Elements of Probability Theory (Part II)

Prof. Americo Cunha Jr

Rio de Janeiro State University – UERJ

americo.cunha@uerj.br

www.americocunha.org

@AmericoCunhaJr @AmericoCunhaJr

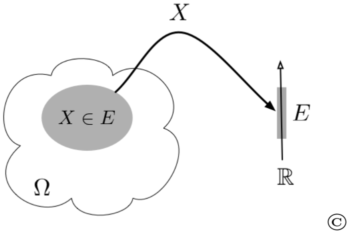
@AmericoCunhaJr @AmericoCunhaJr

More on Probability in Dimension 1

Random variable

A mapping $X : \Omega \to \mathbb{R}$ is called a random variable (RV) if the preimage of every real number under X is a relevant event, i.e.,

$$X^{-1}(x) = \left\{ \omega \in \Omega : \; X(\omega) \leq x \right\} \in \Sigma, \quad \text{for every } x \in \mathbb{R}.$$



A collection of events in Ω is mapped to an interval E on the real line under such mapping.

RV are numerical characteristics of interesting events.

Remark:
A random variable is a function from Ω to $\mathbb{R}$, not a real number.

Examples of random variables

1. Rolling die experiment

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

$$\Sigma = \left\{\phi, \{1, 3, 5\}, \{2, 4, 6\}, \Omega\right\}$$

- $X_1(\omega) = \begin{cases} 1 & \text{if } \omega \text{ is even} \\ 0 & \text{if } \omega \text{ is odd} \end{cases}$

  (random variable)

- $X_2(\omega) = \omega^2$

  (not a random variable)

2. Temperature (in Kelvin) measurement experiment

$$\Omega = [a, b] \subset [0, +\infty)$$

$$\Sigma = \mathcal{B}_{[a,b]}$$ (Borel σ-algebra)

- $X(\omega) = -459.67 + 1.8\,\omega$

  (random variable)

G. Grimmett and D. Welsh, **Probability: An Introduction**. Oxford University Press, 2 edition, 2014.

Probability distribution

The probability distribution of X, denoted by F_X, is defined as the probability of the elementary event $\{X \leq x\}$, i.e.,

$$F_X(x) = \mathcal{P}\{X \leq x\}.$$

F_X is also known as cumulative distribution function (CDF) and has the following properties:

- $0 \leq F_X(x) \leq 1$
- F_X is a monotonic, non decreasing, right continuous function
- $\mathcal{P}\{x_1 < X \leq x_2\} = F_X(x_2) - F_X(x_1) = \int_{x_1}^{x_2} dF_X(x)$
- $\int_{\mathbb{R}} dF_X(x) = 1$
- $F_X(-\infty) = 0$ and $F_X(+\infty) = 1$

Probability density function

If the function F_X is differentiable, its derivative $p_X(x) = dF_X(x)/dx$ is called probability density function (PDF) of X, and one has

$$F_X(x) = \int_{-\infty}^{x} p_X(\xi)\, d\xi.$$

Note also that:

- $p_X(x) \geq 0$ for every $x \in \mathbb{R}$
- $\displaystyle\int_{\mathbb{R}} p_X(x)\, dx = 1$

Remark:
Intuitively, $p_X(x)\, dx$ can be thought of as the probability of X falling within the infinitesimal interval $[x, x + dx]$.

Types of random variables

1. Discrete random variable
   Distribution is discrete.
   Assumes a denumerable number of values.
   Typically associated with counting processes.
2. Continuous random variable
   Distribution is continuous.
   Assumes a non-denumarable number of values.
   Typically associated with measuring processes.
3. Mixed random variable
   Distribution has points of discontinuity.
   Assumes a non-denumarable number of values.
   A "mixture"of the two previous types.
4. Singular random variable
   Distribution is not differentiable at any point.
   It has theoretical interest only.

Function of random variable

The function of random variable $Y = h(X)$, for a random variable X and measurable mapping $h : \mathbb{R} \to \mathbb{R}$, is also a random variable, which has its own probability distribution.

