean. In that case, the absolute value of the complex number is Rayleigh-distributed.

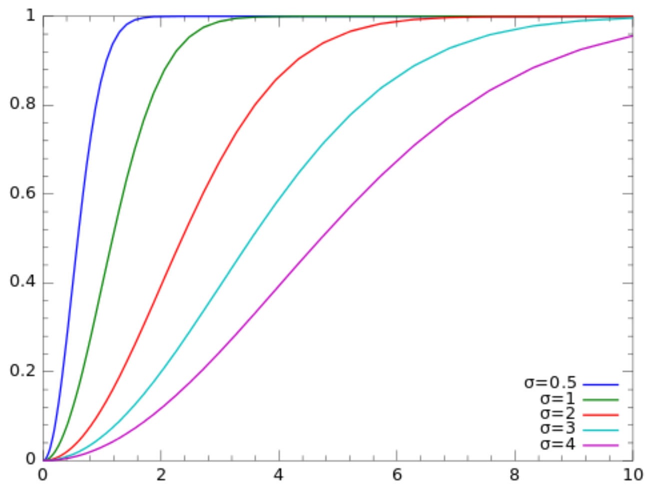
The distribution is named after Lord Rayleigh (/ˈreɪli/).[1]

Rayleigh

Probability density function



Cumulative distribution function



Parameters	scale: $\sigma > 0$
Support	$x \in [0, \infty)$
PDF	$\frac{x}{\sigma^2} e^{-x^2/(2\sigma^2)}$
CDF	$1 - e^{-x^2/(2\sigma^2)}$
Quantile	$Q(F;\sigma) = \sigma\sqrt{-2\ln(1-F)}$
Mean	$\sigma\sqrt{\frac{\pi}{2}}$
Median	$\sigma\sqrt{2\ln(2)}$
Mode	σ
Variance	$\frac{4-\pi}{2}\sigma^2$
Skewness	

Contents

Definition

The probability density function of the Rayleigh distribution is[2]

$$f(x;\sigma) = \frac{x}{\sigma^2} e^{-x^2/(2\sigma^2)}, \quad x \geq 0,$$

where σ is the scale parameter of the distribution. The cumulative distribution function is[2]

$$F(x;\sigma) = 1 - e^{-x^2/(2\sigma^2)}$$

for $x \in [0, \infty)$.

	$\frac{2\sqrt{\pi}(\pi-3)}{(4-\pi)^{3/2}}$
Ex. kurtosis	$-\frac{6\pi^2 - 24\pi + 16}{(4-\pi)^2}$
Entropy	$1 + \ln\left(\frac{\sigma}{\sqrt{2}}\right) + \frac{\gamma}{2}$
MGF	$1 + \sigma t e^{\sigma^2 t^2/2} \sqrt{\frac{\pi}{2}} \left(\operatorname{erf}\left(\frac{\sigma t}{\sqrt{2}}\right) + 1\right)$
CF	$1 - \sigma t e^{-\sigma^2 t^2/2} \sqrt{\frac{\pi}{2}} \left(\operatorname{erfi}\left(\frac{\sigma t}{\sqrt{2}}\right) - i\right)$

Relation to random vector length

Consider the two-dimensional vector $Y = (U, V)$ which has components that are normally distributed, centered at zero, and independent. Then U and V have density functions

$$f_U(x;\sigma) = f_V(x;\sigma) = \frac{e^{-x^2/(2\sigma^2)}}{\sqrt{2\pi\sigma^2}}.$$

Let X be the length of Y. That is, $X = \sqrt{U^2 + V^2}$. Then X has cumulative distribution function

$$F_X(x;\sigma) = \iint_{D_x} f_U(u;\sigma) f_V(v;\sigma)\, dA,$$

where D_x is the disk

$$D_x = \left\{(u,v) : \sqrt{u^2+v^2} < x\right\}.$$

Writing the double integral in polar coordinates, it becomes

$$F_X(x;\sigma) = \frac{1}{2\pi\sigma^2} \int_0^{2\pi} \int_0^x r e^{-r^2/(2\sigma^2)}\, dr\, d\theta = \frac{1}{\sigma^2} \int_0^x r e^{-r^2/(2\sigma^2)}\, dr.$$

Finally, the probability density function for X is the derivative of its cumulative distribution function, which by the fundamental theorem of calculus is

$$f_X(x;\sigma) = \frac{d}{dx} F_X(x;\sigma) = \frac{x}{\sigma^2} e^{-x^2/(2\sigma^2)},$$

which is the Rayleigh distribution. It is straightforward to generalize to vectors of dimension other than 2. There are also generalizations when the components have unequal variance or correlations, or when the vector Y follows a bivariate Student t-distribution.[3]

Properties

The raw moments are given by:

$$\mu_j = \sigma^j 2^{j/2}\, \Gamma\left(1 + \frac{j}{2}\right),$$

where $\Gamma(z)$ is the gamma function.

