NMH'S INCOMPLETE

DICTIONARY OF STATISTICS

NMH's
INCOMPLETE DICTIONARY OF STATISTICS

Nils M Holm

NMH’s Incomplete Dictionary of Statistics

Nils M Holm, 2017

Print and Distribution:
Lulu Press, Inc.; Raleigh, NC; USA

To

Hanne
Günter († 2016)
Franz († 1992)
Andrea
Benjamin
Nana / Nick
Célina

Thanks!

Preface

This book is incomplete for two reasons: there is obviously so much more to statistics than covered in this tiny volume, so it is a work in slow progress. It will probably be extended whenever I learn something new and interesting. Then it focuses on the most obvious and simple applications of each item it describes. Most more complete works and even Wikipedia these days attempt to cover any subject in a broad *and* in-depth fashion but, while focusing on correctness and completeness, they often drown the reader in formulae and proofs without ever saying what a concept is about. This volume attempts to give the reader an intuition about the use and meaning of each concept explained.

This book is primarily a note to myself. As I move on from one topic to another, I often forget things, simple because I never use them again. Then an occasion may come up where some of the past knowledge may come in handy, and at this point I would like to avoid the frustrating process of re-gaining the knowledge—and learning mathematics, in particular, is extraordinarily frustrating due terrible textbooks and highly inconsistent terminology and use of symbols. So I have written this book in order to save myself some future pain.

As usual, this book has been typeset with Troff, Pic, Eqn, and Tbl. All diagrams in the book have been made with the Klong statistics library and plotter. You can find the Klong language and libraries on my homepage at **`t3x.org`**.

Nils M Holm, 2017

A

alpha level (α)
→ *level of significance*

alternative hypothesis (H_A, H_1)
A hypothesis to be tested in a *hypothesis test*. It can be accepted if, and only if, the corresponding *null hypothesis* H_0 must be rejected. For example, if H_A states that a die is loaded, then H_0 would state that the die is fair. The two hypotheses are mutually exclusive.

H_A is called the "alternative hypothesis", because the null hypothesis H_0 is often based on some ideal *model* that would deliver a perfect explanation for a phenomenon. If H_0 must be rejected, H_A provides an alternative explanation.

anticorrelation
The negative *correlation* of two *random variables*. A form of correlation where the value of one variable grows larger as the value of the other variables grows smaller. See figure ACO.

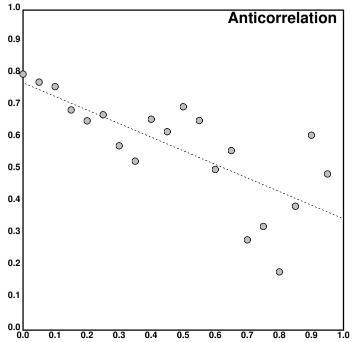


Figure **ACO:** anticorrelation, correlation coefficient: $r = -0.42$

Formally, two variables are anticorrelated, when their *correlation coefficient* is negative.

average
Mostly the *mean*, but see also *median*, *mode*.

B

Bayes theorem (Bayes rule)
→ *reverse conditional probability*

bell curve (Gauss curve)
→ *normal distribution*

Bernoulli trial
A *trial* with exactly two *outcomes* which are typically called "success" and "failure". The *probability* of success (p) and failure ($q = 1 - p$) is constant in an experiment consisting of a series of Bernoulli trials.

The term "success" here describes the *observation* of an *event* and "failure" describes the absence of such an observation.

Bessel's correction
A method to remove *bias* from the *sample variance* s^2 (and *standard deviation* s) by multiplying it by a factor of $\frac{n}{n-1}$. Bessel's correction then results in the following formula for the sample variance:

$$s^2 = \frac{n}{n-1} \cdot \frac{1}{n} \sum_{i=1}^{n} (x - \bar{x})^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x - \bar{x})^2$$

bias ($bias(\hat{\theta})$)
The difference between an *estimator's* expected value (→ *expectation*) and the true value of the *statistic* being estimated. An estimator is called "biased", if its bias is non-zero, otherwise it is called "unbiased". The bias of an estimator $\hat{\theta}$, which estimates the value of a statistic θ, is defined as

$$bias(\hat{\theta}) = E[\hat{\theta} - \theta] = E[\hat{\theta}] - \theta$$

An estimator is called "negatively biased", if $bias(\hat{\theta}) < 0$ and "positively biased", if $bias(\hat{\theta}) > 0$. See also *bias-variance trade-off*.

bias-variance trade-off
In a statistical *model*, the *reducible error* is the sum of the *bias* and *variance* of the model (see *mean squared error*). Hence reducing the bias will increase the variance and vice versa. The trade-off

between bias and variance has to be considered on an individual basis, depending on the application of the model.

In non-linear models, reducing the variance to zero will result in overfitting, i.e. the model will measure the *irreducible error* (or noise) that is inherent in the *observation*. Overfitting causes the bias to grow, because the *estimates* of the model fluctuate heavily between the known *data points* used to create the model.

Using a model that is too simple (underfitting), on the other hand, will create a large variance and a minimal bias. See figure BVT for a visualization.

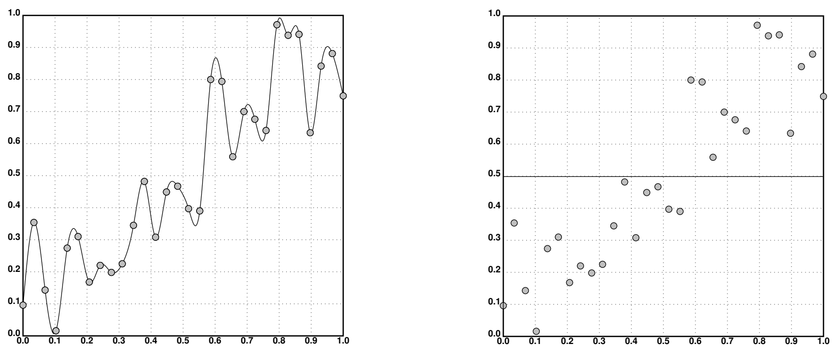


Figure **BVT:** extremes of bias-variance trade-off; left: no variance, all data points are matched perfectly, but huge bias due to modeling the irreducible error; right: no bias (negative and positive differences cancel out each other), but huge variance due to omitting a correlation

bimodal (adj)
A *data set* that has two *modes*.

binomial coefficient (${}_nC_k$)
Pronounced "n choose k". The number of unique k*-combinations* (configurations of k elements) that can be created from a set of n items without *replacement*. Alternative notation: $\binom{n}{k}$

The binomial coefficient is computed as follows:

$${}_nC_k = \frac{n!}{(n-k)! \cdot k!} = \frac{(n-k+1) \cdot \cdots \cdot (n-1) \cdot n}{k!}$$

In a combination, the order of elements does not matter. When ordering plays a role, configurations are called *permutations*. The number of combinations can be expressed using the number of k-

permutations as follows:

$$_nC_k = \frac{_nP_k}{k!}$$

The series $_nC_i$ for $i = 0, \cdots, n$ is exactly equal to the series $2^n f_B(i)$, where f_B is the *probability mass function* of the *binomial distribution* $X \sim B(n, \frac{1}{2})$.

binomial distribution $(X \sim B(n, p))$
A *discrete probability distribution* modeling the *probability* of a specific number of *successes* given a fixed number of *trials* and a fixed probability of success. A variable $X \sim B(n, p)$ is said to be binomially distributed with n trials and a probability of success of p. See figure BDD for details and figure BDC for function plots.

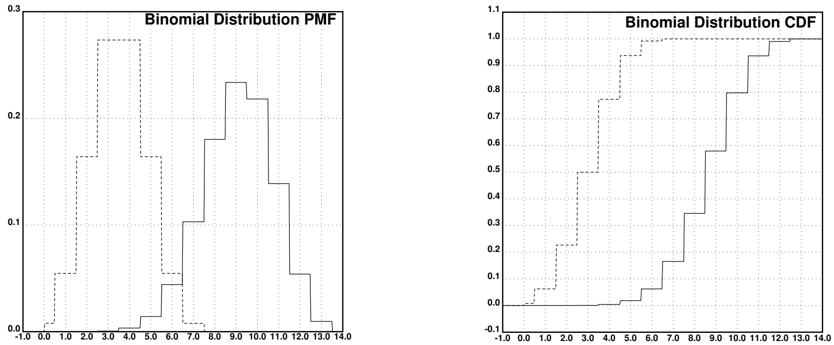


Figure **BDC:** binomial distr. probability functions; left: PMF of $X \sim B(13; 0.7)$ (solid) and $X \sim B(7; 0.5)$ (dashed); right: CDF of same distributions

For example, the probability of x out of 7 children being girls follows the probability distribution $X \sim B(7, 0.5)$, assuming that the probability for a child being a girl is $P(Girl) = 0.5$.

Given this distribution, the probability of three out of seven children being girls would be

$$P(X = 3) = f_B(3) = \binom{7}{3} \cdot 0.5^3 \cdot 0.5^{7-3} \approx 0.273$$

where $\binom{n}{x}$ is the *binomial coefficient*.

The probability $P(X \leq 3)$ of up to three out of seven children being girls would be:

$$F_B(3) = \sum_{w=0}^{3} f_B(w) = 0.5$$

$X \sim B(n, p)$	
PMF	$f(x) = \binom{n}{x} \cdot p^x \cdot q^{n-x}$
CDF	$F(x) = \sum_{i=0}^{x} \binom{n}{i} \cdot p^i \cdot q^{n-i}$
	$F(x) = I_q(n-x, 1+x)$
Statistic	$x \in \mathbf{N}_0, x \leq n$: number of successes
Parameters	$n \in \mathbf{N}_0$: number of trials
	$p \in [0, 1]$: probability of success
	q: $1 - p$ (probability of failure)
μ	np
σ^2	npq
Skewness (γ_1)	$\frac{q-p}{\sqrt{npq}}$
Approximations	$N(np, npq)$ for $np > 5, nq > 5$
	$Poi(np)$ for $n \geq 50$, $p < 0.1$

Figure **BDD:** binomial distribution; $I_x(a, b)$ is the regularized incomplete B (beta) function, a common method for computing the CDF of the distribution

bivariate (adj)
→ *multivariate*

C

category
One of a limited set of values that a *random variable* can assume. Categories are often non-numerical, e.g.: car brands, political parties, or favorite meals. Categorical data are typically modeled using *frequency distributions*, where a frequency is assigned to each category.