Example:

Let $h(x) = x^2$ and $X : \Omega \to [1, 2]$ be a random variable such that

$$F_X(x) = \begin{cases} 0 & \text{if } \; x < 1 \\ 1/2\, x & \text{if } \; 1 \leq x \leq 2 \\ 1 & \text{if } \; x > 2. \end{cases}$$

The composition $Y = X^2$ is a random variable with distribution

$$F_Y(y) = \begin{cases} 0 & \text{if } \; y < 1 \\ 1/2\, \sqrt{y} & \text{if } \; 1 \leq y \leq 4 \\ 1 & \text{if } \; y > 4. \end{cases}$$

Mathematical expectation operator

The mathematical expectation of a random variable X is defined as

$$\mathbb{E}\left\{X\right\} = \int_{\mathbb{R}} x \, dF_X(x).$$

The mathematical expectation is a linear operator since

$$\mathbb{E}\left\{\alpha_1 \, X_1 + \alpha_2 \, X_2\right\} = \alpha_1 \, \mathbb{E}\left\{X_1\right\} + \alpha_2 \, \mathbb{E}\left\{X_2\right\},$$

for any pairs of number α_1, α_2 and random variables X_1, X_2.

Theorem (law of the unconscious statistician):

Given a measurable mapping $h : \mathbb{R} \to \mathbb{R}$ and a random variable X the expected value of $h\,(X)$ is given by

$$\mathbb{E}\left\{h\,(X)\right\} = \int_{\mathbb{R}} h\,(x) \; dF_X(x).$$

Statistical moments

The value $\mathbb{E}\left\{(X-b)^k\right\}$ is called k-th moment around b of the random variable X, for $b \in \mathbb{R}$ and $k = 1, 2, \cdots$.

For $b = \mathbb{E}\{X\}$, $\mathbb{E}\left\{\left(X - \mathbb{E}\{X\}\right)^k\right\}$ is dubbed k-th central moment of the random variable X.

When $b = 0$, $\mathbb{E}\left\{X^k\right\}$ is simply called k-th moment of the random variable X.

Mean value

The mean value of the random variable X is defined as

$$\begin{aligned} \mu_X &= \mathbb{E}\left\{X\right\} \\ &= \int_{\mathbb{R}} x \, dF_X(x) \\ &= \int_{\mathbb{R}} x \, p_X(x) \, dx. \end{aligned}$$

(measure of the central tendency)

Remark:
The mean value μ_X is the constant which best approximate the random variable X. The error of this approximation is the standard deviation σ_X.

Variance

The variance of the random variable X is defined as

$$\begin{aligned}\sigma_X^2 &= \mathbb{E}\left\{(X-\mu_X)^2\right\}\\ &= \int_{\mathbb{R}} (x-\mu_X)^2 \, dF_X(x)\\ &= \int_{\mathbb{R}} (x-\mu_X)^2 \, p_X(x)\, dx.\end{aligned}$$

(measure of dispersion about the mean)

Note that variance can also be written as

$$\sigma_X^2 = \mathbb{E}\left\{X^2\right\} - \left(\mathbb{E}\left\{X\right\}\right)^2 .$$

Remark:
σ_X^2 has the same unit as X^2.

Standard deviation and variation coefficient

Other second-order statistics of X are the standard deviation

$$\sigma_X = \sqrt{\sigma_X^2},$$

and the variation coefficient

$$\delta_X = \sigma_X / \mu_X, \qquad \mu_X \neq 0.$$

(both are measures of dispersion about the mean)

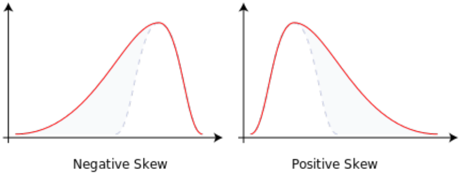
Remark:
σ_X has the same unit as X and δ_X is dimensionless.

Skewness

The skewness of the random variable X is defined as

$$\texttt{Skew}\,[X] = \mathbb{E}\left\{ \left(\frac{X - \mu_X}{\sigma_X} \right)^3 \right\}$$

(measure of asymmetry about the mean)



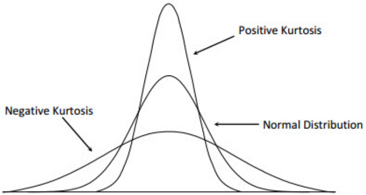
*Picture from https://en.wikipedia.org/wiki/Skewness

Kurtosis

The kurtosis of the random variable X is defined as

$$\texttt{Kurt}\left[X\right] = \mathbb{E}\left\{\left(\frac{X-\mu_X}{\sigma_X}\right)^4\right\}.$$

(measure of "tailedness")



Entropy

The Entropy of p_X is defined as

$$\mathcal{S}\left(p_X\right) = -\mathbb{E}\left\{\ln\left(p_X(X)\right)\right\}$$

which is equivalent to

$$\mathcal{S}\left(p_X\right) = -\int_{\mathbb{R}} p_X(x) \ln\left(p_X(x)\right) dx.$$

It provides a measure for the level of uncertainty of p_X.