The mean of a Rayleigh random variable is thus :

$$\mu(X) = \sigma\sqrt{\frac{\pi}{2}} \ \approx 1.253\ \sigma.$$

The variance of a Rayleigh random variable is :

$$\operatorname{var}(X) = \mu_2 - \mu_1^2 = \left(2 - \frac{\pi}{2}\right)\sigma^2 \approx 0.429\ \sigma^2$$

The mode is σ, and the maximum pdf is

$$f_{\max} = f(\sigma;\sigma) = \frac{1}{\sigma}e^{-1/2} \approx \frac{0.606}{\sigma}.$$

The skewness is given by:

$$\gamma_1 = \frac{2\sqrt{\pi}(\pi-3)}{(4-\pi)^{3/2}} \approx 0.631$$

The excess kurtosis is given by:

$$\gamma_2 = -\frac{6\pi^2 - 24\pi + 16}{(4-\pi)^2} \approx 0.245$$

The characteristic function is given by:

$$\varphi(t) = 1 - \sigma t e^{-\frac{1}{2}\sigma^2 t^2}\sqrt{\frac{\pi}{2}}\left[\operatorname{erfi}\left(\frac{\sigma t}{\sqrt{2}}\right) - i\right]$$

where $\operatorname{erfi}(z)$ is the imaginary error function. The moment generating function is given by

$$M(t) = 1 + \sigma t\, e^{\frac{1}{2}\sigma^2 t^2}\sqrt{\frac{\pi}{2}}\left[\operatorname{erf}\left(\frac{\sigma t}{\sqrt{2}}\right) + 1\right]$$

where $\operatorname{erf}(z)$ is the error function.

Differential entropy

The differential entropy is given by

$$H = 1 + \ln\left(\frac{\sigma}{\sqrt{2}}\right) + \frac{\gamma}{2}$$

where γ is the Euler–Mascheroni constant.

Parameter estimation

Given a sample of N independent and identically distributed Rayleigh random variables x_i with parameter σ,

$\widehat{\sigma}^2 \approx \frac{1}{2N}\sum_{i=1}^N x_i^2$ is the maximum likelihood estimate and also is unbiased.

$\widehat{\sigma} \approx \sqrt{\frac{1}{2N}\sum_{i=1}^N x_i^2}$ is a biased estimator that can be corrected via the formula

$$\sigma = \widehat{\sigma}\frac{\Gamma(N)\sqrt{N}}{\Gamma(N+\frac{1}{2})} = \widehat{\sigma}\frac{4^N N!(N-1)!\sqrt{N}}{(2N)!\sqrt{\pi}}$$ [4]

Confidence intervals

To find the (1 − α) confidence interval, first find the bounds $[a, b]$ where:

$$P(\chi^2_{2N} \leq a) = \alpha/2, \quad P(\chi^2_{2N} \leq b) = 1 - \alpha/2$$

then the scale parameter will fall within the bounds

$$\frac{N\overline{x^2}}{b} \leq \widehat{\sigma}^2 \leq \frac{N\overline{x^2}}{a}$$ [5]

Generating random variates

Given a random variate U drawn from the uniform distribution in the interval (0, 1), then the variate

$$X = \sigma\sqrt{-2\ln U}$$

has a Rayleigh distribution with parameter σ. This is obtained by applying the inverse transform sampling-method.

Related distributions

- $R \sim \mathrm{Rayleigh}(\sigma)$ is Rayleigh distributed if $R = \sqrt{X^2 + Y^2}$, where $X \sim N(0, \sigma^2)$ and $Y \sim N(0, \sigma^2)$ are independent normal random variables.[6] (This gives motivation to the use of the symbol "sigma" in the above parametrization of the Rayleigh density.)
- The magnitude $|z|$ of a standard complex normally distributed variable *z* will have the Rayleigh distribution.
- The chi distribution with *v* = 2 is equivalent to the Rayleigh Distribution with *σ* = 1.
- If $R \sim \mathrm{Rayleigh}(1)$, then R^2 has a chi-squared distribution with parameter N, degrees of freedom, equal to two (*N* = 2)

  $$[Q = R^2] \sim \chi^2(N)\,.$$

- If $R \sim \mathrm{Rayleigh}(\sigma)$, then $\sum_{i=1}^N R_i^2$ has a gamma distribution with parameters N and $2\sigma^2$