CDF
→ *cumulative distribution function*

central limit theorem (CLT)
A central (no pun intended) theorem of *statistics* that states that the *sampling distribution of the mean* $\bar{X}$ will be asymptotically normal (→ *normal distribution*), no matter what the distribution of the underlying *population* looks like. The necessary prerequisites for the CLT are that the population has a finite *mean* μ and a finite and non-zero *variance* σ^2, and the *random variables* of the sampling distribution are *independent and identically distributed*. The sampling distribution of the mean $\bar{X}$ will then be an asymptotically normal distribution with $\mu_{\bar{X}} = \mu$ and $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{k}$ (k being the number of samples).

More precisely, let $S_k = X_1 + X_2 + \cdots + X_k$ be the sum of k random variables (associated with k samples). Then $\frac{S_k}{k}$ is their sampling distributions of the mean. Assuming that this distribution is normal, it can be converted to a *standard normal distribution* as follows:

$$Z_k = \frac{S_k - k\mu}{\sqrt{k\sigma^2}} = \frac{S_k - k\mu}{\sigma\sqrt{k}}$$

As $\frac{S_k}{k}$ is assumed to be asymptotically normal, Z_k should be asymptotically equal to the standard normal distribution Z, i.e.:

$$\lim_{k\to\infty} P(Z_k \le x) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{x} e^{-\frac{w^2}{2}}\, dw$$

where the right-hand side of the equation is exactly the *cumulative distribution function* (CDF) of Z. And this is precisely what the CLT proves. See figure CLT for a visualization.

The CLT is important, because it shows that many distributions can be asymptotically normal under certain conditions, thus frequently allowing to use tools intended for the normal distribution instead of more complex, specialized tools of other distributions.

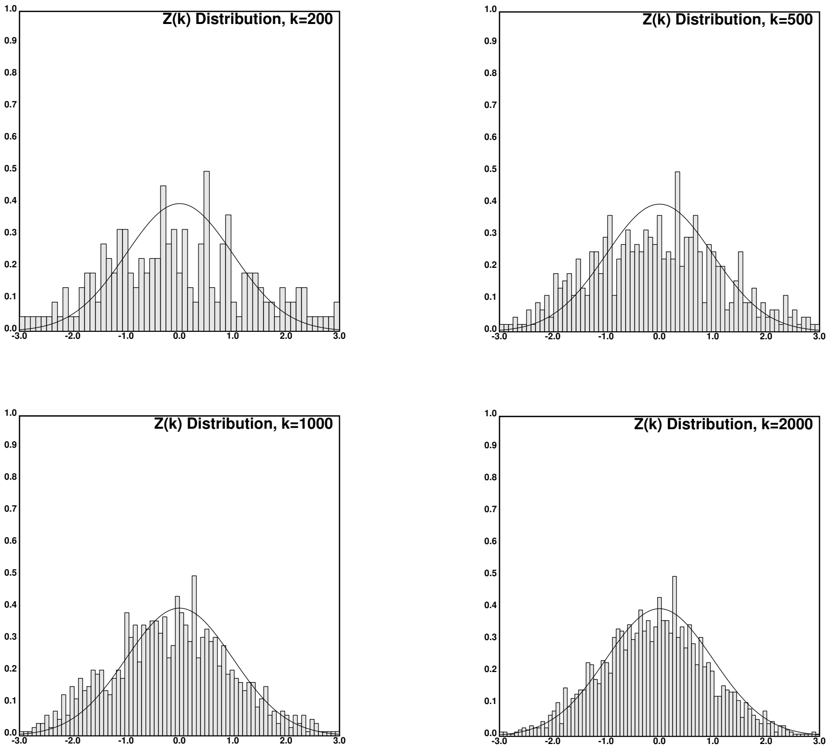


Figure **CLT:** central limit theorem: all four panels show the normalized sampling distribution of the mean, Z_k, with samples drawn from a uniform distribution; upper left panel: $k = 200$, upper right panel: $k = 500$, lower left panel: $k = 1000$, lower right panel: $k = 2000$; shape tends toward the Z-distribution

central tendency
→ *location*

chi-square distribution $(X \sim \chi^2(\nu))$
A *continuous probability distribution* that models the *improbability* of an *outcome* O given an expectation E, where both O and E may be multiple *events*. The χ^2-distribution maps a χ^2 statistic

(→ *chi-square statistic*) to the improbability of O arising by chance given the expectation E.

A *random variable* $X \sim \chi^2(\nu)$ is χ^2-distributed with ν (nu) *degrees of freedom*. Sometimes the notation $X \sim \chi^2_\nu$ is used instead. Often the symbol χ^2 is also used to denote a χ^2-distributed random variable or a χ^2 statistic, which can be confusing.

See figure XSD for a summary of the distribution and figure XSC for sample curves of the associated probability functions.

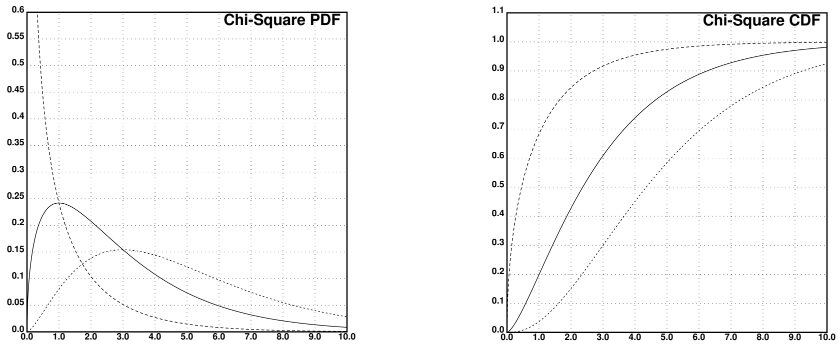


Figure **XSC:** χ^2-distribution probability functions; left: PDF of distributions with $\nu = 1$ degrees of freedom (dashed), $\nu = 3$ (solid), and $\nu = 5$ (dotted); right: CDF with the same parameters

The following example describes a test for goodness of fit. For a test for independence, see *contingency table*.

If 67 white marbles and 33 black marbles are drawn from a box given the expectation that there is a fair chance for drawing either color, the resulting χ^2 statistic would be

$$\chi^2 = \sum \frac{(67-50)^2}{50} = \frac{(67-50)^2}{50} = 5.78$$

Here the expectation would be $E(X) = 50$, because $67 + 33 = 100$ marbles are drawn in total. The statistic could also use the distance $33 - 50$ (white marbles minus expectation) instead, because the squared deviation would be the same.

The χ^2 test for goodness of fit used to find the improbability of the outcome has one degree of freedom here, because one marble count can be derived from the other. The *CDF* of the

χ^2-distribution with one degree of freedom gives

$$F_{\chi^2(1)}(5.78) \approx 0.984$$

So the probability of the outcome to arise by chance given that the expectation is true is $p \approx 1 - 0.984 = 0.016$ and therefore the outcome does not fit the expectation well.

When this result would occur in the *experiment* of a *hypothesis test*, it might be significant (→ *significance*).

$X \sim \chi^2(\nu)$	
PDF	$f(x) = \frac{1}{2^{\frac{\nu}{2}} \cdot \Gamma(\frac{\nu}{2})} \cdot x^{\frac{\nu}{2}-1} \cdot e^{-\frac{x}{2}}$
CDF	$F(x) = \frac{1}{2^{\frac{\nu}{2}} \cdot \Gamma(\frac{\nu}{2})} \cdot \int_0^x w^{\frac{\nu}{2}-1} \cdot e^{-\frac{w}{2}} dw$
	$F(x) = P(\frac{\nu}{2}, \frac{x}{2})$
Statistic	$x \in \mathbf{R}_0^+$: χ^2 statistic
Parameter	$\nu \in \mathbf{N}^+$: degrees of freedom
μ	ν
σ^2	2ν
Skewness (γ_1)	$\sqrt{\frac{8}{\nu}}$
Approximation	$\sqrt{2\chi^2} - \sqrt{2\nu - 1} \sim N(0, 1)$ for $\nu \geq 30$

Figure **XSD:** χ^2-distribution; $P(a, x)$ is the regularized incomplete Γ function, a common method for computing the CDF of the distribution

chi-square statistic (χ^2, X^2)
A *score* that equals the sum of normalized squared deviations between observed *outcomes* and expected outcomes (→ *observation*, *expectation*):

$$\chi^2 = \sum_{i=1}^{n} \frac{(E_i - O_i)^2}{E_i}$$

where each O_i is an outcome and E_i is the corresponding expectation.

The statistic can be used in a χ^2 test for goodness of fit or a χ^2

test for independence. See *chi-square distribution*, *contingency table*. The symbol χ^2 is used for both the distribution and the corresponding statistic. Sometimes the upper case X^2 is used for the statistic to avoid confusion.

CLT
$\rightarrow$ *central limit theorem*

coefficient of determination (r^2)
A normalized measure of discrepancy that expresses the fraction of *variance* that is explained by a statistical *model* ("explained variance"). The coefficient is usually denoted by the symbol r^2. Its value is in the interval $[0,1]$. A large fraction of explained variance indicates a good fit of the model to the data. The coefficient of determination is calculated as follows:

$$r^2 \;=\; \frac{TSS - RSS}{TSS} \;=\; 1 - \frac{RSS}{TSS} \;=\; 1 - \frac{\sum(y - f(x))^2}{\sum(y - \bar{y})^2}$$

where y is an actual outcome, $\bar{y}$ is the mean of outcomes, and $f(x)$ is a prediction based on an *independent variable* x made by the model f. The *residual sum of squares* (RSS) measures the unexplained variance and the *total sum of squares* (TSS) the total variance. Their difference is the explained variance. See figure COD. See also: *linear regression*.