Remark:

For discrete random variables $\mathcal{S}\left(P_X\right) = -\sum_k P_k \ln\left(P_k\right)$.

Second-order random variables

The mapping X is a second-order random variable if the expectation of its square (second-order moment) is finite, i.e.,

$$\mathbb{E}\left\{X^2\right\} < +\infty.$$

In consequence,

$$\mathbb{E}\left\{X\right\} < +\infty,$$

and hence

$$\sigma_X^2 = \mathbb{E}\left\{X^2\right\} - \left(\mathbb{E}\left\{X\right\}\right)^2 < +\infty.$$

(mean and variance are also finite)

Probability Distributions

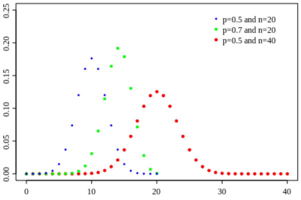
Binomial

- Notation: $B(n, p)$
- Support: $\{0, 1, 2, \cdots, n\}$
- Parameters:
  - $n \in \mathbb{N}$ — number of trials
  - $p \in [0, 1]$ — success probability
- PMF:

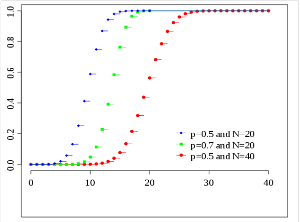
$$P_k = \binom{n}{k} \, p^k \, (1-p)^{n-k}$$

- Statistics:
  - $\mu = n\,p$
  - $\sigma^2 = n\,p\,(1-p)$

Probability Mass Function



Cumulative Distribution Function



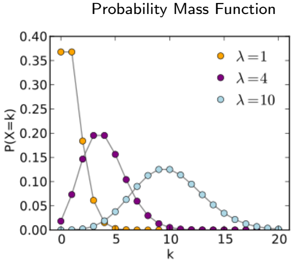
Binomial distribution — Wikipedia, The Free Encyclopedia, 2021.

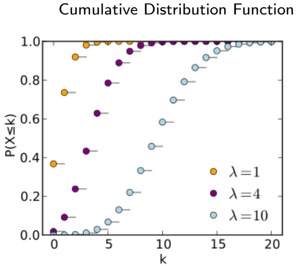
Poisson

- Notation: $Poisson(\lambda)$
- Support: $\{0, 1, 2, 3, \cdots\}$
- Parameter:
  - $\lambda > 0$ — mean
- PMF:

$$P_k = \frac{\lambda^k \, e^{-\lambda}}{k!}$$

- Statistics:
  - $\mu = \lambda$
  - $\sigma^2 = \lambda$





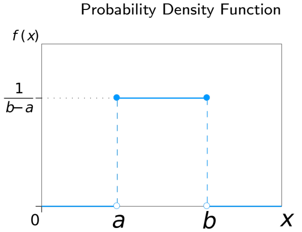
Poison distribution — Wikipedia, The Free Encyclopedia, 2021.

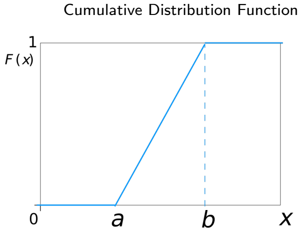
Uniform

- Notation: $\mathcal{U}(a, b)$
- Support: $[a, b]$
- Parameters:
  - $-\infty < a < b < +\infty$ — boundaries
- PDF:

$$p_X(x) = \frac{1}{b-a}\,\mathbb{1}_{[a,b]}(x)$$

- Statistics:
  - $\mu = \frac{1}{2}\,(a+b)$
  - $\sigma^2 = \frac{1}{12}\,(b-a)^2$





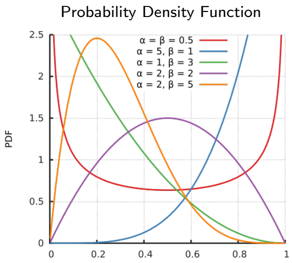
Uniform distribution — Wikipedia, The Free Encyclopedia, 2021.