$$\left[Y = \sum_{i=1}^N R_i^2\right] \sim \Gamma(N, 2\sigma^2).$$

- The Rice distribution is a noncentral generalization of the Rayleigh distribution: $\mathrm{Rayleigh}(\sigma) = \mathrm{Rice}(0, \sigma)$.
- The Weibull distribution with the "shape parameter" *k*=2 yields a Rayleigh distribution. Then the Rayleigh distribution parameter σ is related to the Weibull scale parameter according to $\lambda = \sigma\sqrt{2}.$
- The Maxwell–Boltzmann distribution describes the magnitude of a normal vector in three dimensions.
- If X has an exponential distribution $X \sim \mathrm{Exponential}(\lambda)$, then $Y = \sqrt{X} \sim \mathrm{Rayleigh}(1/\sqrt{2\lambda})$.
- The half-normal distribution is the univariate special case of the Rayleigh distribution.

Applications

An application of the estimation of σ can be found in magnetic resonance imaging (MRI). As MRI images are recorded as complex images but most often viewed as magnitude images, the background data is Rayleigh distributed. Hence, the above formula can be used to estimate the noise variance in an MRI image from background data.[7] [8]

The Rayleigh distribution was also employed in the field of nutrition for linking dietary nutrient levels and human and animal responses. In this way, the parameter σ may be used to calculate nutrient response relationship.[9]

See also

- Rayleigh fading
- Rayleigh mixture distribution
- Circular error probable

References

1. "The Wave Theory of Light", *Encyclopedic Britannica* 1888; "The Problem of the Random Walk", *Nature* 1905 vol.72 p.318
2. Papoulis, Athanasios; Pillai, S. (2001) *Probability, Random Variables and Stochastic Processes*. ISBN 0073660116, ISBN 9780073660110
3. Röver, C. (2011). "Student-t based filter for robust signal detection". *Physical Review D*. **84** (12): 122004. arXiv:1109.0442 (https://arxiv.org/abs/1109.0442). Bibcode:2011PhRvD..84l2004R (https://ui.adsabs.harvard.edu/abs/2011PhRvD..84l2004R). doi:10.1103/physrevd.84.122004 (https://doi.org/10.1103%2Fphysrevd.84.122004).
4. Siddiqui, M. M. (1964) "Statistical inference for Rayleigh distributions", *The Journal of Research of the National Bureau of Standards, Sec. D: Radio Science*, Vol. 68D, No. 9, p. 1007 (https://archive.org/details/jresv68Dn9p1005)
5. Siddiqui, M. M. (1961) "Some Problems Connected With Rayleigh Distributions", *The Journal of Research of the National Bureau of Standards; Sec. D: Radio Propagation*, Vol. 66D, No. 2, p. 169 (http://nvlpubs.nist.gov/nistpubs/jres/66D/jresv66Dn2p167_A1b.pdf)

6. Hogema, Jeroen (2005) "Shot group statistics" (https://web.archive.org/web/20131105232146/http://home.kpn.nl/jhhogema1966/skeetn/ballist/sgs/sgs.htm#_Toc96439743)
7. Sijbers, J.; den Dekker, A. J.; Raman, E.; Van Dyck, D. (1999). "Parameter estimation from magnitude MR images". *International Journal of Imaging Systems and Technology*. **10** (2): 109–114. CiteSeerX 10.1.1.18.1228 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.18.1228). doi:10.1002/(sici)1098-1098(1999)10:2<109::aid-ima2>3.0.co;2-r (https://doi.org/10.1002%2F%28sici%291098-1098%281999%2910%3A2%3C109%3A%3Aaid-ima2%3E3.0.co%3B2-r).
8. den Dekker, A. J.; Sijbers, J. (2014). "Data distributions in magnetic resonance images: a review". *Physica Medica*. **30** (7): 725–741. doi:10.1016/j.ejmp.2014.05.002 (https://doi.org/10.1016%2Fj.ejmp.2014.05.002). PMID 25059432 (https://pubmed.ncbi.nlm.nih.gov/25059432).
9. Ahmadi, Hamed (2017-11-21). "A mathematical function for the description of nutrient-response curve" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5697816). *PLOS ONE*. **12** (11): e0187292. Bibcode:2017PLoSO..1287292A (https://ui.adsabs.harvard.edu/abs/2017PLoSO..1287292A). doi:10.1371/journal.pone.0187292 (https://doi.org/10.1371%2Fjournal.pone.0187292). ISSN 1932-6203 (https://www.worldcat.org/issn/1932-6203). PMC 5697816 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5697816). PMID 29161271 (https://pubmed.ncbi.nlm.nih.gov/29161271).