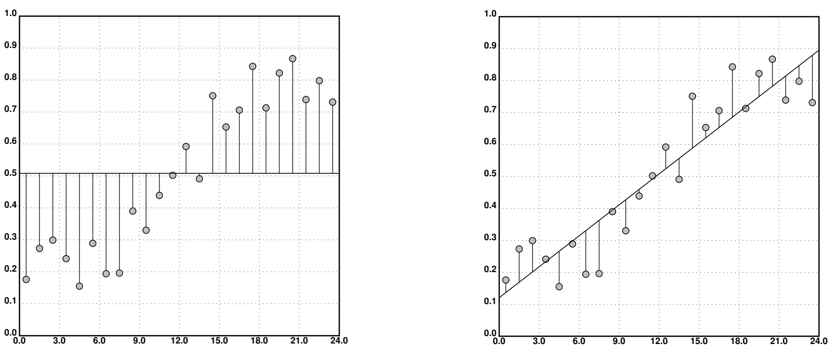


Figure **COD:** left: total variance (TSS); right unexplained variance (RSS) given a linear model; their difference is the explained variance (TSS−RSS) of the model

combination
$\rightarrow$ k*-combination*

complement ($\bar{A}$, A')
An *event* that does not occur. When the *probability* of an event A to occur is $P(A)$, then the complementary probability of the event is $P(\bar{A}) = 1 - P(A)$. Therefore, every event A must either occur or not occur. The events A and $\bar{A}$ are called "complementary events". Probability is often expressed using the letter p and, similarly, complementary probability is expressed using $q = 1 - p$.

For example, rolling a 3 with a six-sided die has a probability of $p = \frac{1}{6}$, so the probability of not rolling a 3 is $q = 1 - \frac{1}{6} = \frac{5}{6}$.

conditional probability
The *probability* of an *event* B in the case that the event A already has occurred:

$P(B|A)$

This is also called the probability of "B given A". Iff A and B are independent (→ *independence*), the probability of $P(B|A)$ is equal to $P(B)$. The probability $P(A|B)$ is called the *reverse conditional probability*. It is computed using *Bayes theorem*. Conditional probability can be visualized using a *decision tree*.

confidence interval (CI)
An interval estimate (→ *estimator*) that specifies a range of values so that there is a given *probability* for an unknown parameter to lie within that range. The given probability is known as the *level of confidence*, c. A confidence interval is specified as

$P(a \leq X \leq b) = c$

i.e. there is a probability of c for the *random variable* X to take a value between (and including) a and b, where X follows the *probability distribution* of the unknown parameter. Typical values of c are 0.9, 0.95, 0.99, and 0.999. With increasing levels of confidence, the range of the confidence interval also increases.

Due to the *central limit theorem*, the *random variable* $\bar{X}$, which is the *sampling distribution of the mean*, is normally distributed (→ *normal distribution*). So when the unknown parameter is μ, the *population mean*, and $X = \bar{X}$, then X can be normalized (→ *standard normal distribution*), thereby transforming the above formula to

$$P(-z \le Z \le z) \; = \; P(\frac{a-\hat{\mu}}{\sigma} \le Z \le \frac{b-\hat{\mu}}{\sigma}) \; = \; c$$

At this point, the *quantile function* F_Z^{-1} of Z can be used to find the values for $\frac{a-\hat{\mu}}{\sigma}$ and $\frac{b-\hat{\mu}}{\sigma}$ and hence for a and b:

$$z \; = \; F_Z^{-1}(c + \frac{\alpha}{2})$$

Because the normal distribution is a *two-tailed* distribution, half the size of the *critical region* has to be added to the confidence level c (where $\alpha = 1 - c$). Once z is known, the confidence interval can be specified as

$$P(-z\sigma + \hat{\mu} \le X \le z\sigma + \hat{\mu}) \; = \; c$$

The meaning of this statement is that there is a $\alpha = 1 - c$ probability for the parameter μ to lie outside of the given interval while the sample X was skewed by chance. See also: *hypothesis test*, *significance*, *level of significance*.

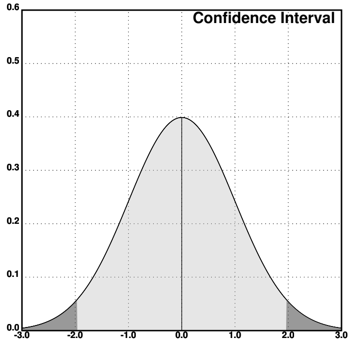


Figure **CFI:** confidence interval on the standard normal distribution with confidence level $c = 0.95$; light gray area: (probability of the) confidence interval; dark gray area: (probability of the) critical regions; significance level $\alpha = 0.05$, two critical regions with probability $\frac{\alpha}{2} = 0.025$.

Example: Given a sampling distribution of the mean $\bar{X}$ of diameters of marbles with estimated mean $\hat{\mu} = 12.5mm$ and $\sigma = 0.2mm$, the confidence interval for the *mean* at a $c = 0.95$ level of confidence would be computed as follows:

$$P(-z \le Z \le z) \; = \; P(\frac{a-12.5}{0.2} \le Z \le \frac{b-12.5}{0.2}) \; = \; 0.95$$

$$z \; = \; F_Z^{-1}(0.95 + \frac{0.05}{2}) \; \approx \; 1.96$$

$$P(-1.96 \cdot 0.2 + 12.5 \ \leq \ X \ \leq \ 1.96 \cdot 0.2 + 12.5) \ \approx \ 0.95$$

$$P(12.11 \leq \bar{X} \leq 12.89) \ \approx \ 0.95$$

I.e., the probability for the true mean μ of the diameter to not lie within the given interval (and the sample being skewed by chance) would be $1 - 0.95 = 0.05$. See figure CFI for an illustration.

confidence level
→ *level of confidence*

contingency table
A *multivariate frequency table*. The most common variant of the contingency table is the *bivariate* version, which lists the categories of one variable in the top row and those of the other in the left column. In addition it lists the marginal total of each category at the opposite end of the corresponding row/column and a grand total in the lower right corner. See figure CTT.

	X_1	$\cdots$	X_k	
Y_1	X_1Y_1		X_kY_1	$\sum_{Y_1}$
$\cdots$				
Y_h	X_1Y_h		X_kY_h	$\sum_{Y_h}$
	$\sum_{X_1}$	$\cdots$	$\sum_{X_k}$	$\sum_{X,Y}$

Figure **CTT:** contingency table of the size $k \times h$; the variable X has k categories and Y has h categories

The fields of the table list the *intersections* of the categories. For example, the field X_1Y_1 lists the number of items that belong to category X_1 *and* to category Y_1. The grand total $\sum_{X,Y}$ of the table is the sum of either $\sum_X$ or $\sum_Y$, i.e. the sum of all fields in the table.

Various *probabilities* can be derived from the table, for instance:

$$P(X_c) \ = \ \sum_{X_c} / \sum_{X,Y}$$

$$P(X_c \cap Y_d) \ = \ P(X_cY_d) \ = \ \frac{X_cY_d}{\sum_{X,Y}}$$

$$P(X_cY_d \mid X_c) \ = \ \frac{X_cY_d}{\sum_{X_c}}$$

The *expectation* of the contingency table is a table in which the variables are *independent*. Its fields are calculated as follows:

$$E(X_iY_j) = \frac{\sum_{X_i} \sum_{Y_j}}{\sum_{X,Y}}$$

Because the table rows and columns implement *conditional probabilities* and the expectation assumes that the variables X and Y are *independent*, the expectation will be close to the corresponding value in the table if, and only if, X and Y are weakly correlated (→ *correlation*).

The accumulated sum of deviations from the expectation can be used to perform a χ^2 test (→ *chi-square distribution*, *hypothesis test*) with $\nu = (k-1) \cdot (h-1)$ *degrees of freedom*. This test estimates (→ *estimate*) the probability of the variables X and Y being independent, i.e. it is a "test of independence". The χ^2 test *score* is calculated as follows:

$$\chi^2 = \sum_{i=1}^{k} \sum_{j=1}^{h} \frac{[E(X_iY_j) - X_iY_j]^2}{E(X_iY_j)}$$

For this method to work, no field in the table should have a value less than 5.

Example: The following table lists people by behavioral priority (X) and gender (Y). Expectations are given in parentheses.

		Priority (X)		
		Assertiveness	Empathy	Total
Gender (Y)	Male	87 (82.3)	13 (17.7)	100
	Female	57 (61.7)	18 (13.3)	75
	Total	144	31	175

Given this table, for instance, the probability of

- being a woman is $P(Y = Female) = 75/175 \approx 0.43$
- preferring empathy given being male is
  $P(X = Empathy \mid Y = Male) = 13/100 \approx 0.13$,

The χ^2 score of the table is $\chi^2 \approx 3.59$ with one degree of freedom, giving a probability of $1 - F_{\chi^2(1)}(3.59) \approx 0.06$ for the variables being independent while the *observations* in the table arose by chance (F_{χ^2} being the *cumulative distribution function* of the χ^2-distribution).

continuous (adj)
Something that is not composed of individual units. For instance, an interval of real numbers or a period of time cannot be broken down into inseparable, individual parts. Even when divided into smaller parts, these parts can always be divided into still smaller parts. Continuous *probability distributions* assign probabilities to continuous measures, like time, size, weight, etc. Compare: *discrete* distribution.

continuous uniform distribution $(X \sim U(a, b))$
A *probability distribution* that models equal *probability* across its *sample space*, i.e. its *probability density function* is a constant function, so all equally sized intervals in its distribution are associated with the same probability. See figure CUD for details, figure CUC for the curves of the distribution functions.

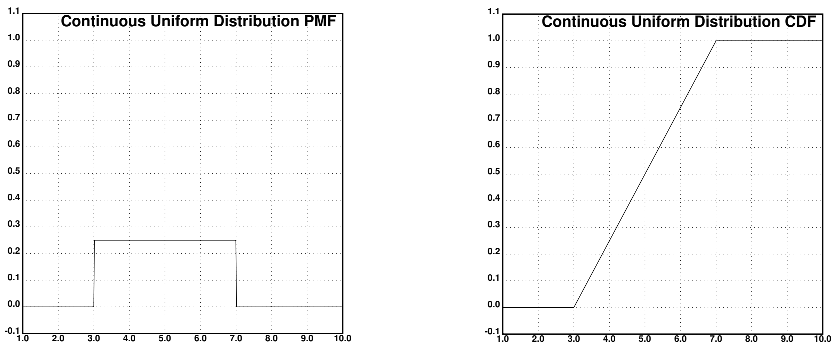


Figure **CUC:** continuous uniform distribution probability functions; left: PDF of $X \sim U(3, 7)$, right: CDF of same distribution

Example: The probability of randomly drawing a real number between one and ten is uniformly distributed. The corresponding random variable follows the uniform distribution $X \sim U(1, 10)$. So, for instance, the probability of drawing a number in the interval $[5, 7]$ would be:

$$F_U(7) - F_U(5) = \frac{6}{9} - \frac{4}{9} = \frac{2}{9}$$

correlation
An indicator for some form of relationship between two *random variables*. Variables that are not correlated are said to be uncorrelated or independent ($\rightarrow$ *independence*). Formally, the

correlation coefficient is used to determine the degree of correlation of two variables. Variables are rarely completely uncorrelated, but the correlation can be very weak. A visual hint about the correlation of two variables can be obtained by creating a *scatter plot* of the variables. When the value of one variable grows smaller as the value of the other variable grows larger, this is called "negative correlation" or *anticorrelation*.