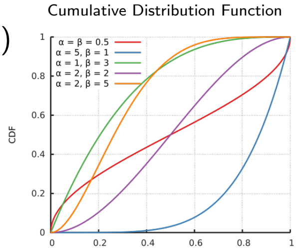
Beta

- Notation: $Beta(\alpha, \beta)$
- Support: $[0, 1]$
- Parameters:
  - α — shape parameter
  - β — shape parameter
- PDF:

$$p_X(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\, x^{\alpha-1}\,(1-x)^{\beta-1}\,\mathbb{1}_{[0,1]}(x)$$

- Statistics:
  - $\mu = \dfrac{\alpha}{(\alpha+\beta)}$
  - $\sigma^2 = \dfrac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$





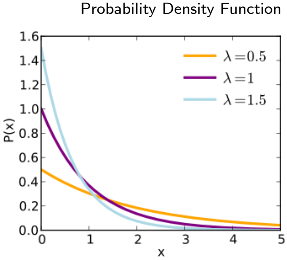
Beta distribution — Wikipedia, The Free Encyclopedia, 2021.

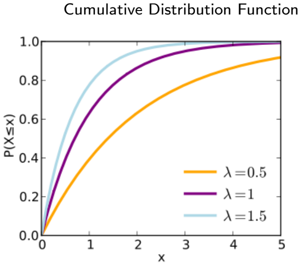
Exponential

- Notation: $Exp(\lambda)$
- Support: $[0, +\infty)$
- Parameter:
  - λ — rate parameter
- PDF:

$$p_X(x) = \lambda\, e^{(-x/\lambda)}\, \mathbb{1}_{[0,+\infty)}(x)$$

- Statistics:
  - $\mu = \lambda^{-1}$
  - $\sigma^2 = \lambda^{-2}$





Exponential distribution — Wikipedia, The Free Encyclopedia, 2021.

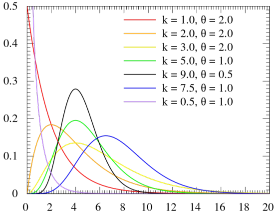
Gamma

- Notation: $Gamma(k, \theta)$
- Support: $(0, +\infty)$
- Parameters:
  - k — shape parameter
  - θ — scale parameter
- PDF:

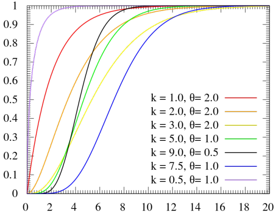
$$p_X(x) = \frac{1}{\Gamma(k)\,\theta^k}\, x^{k-1}\, e^{\left(-x/\theta\right)}\, \mathbb{1}_{(0,+\infty)}(x)$$

- Statistics:
  - $\mu = k\,\theta$
  - $\sigma^2 = k\,\theta^2$

Probability Density Function



Cumulative Distribution Function



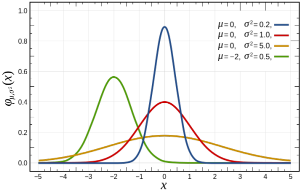
Gamma distribution — Wikipedia, The Free Encyclopedia, 2021.

Gaussian

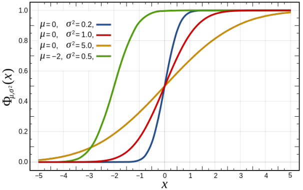
- Notation: $\mathcal{N}(\mu, \sigma^2)$
- Support: $(-\infty, +\infty)$
- Parameters:
    - μ — mean
    - σ^2 — variance
- PDF:

$$p_X(x) = \frac{1}{\sqrt{2\,\pi\,\sigma^2}} \exp\left\{ -\frac{(x-\mu)^2}{2\,\sigma^2} \right\}$$

Probability Density Function



Cumulative Distribution Function



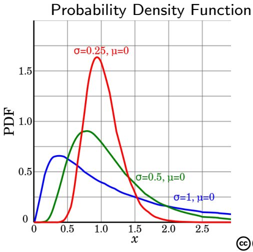
Normal distribution — Wikipedia, The Free Encyclopedia, 2021.