$X \sim U(a,b)$	
PDF	$f(x) = \begin{cases} 0 & \text{if } x < a \\ \frac{1}{b-a} & \text{if } a \le x \le b \\ 0 & \textit{if } x > b \end{cases}$
CDF	$F(x) = \begin{cases} 0 & \text{if } x < a \\ \frac{x-a}{b-a} & \text{if } a \le x \le b \\ 1 & \textit{if } x > b \end{cases}$
Statistic	$x \in \mathbf{R}$: outcome
Parameters	$a, b \in \mathbf{R}; a \le b$: range
μ	$\frac{a+b}{2}$
σ^2	$\frac{(b-a)^2}{12}$
Skewness (γ_1)	0

Figure **CUD:** continuous uniform distribution

correlation coefficient (Pearson's r, r, ρ)
A normalized measure indicating the degree of *correlation* between two *random variables* X and Y, where $r \approx 0$ means that the variables are asymptotically uncorrelated, $r = 1$ indicates perfect correlation $(X = Y)$ and $r = -1$ indicates perfect anticorrelation $(X = -Y)$. The correlation coefficient is defined as:

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sigma_X \sigma_Y}$$

where $cov(X,Y)$ is the *covariance* of X and Y and σ_X and σ_Y are the *standard deviations* of X and Y, respectively.

A similar operator exists for *samples* but, since no *Bessel's correction* is needed (→ *covariance*), it is in fact equal to the above. Formally, it is defined as:

$$r_{x,y} = \frac{s_{x,y}}{s_x s_y}$$

covariance
A measure of the *correlation* of two *random variables* formed by the *mean* of the product of their deviations:

$$cov(X,Y) = \frac{1}{n}\sum(X-\mu_X)(Y-\mu_Y)$$

When the covariance of two variables is a positive value, the variables are (positively) correlated and when the covariance is negative, they are anticorrelated or negatively correlated (→ *anticorrelation*). To express the degree of correlation, the normalized covariance is used. See *correlation coefficient*.

The *sample* covariance is defined as

$$s_{x,y} = \frac{1}{n}\sum(x-\bar{x})(y-\bar{y})$$

Because covariance is a measure of correlation rather than *dispersion* (→ *variance*), no *Bessel's correction* is performed.

critical region
The interval (or intervals) on the domain of a *probability distribution* that lie outside of a *confidence interval* (CI). The probability of a *random variable* X taking a value in the critical region is $\alpha = 1 - c$, where c is the *level of confidence* associated with the CI. The parameter a is called the *level of significance*.

When the probability distribution in which the CI lies is a *one-tailed* distribution, then there is one critical region, and when the distribution is *two-tailed*, there are two critical regions, one in each tail of the distribution, each with a probability of $\frac{\alpha}{2}$.

See *hypothesis test*, *confidence interval*, and z*-test for location* for further explanation and illustrations.

cross validation (CV)
The process of measuring the *cross validation error* (CVE) of a *model*.

cross validation error (CVE)
An absolute measure of discrepancy for the predictions made by a given *model*, where smaller values indicate a better fit of the model to the data. It differs from the *mean squared error* (MSE), because it measures the *test error* instead of the *training error*, i.e. it approximates the quality of the predictions made by a model. There are different ways to compute the CVE. A simple variant is the "leave one out CV" (or LOOCV), which is computed as follows:

$$CVE = \frac{1}{n}\sum_{i=1}^{n}(f_i(x_i) - y_i)^2$$

where f_i is a model fitted on the data set S_i which is the original data set S with the i^{th} value left out, i.e.:

$$S_i = \{x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n\}$$

So the LOOCV uses a subset S_i of S to fit a model and then computes the *error* of the omitted element $\hat{y}_i = f_i(x_i)$, whose actual value y_i is known from the data. This process is repeated for each pair $x_i, y_i \in S$. The CVE is the mean of all of the computed errors. Because LOOCV has to fit the model f n times, it is expensive to compute.

Other means of computing the CVE are the LpOCV, which leaves out p elements instead of one, and the validation set approach, which divides the set S into a "training set" and a "test set" and then uses the training set for model fitting and the test set for computing the CVE. The latter method is biased, though, (→ *bias*) because the CVE here depends on the choice of the test set.

cumulative distribution function (CDF, F_X, F)
The integral of the *probability density function* (PDF) or a sum over the *probability mass function* (PMF) of a *probability distribution*. It is denoted by the symbol F_X (or just F, if X is implied), where X is the *random variable* X or the probability distribution of X. Each distribution has its own specific CDF.

The CDF $F_X(x)$ accumulates the *probabilities* of all values in the *sample space* of X between (and including) the minimum $min(X)$ of X and its argument x, i.e. it expresses the probability $P(X \leq x)$ of X taking a value no larger than x. For *discrete* random variables, it is defined as the sum

$$F(x) = \sum_{w=min(X)}^{x} f(w)$$

and for *continuous* variables, it is the integral

$$F(x) = \int_{min(X)}^{x} f(w)dw$$

The CDF is often used to describe the probability of a random variable taking a value that falls into a specific interval in the sample space of a random variable. The probability $P(a \leq X \leq b)$ of X taking a value between (and including) a and b is

$$P(a \leq X \leq b) = F(b) - F(a)$$

The probability $F_X(x)$ of a continuous random variable X can be visualized as the area under the corresponding PDF curve from $min(X)$ to x. Its value is the proportion of the area between $min(X)$ and x and the area under the entire curve. Because the entire curve equals the sample space of X, the area under the whole PDF curve always equals $\int_{min(X)}^{max(X)} f(x)dx = 1$. Hence the value $F_X(x)$ is exactly the probability $P(X \leq x)$.

Example: the probability of a z-*score* falling within the interval $[1, 2]$ is

$$\begin{aligned} P(1 \leq Z \leq 2) &= F_Z(2) - F_Z(1) \\ &= \int_{-\infty}^{2} f_Z(x)dx - \int_{-\infty}^{1} f_Z(x)dx \\ &\approx 0.977 - 0.841 = 0.136 \end{aligned}$$

where F_Z is the CDF of the *standard normal distribution*. See figure CDX for an illustration.

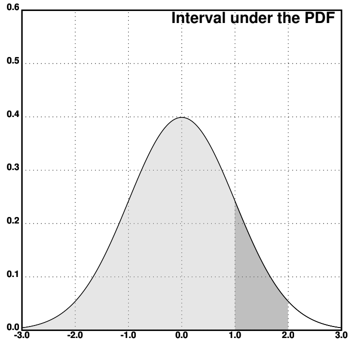


Figure **CDX:** interval under the PDF; complete gray area: $F_Z(2)$; light gray area: $F_Z(1)$; dark gray area: $F_Z(2) - F_Z(1)$; note: the domain of F_Z is $(-\infty, \infty)$, so not the entire curve is shown

D

data generating function (DGF)
An assumed function that generates the data of a *data set* without any *reducible error*, i.e. a function that matches any observed *data points* perfectly within the limits set by the inherent *variance* of the data generating process (the *irreducible error*). The data generating function is an ideal function and cannot normally be observed in practice. The purpose of a statistical *model* is to match the data generating function as closely as possible.

data point
A value contained in a *data set*.

data set
A collection S of values $s_1, \cdots, s_n$. A data set is different from a mathematical set, because it may contain the same value multiple times, i.e. $s_i = s_j$ where $i \neq j$ is possible. In mathematical terms, a data set is vector. When a data set contains values describing some quantifiable property of a set of *specimens* or the *outcomes* of an experiment, it is called a *sample*.

decision tree
A tree diagram with *probabilities* and *conditional probabilities* on its edges. See figure DCT. The leaf nodes usually contain conclusions. In the figure, they list joint probabilities.

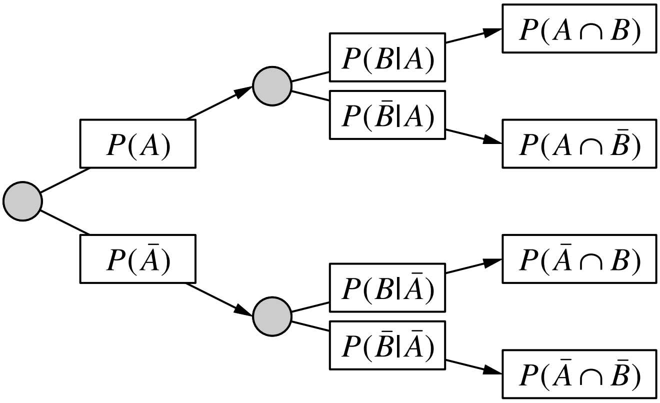


Figure **DCT:** decision tree

The figure shows a decision tree with two variables and hence two levels. The first level, originating from the root node to the left, determines whether or not the *event* A occurs. When the event occurs, the $P(A)$ branch is followed, else the $P(\bar{A})$ branch (→ *complement*) is taken. At the next level A versus $\bar{A}$ is already known, so it lists the probabilities of $P(B|A)$ ("B given A") and $P(\bar{B}|A)$ in the upper branch and $P(B|\bar{A})$ and $P(\bar{B}|\bar{A})$ in the lower one.

At each vertex, the probability of all branches sums up to $p = 1$. For instance, $P(A) + P(\bar{A}) = 1$ and $P(B|A) + P(\bar{B}|A) = 1$. The joint probabilities (like $P(A \cap B)$) are computed by multiplying the probabilities that lead to their respective leaf vertexes (e.g. $P(A \cap B) = P(A) \cdot P(B|A)$.

Decision trees are useful for visualizing conditional probability and *reverse conditional probability*.

degrees of freedom (ν (nu), d.f.)
The number of values that may vary in the computation of a *statistic*. For example, the computation of the *expectation* of rolling a six-sided die involves 5 degrees of freedom, because there are 6 possible *outcomes* $s_1, \cdots, s_6$, but only the frequencies of 5 of them are free to vary, because in order to compute the expectation, the number of trials (n) has to be fixed, and

$$n = s_1 + s_2 + s_3 + s_4 + s_5 + s_6$$

Varying all 6 frequencies would not necessarily add up to n, so only 5 of the parameters contribute a degree of freedom.

Another, more precise way to express the degrees of freedom is the number of outcomes minus the number of necessary relations between the outcomes. Degrees of freedom are often required in combination with certain *probability distributions*, like the χ^2-distribution (→ *chi-square distribution*) or the *t-distribution*. The required number of degrees of freedom is explained in the entry about the corresponding distribution.

dependent variable
A variable whose value will be determined by a *model* or function and hence cannot be chosen freely. Examples would be the y in

$y = f(x)$ or the *explained variable* in statistical models. Compare: *independent variable*.

discrete (adj)
Something that is composed of separate, individual units. For example, a number is composed of digits and a flock is composed of birds. It does not make any sense to talk about fragments of digits or birds. Discrete *probability distributions* assign probabilities to *frequencies*. Compare: *continuous* distribution.

discrete uniform distribution ($X \sim U(a,b)$)
A *discrete probability distribution* that models equal *probability* across its *sample space*, i.e. each of its *outcomes* has the same probability of occurring. See figure DUD for details, figure DUC for the curves of the distribution functions.

$X \sim U(a,b)$	
PMF	$f(x) = \begin{cases} 0 & \text{if } x < a \\ \frac{1}{b-a+1} & \text{if } a \le x \le b \\ 0 & \textit{if } x > b \end{cases}$
CDF	$F(x) = \begin{cases} 0 & \text{if } x < a \\ \frac{x-a+1}{b-a+1} & \text{if } a \le x \le b \\ 1 & \textit{if } x > b \end{cases}$
Statistic	$x \in \mathbf{Z}$: outcome
Parameters	$a, b \in \mathbf{Z}; a \le b$: range
μ	$\frac{a+b}{2}$
σ^2	$\frac{(b-a+1)^2 - 1}{12}$
Skewness (γ_1)	0

Figure **DUD:** discrete uniform distribution

Example: The probability of each face of a six-sided die to show up is uniformly distributed. The corresponding random variable follows the uniform distribution $X \sim U(1,6)$. Hence

$$f_U(1) = f_U(2) = \cdots = f_U(6) = \frac{1}{6-1+1} = \frac{1}{6}$$

and, for instance, the probability of rolling anything smaller than a 4 would be:

$$F_U(3) = \frac{3-1+1}{6-1+1} = \frac{1}{2}$$

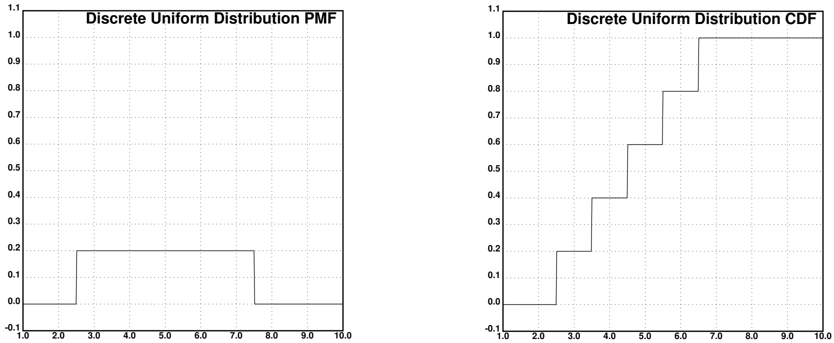


Figure **DUC:** discrete uniform distribution probability functions; left: PMF of $X \sim U(3,7)$, right: CDF of same distribution

dispersion

The extent to which *observations* differ from a given *expectation*. Most common measures of dispersion include the *range*, the *interquartile range*, the *variance*, and the *standard deviation*. Larger values of a measure of dispersion typically indicate that data points are spread out farther from the *central tendency*. Figure DIS illustrates the dispersion of data points around a central tendency, in this case the *mean*.

distribution parameter

A parameter influencing the shape or *location* of the density curve (→ *probability density function*) of a *probability distribution*. Shape parameters influence the form of the curve, like the σ parameter of the *normal distribution* or the ν parameter of the χ^2 (→ *chi-square distribution*) and *t-distribution*. Location parameters typically specify the location of the origin or *mean* of the distribution on the x-axis of a Cartesian coordinate system.

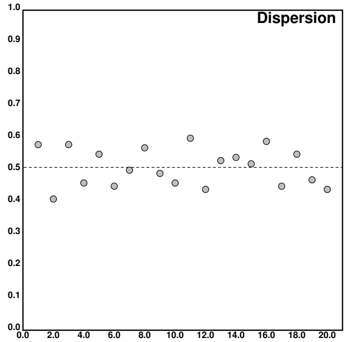


Figure **DIS:** dispersion, dashed line = mean

E

error
The difference between an *estimate* and a corresponding *observation*. In most practical cases, there is no known perfect *model* for describing a series of observations, which is why there is some inherent "noise" in each *sample*, which is called the "irreducible error". No matter how well a model fits the given data, there will always be some variance due to the irreducible error. The "reducible error" is the difference between the total error and the irreducible error. The process of statistical modeling attempts to minimize the reducible error.

Methods for measuring the error of an estimate include, among others, the *standard error of the estimate*, the *mean squared error* (MSE), and the *residual sum of squares* (RSS). See *model* for a visualization. See *type I error* and *type II error* for other uses.

error of the first kind
→ *type I error*

error of the second kind
→ *type II error*

estimate
A value produced by an *estimator*.

estimator ($\hat{\theta}$)
A function that is used to infer (→ *inference*) an estimate of an unknown parameter of a statistical *model*, like the *mean* of a *population* or the *data generating function* of a *data set* or *probability distribution*. The estimator $\hat{\theta}$ estimates the unknown parameter θ. When θ is a *random variable* the estimator itself is a function $\hat{\theta}(X)$ on that random variable, and the estimate for a specific value $X = x$ is demoted by $\hat{\theta}(x)$. Sometimes $\hat{\theta}$ itself is treated as a random variable. The *error* of the estimator is $\hat{\theta}(x) - \theta$.

A "hat" is typically used to indicate an estimator. For example, the estimator of the population mean is equal to the sample mean:

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

The parameter θ would in this case be the true population mean, μ.

event
A set of *outcomes* with a *probability* assigned to them. In each trial, an event does happen or does not happen ($\rightarrow$ *Bernoulli trial*). For example, rolling a die has the *sample space* $S = \{1, 2, 3, 4, 5, 6\}$, so, for instance, the event $X = 3$ of rolling a 3 has the probability $P(X = 3) = \frac{1}{6}$ assigned to it, because the probability for each outcome to occur is $\frac{1}{6}$. The probability assigned to the event of rolling an odd number would be $P(X \in \{1, 3, 5\}) = \frac{3}{6}$.

The probability assigned to the event of a die having a weight between 10 and 20 grams would be $P(10 \leq X \leq 20)$, but an *experiment* would need to be conducted in order to find out what the actual probability is.

excess kurtosis (ex. kurtosis)
$\rightarrow$ *kurtosis*

expectation ($E(X)$, μ)
The *average* value of all possible values of a *random variable*, weighted by *probability*. The expectation is also referred to as the *mean* and then the letter μ is used to refer to it. For the computation of the mean value of a *data set*, see *mean*. The expectation of a numeric frequency distribution is computed as follows:

$$E(X) = \sum_{i=1}^{n} x_i \cdot P(x_i)$$

For instance, when buying lottery tickets at a cost of 1 currency unit each, the random variable X described by the following frequency table may list the probabilities of winning a prize:

X	-1	10	50	100
$P(X)$	0.9655	0.03	0.004	0.0005

Then the expectation $E(X)$ of the variable is:

$$0.9655 \cdot -1 \ + \ 0.03 \cdot 10 \ + \ 0.004 \cdot 50 \ + \ 0.0005 \cdot 100 \ \approx \ -0.42$$

That is, the average return from buying a ticket will be -0.42 currency units.

The expectation of a *probability distribution* is the value x for which the *cumulative distribution function* (CDF) F of the distribution yields $F(x) = 0.5$, i.e. the point x at which the area under the CDF curve is split in half or, more formally, where $P(X \leq x) = 0.5$ and $P(X \geq x) = 0.5$. The expectation of a probability distribution can be computed using the *quantile function* F^{-1} of that distribution or by using a closed formula that is specific to the distribution.

experiment
A fixed number of *trials* resulting in a *sample* (a set of *observations*). Each observation in an experiment must be contained in the *sample space* of the experiment. For instance, the sample space of tossing a coin is $S = \{heads, tails\}$, so only these *outcomes* can be considered. When the coin lands on its edge or disappears, the trial must be discarded.

More generally, any process that generates a sample may be called an experiment. For example, asking questions to people on the street or selecting and quantifying random *specimens* of a *population* are experiments.

explained variable
See *dependent variable*, *linear regression*, *regression analysis*.

explained variance
→ *coefficient of determination*

explanatory variable
See *independent variable*, *linear regression*, *regression analysis*.

F

failure
→ *Bernoulli trial*

frequency
The number of times an *outcome* appears in an *experiment*. See *frequency table*.

frequency distribution
A *random variable* mapping a *category* to a *frequency*. The most common way to express a frequency distribution is the *frequency table*, the most common way to visualize it is the *histogram*. Since the frequency distribution assigns frequencies to categories, it is a special case of the *discrete probability distribution*.

frequency table
A table listing the frequencies of *outcomes*. A frequency table is typically used when the outcomes of an *experiment* fall into *categories*. The following table lists the frequencies of vehicles passing a point on a road.