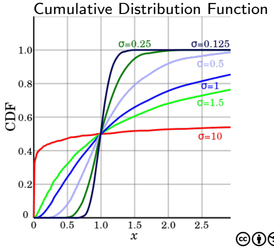
Log-normal

- Notation: $\ln \mathcal{N}(m, s^2)$
- Support: $(0, +\infty)$
- Parameters:
  - $m \in \mathbb{R}$ — location parameter
  - $s > 0$ — scale parameter
- PDF:

$$p_X(x) = \frac{1}{x \sqrt{2\,\pi\, s^2}} \exp\left\{ -\frac{(\ln x - m)^2}{2\, s^2} \right\}$$

- Statistics:
  - $\mu = e^{m + s^2/2}$
  - $\sigma^2 = (e^{s^2} - 1)\, e^{2\,m + s^2}$





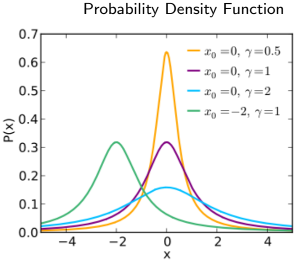
Log-normal distribution — Wikipedia, The Free Encyclopedia, 2021.

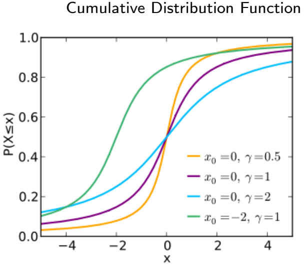
Cauchy

- Notation: $\mathit{Cauchy}(x_0, \gamma)$
- Support: $(-\infty, +\infty)$
- Parameters:
  - x_0 — location parameter
  - γ — scale parameter
- PDF:

$$p_X(x) = \frac{1}{\pi\,\gamma} \left[\frac{\gamma^2}{(x - x_0)^2 + \gamma^2} \right]$$

- Statistics:
  - $\mu =$ undefined
  - $\sigma^2 =$ undefined (not 2nd-order RV)





Cauchy distribution — Wikipedia, The Free Encyclopedia, 2021.

Characterization of a probability distribution

Given the mean μ and standard deviation σ of a random variable.
Is the distribution well-defined?

Characterization of a probability distribution

Given the mean μ and standard deviation σ of a random variable.
Is the distribution well-defined?

No! Multiple distributions may have the same statistics.

Characterization of a probability distribution

Given the mean μ and standard deviation σ of a random variable.
Is the distribution well-defined?

No! Multiple distributions may have the same statistics.

$$p_X(x) = \frac{1}{\sqrt{6\,\pi}} \exp\left\{-\frac{(x-1)^2}{6}\right\} \quad \text{and} \quad p_X(x) = \frac{1}{6}\,\mathbb{1}_{[-2,4]}(x)$$

have $\mu = 1$ and $\sigma = \sqrt{3}$.

Characterization of a probability distribution

Given the mean μ and standard deviation σ of a random variable.
Is the distribution well-defined?

No! Multiple distributions may have the same statistics.

$$p_X(x) = \frac{1}{\sqrt{6\,\pi}} \exp\left\{-\frac{(x-1)^2}{6}\right\} \quad \text{and} \quad p_X(x) = \frac{1}{6}\,\mathbb{1}_{[-2,4]}(x)$$

have $\mu = 1$ and $\sigma = \sqrt{3}$.

What type of information determines a distribution?

Characterization of a probability distribution

Given the mean μ and standard deviation σ of a random variable. Is the distribution well-defined?

No! Multiple distributions may have the same statistics.

$$p_X(x) = \frac{1}{\sqrt{6\,\pi}} \exp\left\{-\frac{(x-1)^2}{6}\right\} \quad \text{and} \quad p_X(x) = \frac{1}{6}\,\mathbb{1}_{[-2,4]}(x)$$

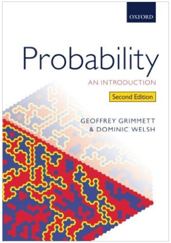
have $\mu = 1$ and $\sigma = \sqrt{3}$.

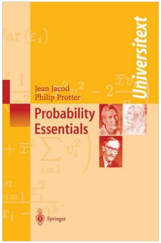
What type of information determines a distribution?