Vehicles by type						
Type	Bicyle	Motorcycle	Car	Truck	Other	**Total**
Frequency	57	11	90	17	5	180

The probability of each category in the table can be computed by dividing the frequency of the category by the total sum of frequencies. Hence the frequency table implements a *random variable*. In the above table the probability for a car to pass the point would be $P(X = Car) = \frac{90}{180} = 0.5$.

A common way to visualize a frequency table is the *histogram*. *Multivariate* frequency tables exist, see *contingency table*.

G

Gauss curve (bell curve)
→ *normal distribution*

Gaussian distribution
→ *normal distribution*

geometric distribution ($X \sim Geo(p)$)
A *discrete probability distribution* that models the *probability* of a given number of *failures* up to (and including) the first *success* in a fixed number of *Bernoulli trials*. The *random variable* $X \sim Geo(p)$ is geometrically distributed with a probability of success of p.

$X \sim Geo(p)$	
PMF	$f(x) = q^{x-1} \cdot p$
CDF	$F(x) = 1 - q^x$
Statistic	$x \in \mathbf{N}^+$: number of trials
Parameters	$p \in [0, 1]$: probability of success
	q: $1 - p$
μ	$\frac{1}{p}$
σ^2	$\frac{q}{p^2}$
Skewness (γ_1)	$\frac{2-p}{\sqrt{q}}$

Figure **GED:** geometric distribution

For example, the geometrically distributed random variable $X \sim Geo(0.5)$ models the number of coin flips up to getting "heads" for the first time. Given that, the probability of getting heads immediately would be $f_X(1) = 0.5^0 \cdot 0.5 = 0.5$ and the probability of success on the third trial would be $f_X(3) = 0.5^2 \cdot 0.5 = 0.125$. See figure GEC for the corresponding plot.

The CDF of the geometric distribution models the case of not getting x failures in a row, i.e. the probability of getting success at least once in the given number of trials. E.g.: the probability of getting heads at least once in three coin flips would be

$F_X(3) = 1 - 0.5^3 = 0.875.$

For a distribution modeling the probability of a number of successes given a fixed number of trials, see *binomial distribution*.

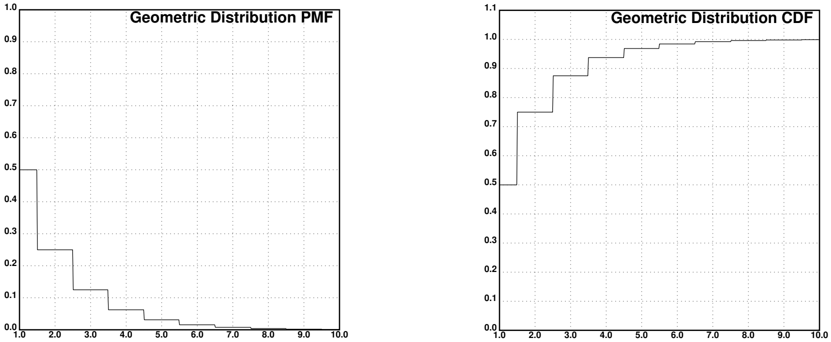


Figure **GEC:** geometric distribution probability functions; left: PMF; right: CDF; both with $p = 0.5$

gradient
→ *slope*

H

histogram

A common way to visualize *frequency distributions* and *discrete probability distributions*. A histogram consists of a bar for each *outcome* (value or *category*) or set of outcomes of the distribution where the height of the bar indicates the frequency of the outcome. See figure HSG (left) for an example.

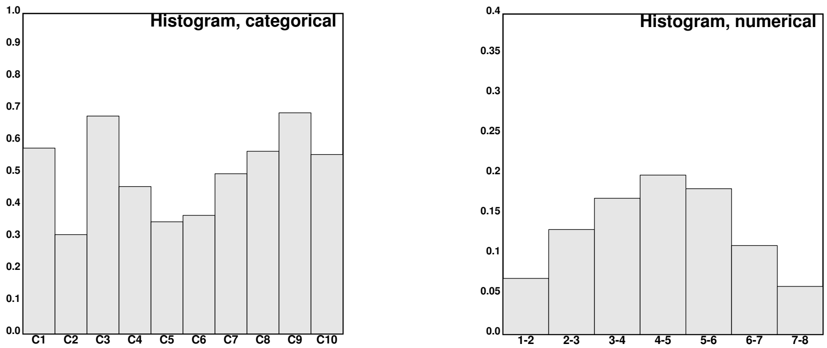


Figure **HSG:** left: histogram from categorical data of the categories C1 ... C10; right: histogram from 8-quantiles of normally distributed data

A histogram can also be created from a *continuous* probability distribution, for example by plotting its q-quantiles ($\rightarrow$ *quantile*). See figure HSG (right).

hypothesis test

A method for gaining confidence in a hypothesis (H_A) by succeeding to reject a contradictory hypothesis, the *null hypothesis* (H_0). The amount of confidence gained by performing the test depends on the *power of the test* which, in turn, depends on the chosen *level of significance*. For instance, when succeeding to reject a null hypothesis at an $\alpha = 0.01$ level of significance, the *probability* of (erroneously) rejecting a true null hypothesis is 1%.

The following procedure is used to conduct a hypothesis test:

(1) A null hypothesis H_0 is stated.

(2) An *alternative hypothesis* H_A is stated.

(3) A level of significance α is stated.

(4) An *experiment* is conducted, resulting in *observations*.

(5) A *test statistic* T is computed from the observations.

(6) If $P(T) > 1-\alpha$, H_0 is rejected and H_A may be accepted.

Hypothesis tests use specific *probability distributions*, like the χ^2-distribution (→ *chi-square distribution*) or the t*-distribution* to calculate the *improbability* of an observation given a null hypothesis. I.e., the value $P(T) = F_X(T)$ of the *cumulative distribution function* (CDF) of the distribution X is exactly the improbability of the observation from which the test statistic T was computed. The *complement* $1-T$ of T is commonly called the *p-value*.

The level of significance α is the complement of the threshold where an observation is deemed so improbable that the null hypothesis has to be rejected. For instance, if $\alpha = 0.05$, then a value of $F_X(T) \geq 0.95$ would lead to the rejection of H_0. (There are also left-tailed tests, where the condition for rejecting H_0 would be $F_X(T) \leq \alpha$).

The intervals between $1-\alpha$ and 1 and/or 0 and α are called the *critical regions* of a test. A test statistic T whose probablity falls within a critical region means that H_0 has to be rejected. The interval outside of the critical region(s) may be considered to be a *confidence interval* at a *confidence level* of c for the observed data to be explained by the null hypothesis.

Some tests use *two-tailed* probability distributions. In this case, there are two critical regions and, subsequently, two conditions for rejecting H_0: $F_X(T) \leq \frac{\alpha}{2}$ and $F_X(T) \geq c + \frac{\alpha}{2}$. See figure HTD for illustrations. For an illustration of a left-tailed test, see z*-test for location*.

When the null hypothesis H_0 can be rejected during a hypothesis test with a significance level of α, the result of the test is stated as "the null hypothesis could be rejected at a α level of significance" or, alternatively, "the alternative hypothesis could be accepted at a α level of significance". Both wordings indicate that there is a probability of $p = \alpha$ of the null hypothesis still being valid while the data gathered in the experiment arose by chance. In other words, there is a probability of $p = \alpha$ of committing a *type I error*.

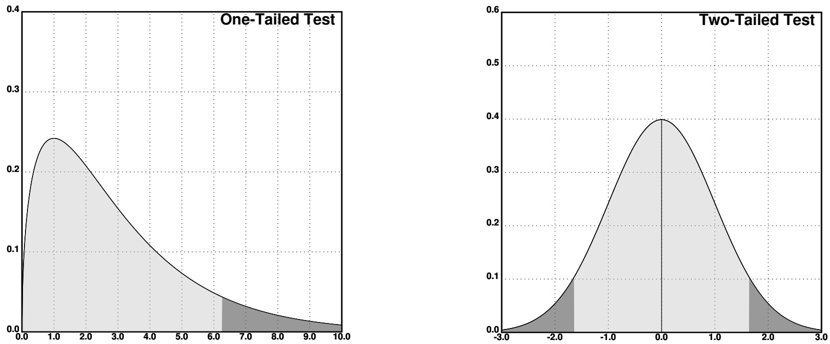


Figure **HTD:** hypothesis test distributions; left: one-tailed $\chi^2(3)$-distribution with one critical region with probability $p = \alpha$; right: two-tailed Z-distribution with two critical regions with probability $p = \frac{\alpha}{2}$ each; both panels: light gray area: confidence interval, dark gray area(s): critical region(s); $\alpha = 0.1$, $c = 0.9$

When the null hypothesis cannot be rejected during a test, the test has no result. In particular, failure to reject H_0 does not mean that the null hypothesis is valid or the alternative hypothesis is invalid. A probability for the test statistic T that is close to $1 - \alpha$ often indicates that further research is advisable.

Example: It is to be shown that a six-sided die is not balanced, i.e. some faces will show up more often than others. Then the null hypothesis H_0 would be that the die is balanced, so that in a series of *trials* that is a multiple of 6 (total of faces) each face shows up the same number of times. The alternative hypothesis H_A would be that the die is not balanced, so some faces show up more often than others. The experiment would be to cast the die $n = 6m$ times, expecting the following *frequency distribution*, which describes an ideal model of a balanced die. Setting $m = 20$:

Eyes	1	2	3	4	5	6	Total
Expected Frequency	20	20	20	20	20	20	120

The level of significance is set at $\alpha = 0.1$, giving a threshold of $p = 0.9$. The actual experiment may then yield the following outcomes:

Eyes	1	2	3	4	5	6	Total
Observed Frequency	27	15	23	16	26	13	120

From the observation and the expectation a χ^2 statistic (→ *chi-square statistic*) can be created:

$$X^2 = \sum_{i=1}^{6} \frac{(O_i - 20)^2}{20} = 9.2$$

where each O_i is an observed frequency. The CDF of the χ^2-distribution with 5 *degrees of freedom* can then be used to compute the improbability of the outcome of the experiment while H_0 holds:

$$F_{\chi^2(5)}(9.2) \approx 0.899$$

Because $0.899 < 0.9$, the null hypothesis cannot be rejected at a $\alpha = 0.1$ level of significance. The data from the experiment is insufficient to conclude that the die is not balanced. However, the test result is so close to the threshold that further experiments seem advisable.

I,J

iid
→ *independent and identically distributed*

improbability (q)
The complement of the *probability* p, i.e. $1-p$. Improbability is often denoted by the letter q.

independence
Two *events* A and B are independent, if the occurrence of one does not influence the other. Formally, A and B are independent, iff

$$P(A \cap B) \ = \ P(A) \cdot P(B)$$

that is, if their joint *probability* (see *intersection*) equals the product of their probabilities. Because

$$P(A \cap B) = P(A) \cdot P(B|A),$$

this statement reduces to

$$P(B) = P(B|A)$$

I.e.: A and B are independent, iff $P(B)$ equals $P(B)$ given $P(A)$ (see *conditional probability*).

independent and identically distributed (adj) (i.i.d., iid)
A set of *random variables* is called iid, if its variables are independent (→ *independence*) and have the same *mean* μ and *variance* σ^2. This is a prerequisite for the *central limit theorem* (CLT).

independent variable
A variable whose value can be chosen freely (sometimes within certain constraints) or controlled. For example, the x in $y = f(x)$ or the *explanatory variable* in statistical *models* are independent. Compare: *dependent variable*.

inference
The process of deducing statistics of a population (like its *mean* or *variance*) on the basis of directly measurable quantities (like corresponding parameters of *samples*) and making judgments

about the accuracy of the inferred data. See *estimator*, *confidence interval*. Also used for the process of understanding the relationship between different *random variables*, for example in statistical modeling (→ *model*) and *regression analysis*.

intercept
The point where the graph of a function of the form $y = ax + b$ intersects with the y-axis of the Cartesian coordinate system. Because at this point $x = 0$, the intercept is equal to $y = 0x + b = b$. The intercept is one of the two coefficients that form a *regression line*, the other one being the *slope*.

interquartile range
(IQR) A measure of *dispersion* that is formed by subtracting the first *quartile* of a *sample* from its third quartile:

$$IQR = Q_3 - Q_1$$

Because the IQR cuts off (roughly) the upper and lower 25% of a sample, it is less sensitive to *outliers* than the *range*.

Example: the data set $x = \{1, 5, 7, 9, 10, 12, 15\}$ has the range $15 - 1 = 14$ and the IQR $12 - 5 = 7$, which ignores the outliers 1 and 15. The three quartiles of the data set are: $Q_1 = 5$, $Q_2 = 9$, and $Q_3 = 12$:

$$\{1, \mathbf{5}, 7, \mathbf{9}, 10, \mathbf{12}, 15\}$$

intersection
The intersection of two *events* A and B is the event of A and B occurring at the same time. The *probability* of the intersection is:

$$P(A \cap B) = P(A) \cdot P(B|A)$$

where $P(B|A)$ is the *conditional probability* of B given A. If the events A and B are independent (→ *independence*), then the above formula reduces to

$$P(A \cap B) = P(A) \cdot P(B)$$

See also: *union*.

interval estimator
An *estimator* that estimates the probability of an unknown parameter to lie within a given range. See: *confidence interval*. Compare: *point estimator*.

IQR
$\rightarrow$ *interquartile range*

irreducible error
An *error* that is inherent in a data generation process ($\rightarrow$ *data generating function*) or *population* and cannot be compensated for in a *model*.

K

k-combination
A set of k elements chosen from a set of n elements without *replacement*, i.e. each element can only be chosen once. In a combination, the order of elements does not matter, so the combinations AB and BA are equal. For instance, the 3-combinations of the set $S = \{A, B, C, D\}$ would be

$$ABC \quad ABD \quad ACD \quad BCD$$

The number of k-combinations from an n-element set is computed by the *binomial coefficient*. See also: *k-permutation*.

k-permutation
An ordered set of k elements chosen from a set of n elements without *replacement*, i.e. each element can only be chosen once. In a permutation, the order of elements does matter, so the combinations AB and BA are different. For instance, the 2-permutations of the set $S = \{A, B, C\}$ would be

$$AB \quad BA \quad AC \quad CA \quad BC \quad CB$$

The number of k-permutations from an n-element set is computed as follows:

$$_nP_k = \frac{n!}{(n-k)!}$$

Sometimes, the n-permutations of an n-element set are referred to as "the permutations of the set". In this case, the number of permutations is simply

$$_nP_n = n!$$

See also: *k-combination.*

kurtosis (K)
A measure for the "tailedness" of a curve, usually used to describe the *probability density functions* of *probability distributions*. A curve with heavy *tails* is called platykurtic, a curve with tiny tails is called leptokurtic. Originally, the *normal distribution* was defined to have a kurtosis of $K = 3$. Lower values indicate platykurtic curves and higher values indicate leptokurtic curves. Sometimes

the normal distribution itself is called “mesokurtic”. The now more common “excess kurtosis” (ex. kurtosis) defines the normal distribution to have a kurtosis of $K = 0$, so positive values indicate lighter and negative values indicate heavier tails.

Strictly speaking, kurtosis measures the proportion of outliers in a distribution and not the shapes of its curves. However, the shape is often related to the amount of outliers. Sample curves are shown in figure KUR.

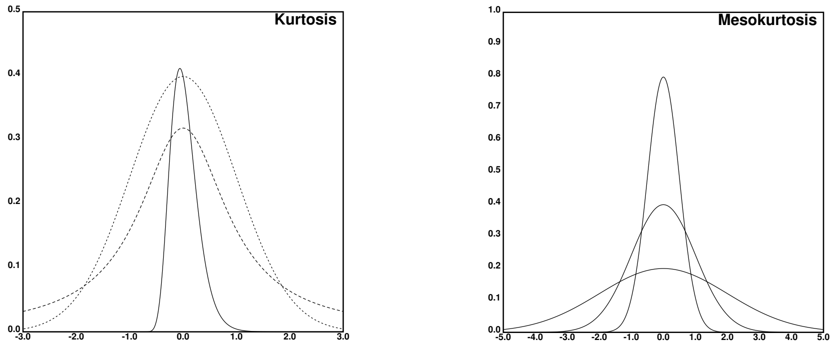


Figure **KUR:** kurtosis; left: leptokurtic scaled ($\frac{1}{4}$) lognormal PDF with $\sigma = 0.25$ and $K \approx 2.3$ (solid); platykurtic t-distribution PDF with $\nu = 1$ and $K = -2.0$ (dashed); mesokurtic standard normal PDF with $K = 0$ (dotted); right: mesokurtic normal PDFs with $\sigma = 0.25$, $\sigma = 1$, and $\sigma = 4$, all with $K = 0$.

L

leptokurtic
→ *kurtosis*

leave one out CV (LOOCV)
→ *cross validation error*

level of confidence (c)
The *probability* for an unknown parameter to fall within a specific *confidence interval*. Sometimes denoted by the letter c, where $c = 0.95$ would indicate a 95% confidence for the parameter to be located in the given interval. Compare: *level of significance*. See also: *hypothesis test*.

level of significance (α)
The acceptable *probability* for rejecting a valid *null hypothesis*, i.e. committing a *type I error*. For example, a level of significance of $\alpha = 0.05$ in a *hypothesis test* would indicate that a 5% chance of incorrectly rejecting a valid *null hypothesis* would be acceptable.

The level of significance is the threshold between a *confidence interval* and a *critical region*. Whenever the result of a test falls within the critical region (which has a probability of exactly $p = \alpha$), the null hypothesis has to be rejected.

linear regression
A simple form of *regression analysis* that attempts to find a linear relationship between two *data sets* or *random variables* X and Y. Linear regression attempts to find the optimal parameters a and b of the function $y = ax + b$ describing the relation between X and Y, i.e. the values for a and b that result in the least *residual sum of squares* (RSS). Visually, linear regression is a line through a *scatter plot* that is as close as possible to all *data points*. The parameter a determines the *slope* of the *regression line*, and b determines its *intercept*. The slope b is calculated by the formula

$$a = \sum \frac{(x-\bar{x})(y-\bar{y})}{(x-\bar{x})^2}$$

where each x is a data point of X, each y is a data point of Y, and $\bar{x}$ and $\bar{y}$ are the *sample means* of the corresponding data sets or variables. Because $\bar{y} = a\bar{x} + b$, the intercept can then be derived from the slope:

$$b = \bar{y} - a\bar{x}$$

Given the *model* $y = ax + b$, a prediction for a value of Y can be made based on a value of X, where X is called the regressor (also: explanatory variable or *independent variable*) and Y the regressand (also: explained variable, *dependent variable*). The accuracy of the predictions made by a linear regression model can be measured by the *coefficient of determination* (r^2), the *mean squared error* (MSE), or the *cross validation error* (CVE).

location
A measure of location is single value that describes a *data set* or *probability distribution*. Typically this is an *average* value of the distribution, most commonly the *mean*. See also: *median*, *mode*, *expectation*.

lognormal distribution $(X \sim Lognormal(\mu, \sigma^2))$
A *continuous probability distribution* that models various natural phenomena which either depend on multiple independent (→ *independence*) *random variables* or randomly partition a finite *population*. Examples for the latter would be the distribution of the sizes of cities or the sizes of the pieces of a shattered vase. A simple example for the former would be the distribution of volumes of cuboids with heights, widths, and depths drawn randomly from a finite interval.

A random variable $X \sim Lognormal(\mu, \sigma^2)$ is lognormally distributed, if the variable $Y \sim ln(X)$ is normally distributed (→ *normal distribution*). Analogously, $X \sim e^Y$ is lognormally distributed, if Y is normally distributed. Sometimes lognormal variables are written as $X \sim e^{\mu+\sigma Z}$, where Z is the *standard normal distribution* and μ and σ are the *mean* and *standard deviation* of Z.

The relationship between normally and lognormally distributed variables is illustrated by the example of the volumes of random cuboids with *discrete* edge lengths in the range $\{1, \cdots, 10\}$. The distribution of cuboid volumes tends toward a lognormal distribution $X \sim Lognormal(ln(16.5), ln(24.75))$. The *frequency distribution* of cuboid volumes is shown in figure RVL. Because the volume of a cuboid is the product of its edge lengths, the logarithm of the volume is the sum of the edge lengths. The sum of lengths then follows a normal distribution $ln(X) \sim N(16.5, 24.75)$, also shown in figure RVL. The parameters of the lognormal distribution are shown on figure LGD, sample plots of its probability functions can be found in figure LGC.