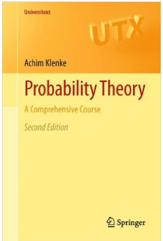
- cumulative distribution function
- probability density function (if exists)
- quantile function
- characteristic function
- moment-generating function (if exists)

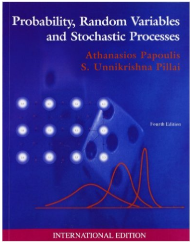
References

- G. Grimmett and D. Welsh, **Probability: An Introduction**. Oxford University Press, 2 edition, 2014.
- J. Jacod and P. Protter, **Probability Essentials**. Springer, 2nd edition, 2004.
- A. Klenke, **Probability Theory: A Comprehensive Course**. Springer, 2nd edition, 2014.
- A. Papoulis and S. U. Pillai, **Probability, Random Variables and Stochastic Processes**. McGraw-Hill Europe; 4th edition, 2002.









How to cite this material?

A. Cunha Jr, *Elements of Probability Theory (Part II)*, Rio de Janeiro State University – UERJ, 2021.

@AmericoCunhaJr @AmericoCunhaJr

@AmericoCunhaJr @AmericoCunhaJr

© copyrighted material (1/3)

Content excluded from our Creative Commons license

- Random variable:
  http://theanalysisofdata.com/probability/2_1.html
- Skewness:
  Wikimedia Commons, File:Negative and positive skew diagrams (English).svg — Wikimedia Commons, the free media repository https://commons.wikimedia.org/w/index.php?title=File:Negative_and_positive_skew_diagrams_(English).svg
- Kurtosis:
  E. Patitsas et al., Evidence That Computer Science Grades Are Not Bimodal, ICER '16 September 08-12, 2016, Melbourne, VIC, Australia, http://dx.doi.org/10.1145/2960310.2960312
- Poison distribution PMF:
  Wikimedia Commons, File:Poisson pmf.svg — Wikimedia Commons, the free media repository https://commons.wikimedia.org/w/index.php?title=File:Poisson_pmf.svg
- Poison distribution CDF:
  Wikimedia Commons, File:Poisson cdf.svg — Wikimedia Commons, the free media repository https://commons.wikimedia.org/w/index.php?title=File:Poisson_cdf.svg

© copyrighted material (2/3)

Content excluded from our Creative Commons license

- Uniform distribution PDF:
  Wikimedia Commons, File:Uniform Distribution PDF SVG.svg — Wikimedia Commons, the free media repository
  https://commons.wikimedia.org/w/index.php?title=File:Uniform_Distribution_PDF_SVG.svg
- Uniform distribution CDF:
  Wikimedia Commons, File:Uniform cdf.svg — Wikimedia Commons, the free media repository
  https://commons.wikimedia.org/w/index.php?title=File:Uniform_cdf.svg
- Exponential distribution PDF:
  Wikimedia Commons, File:Exponential probability density.svg — Wikimedia Commons, the free media repository
  https://commons.wikimedia.org/w/index.php?title=File:Exponential_probability_density.svg
- Exponential distribution CDF:
  Wikimedia Commons, File:Exponential cdf.svg — Wikimedia Commons, the free media repository
  https://commons.wikimedia.org/w/index.php?title=File:Exponential_cdf.svg
- Gamma distribution PDF:
  Wikimedia Commons, File:Gamma distribution pdf.svg — Wikimedia Commons, the free media repository
  https://commons.wikimedia.org/w/index.php?title=File:Gamma_distribution_pdf.svg
- Gamma distribution CDF:
  Wikimedia Commons, File:Gamma distribution cdf.svg — Wikimedia Commons, the free media repository
  https://commons.wikimedia.org/w/index.php?title=File:Gamma_distribution_cdf.svg

© copyrighted material (3/3)

Content excluded from our Creative Commons license

- Log-normal distribution PDF:
  Wikimedia Commons, File:PDF-log normal distributions.svg — Wikimedia Commons, the free media repository
  https://commons.wikimedia.org/w/index.php?title=File:PDF-log_normal_distributions.svg
- Log-normal distribution CDF:
  Wikimedia Commons, File:CDF-log normal distributions.svg — Wikimedia Commons, the free media repository
  https://commons.wikimedia.org/w/index.php?title=File:CDF-log_normal_distributions.svg
- Cauchy distribution PDF:
  Wikimedia Commons, File:Cauchy pdf.svg — Wikimedia Commons, the free media repository
  https://commons.wikimedia.org/w/index.php?title=File:Cauchy_pdf.svg
- Cauchy distribution CDF:
  Wikimedia Commons, File:Cauchy cdf.svg — Wikimedia Commons, the free media repository
  https://commons.wikimedia.org/w/index.php?title=File:Cauchy_cdf.svg