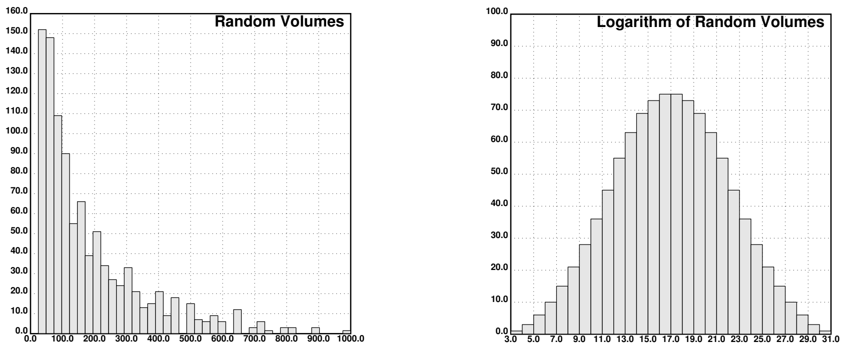


Figure **RVL:** random volumes; left: lognormal distribution of random volumes $X \sim Lognormal(ln(16.5), ln(24.75))$, right: corresponding normal distribution $Y \sim N(16.5, 24.75)$.

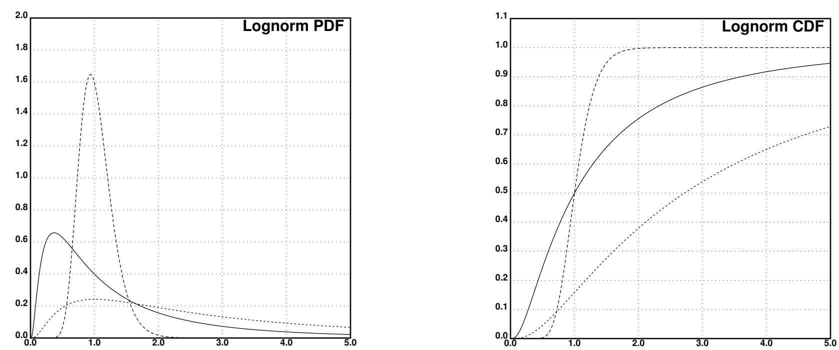


Figure **LGC:** lognormal distribution probability functions: left: PDF with $\mu = 0$, $\sigma^2 = 1$ (solid), $\mu = 0$, $\sigma^2 = 0.0625$ (dashed); $\mu = 1$, $\sigma^2 = 1$ (dotted), right: CDF with same parameters

Example: if the *average* city has $100\,000$ inhabitants and the *variance* is $10\,000^2$, then the distribution of city sizes follows a lognormal distribution $X \sim Lognormal(ln(10^5), ln(10^8))$. Therefore, the probability for a city to have between $1\,000$ and $10\,000$ inhabitants is

$$F_X(10^4) - F_X(10^3) \approx 0.154$$

and the probability for a city to have between one million and two million inhabitants is

$$F_X(2 \cdot 10^6) - F_X(10^6) \approx 0.053$$

$X \sim Lognormal(\mu, \sigma^2)$	
PDF	$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \cdot e^{-\frac{(ln(x)-\mu)^2}{2\sigma^2}}$
CDF	$F(x) = \phi(\frac{ln(x)}{\sigma})$
	$F(x) = \frac{1}{2} + \frac{1}{2} \cdot erf\left(\frac{ln(x)-\mu}{\sigma\sqrt{2}}\right)$
Statistic	$x \in \mathbf{R}_0^+$: raw score
Parameters	$\mu \in \mathbf{R}_0^+$: shape
	$\sigma^2 \in \mathbf{R}^+$: shape
μ	$e^{\mu+\frac{\sigma^2}{2}}$
σ^2	$(e^{\sigma^2} - 1) \cdot e^{2\mu+\sigma^2}$
Skewness (γ_1)	$(e^{\sigma^2} + 2) \cdot \sqrt{e^{\sigma^2} - 1}$

Figure **LGD:** lognormal distribution; ϕ is the CDF of the normal distribution; erf is the Gauss error function; the shape parameters are the variance and location (mean) of $ln(X)$; where $ln(X)$ is the normal distribution $\mu + \sigma Z$ upon which $X \sim e^{\mu+\sigma Z}$ is based

LOOCV (leave one out CV)
→ *cross validation error*

M

mean $(\mu, \bar{x})$
A measure of *location* formed by adding all values in a *data set* and dividing by the number of values (n):

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

The letter μ denotes the *population* mean and $\bar{x}$ the *sample* mean. The mean of a *random variable* X is called its *expectation*.

Example: $\frac{1+2+3+4+5}{5} = \frac{15}{5} = 3$

See also: *expectation*, often used synonymously.

mean squared error (MSE)
An absolute measure for the accurary of an *estimator*, where smaller values indicate a tighter fit of the *model* generating the estimate to the observed data (→ *observation*). The MSE is computed as follows:

$$MSE = E[(\hat{Y} - Y)^2] = \frac{1}{n}\sum_{i=1}^{n}(\hat{Y}_i - Y_i)^2$$

where each $\hat{Y}_i$ is an estimate and each Y_i is the corresponding observed *data point*. It can be shown that the MSE is composed of the *bias* and the *variance* of the model generating $\hat{Y}$:

$$\begin{aligned} MSE = E[(\hat{Y} - Y)^2] &= (E[\hat{Y}] - Y)^2 + E[(\hat{Y} - E[\hat{Y}])^2] \\ &= bias(\hat{Y}, Y)^2 + var(\hat{Y}) \end{aligned}$$

median (sometimes M)
A measure of *location* formed by the value in the middle of an *ordered data set*. If the number of elements is even, i.e. if there is no middle element, the median is the *mean* of both middle elements.

Examples:

$$median(\{1, 2, 3, 4, 5\}) = 3$$

$$median(\{1, 2, 3, 4, 5, 6\}) = \frac{3+4}{2} = 3.5$$

For a formal definition, see *quantile*.

mesokurtic
→ *kurtosis*

mode
A measure of *location* that is equal to the value that occurs most often in a *data set*. When there are multiple values that all occur most frequently, the data set is called “multimodal”. If there are two modes, it is called “bimodal”.

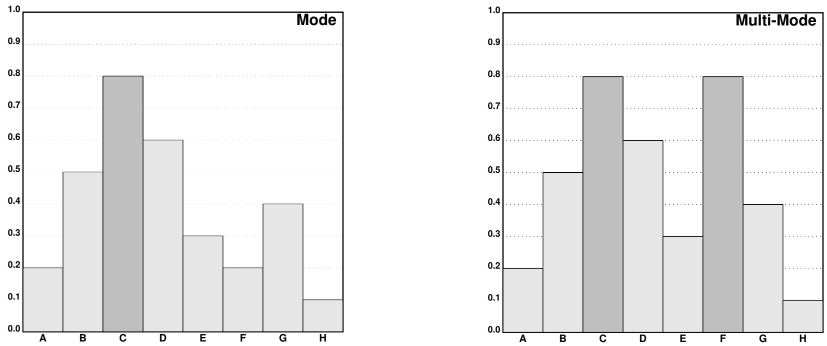


Figure **MDS:** mode (left) and multiple modes (right)

Modes of small data sets can easily be found by plotting their *histograms*. See figure MDS.

model
A function that estimates (→ *estimate*) the distribution of *data points*. More precisely, the existence of a *data generating function* is assumed, which models the generated data perfectly (but causes some *variance* due to the *irreducible error*). The *model* is then a function that resembles the data generating function as closely as possible. The data generating function is generally unknown. *Probability distributions* can be considered to be statistical models.

The *error* of a model is the fraction by which the model fails to predict *observations*. The error is measured by the *mean squared error*, which is in turn composed of the *bias* and *variance* and, in case of statistical modeling, a normally distributed irreducible error, ε:

$$MSE = E[(\hat{Y} - Y)^2] + \varepsilon$$

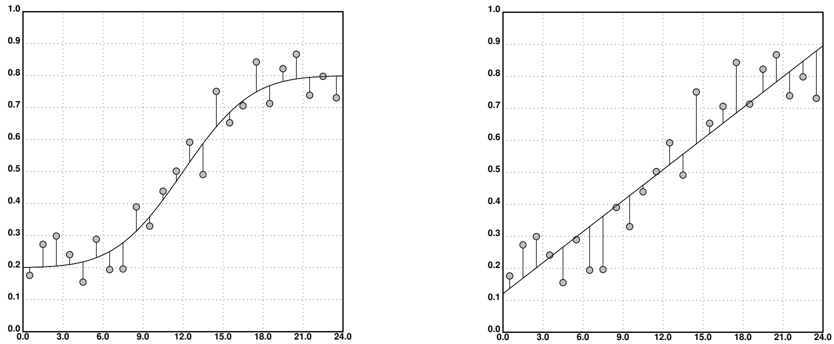


Figure **MOD:** left: data generating function, error bars show the irreducible error; right: linear model, the error bars show the total (reducible and irreducible) error; actual data points are shown as dots

The process that creates statistical models is called *regression analysis*. A simple form of statistical modeling is the *linear regression*. Figure MOD shows a data generating function and a linear model describing the same data set. The accuracy of a model can be examined using the *coefficient of determination*, the mean squared error (MSE), or the *cross validation error* (CVE).

multimodal (adj)
A data set that has multiple *modes*.

multivariate (adj)
A multivariate process is a process involving multiple *random variables*. See, for example: *contingency table*.

N

non-discrete
→ *continuous*

normal distribution (Gaussian distribution, $X \sim N(\mu, \sigma^2)$)
The most common *continuous probability distribution*. A *random variable* $X \sim N(\mu, \sigma^2)$ is normally distributed with mean μ and variance σ^2. The normal distribution plays a central role in statistics due to the *central limit theorem*, which basically says that the *average* of a large number of *iid* probability distributions, no matter of what type, converges toward the normal distribution.

The normal distribution is used to model a lot of natural phenomena, like body weight, intelligence quotients, etc. A measurement or *score*, like body weight or IQ, which follows a specific normal distribution, is called a “raw score”. In order to find out how common a raw score is, it is converted to a z-*score* of the normalized *standard normal distribution* with mean $\mu = 0$ and variance $\sigma^2 = 1$.

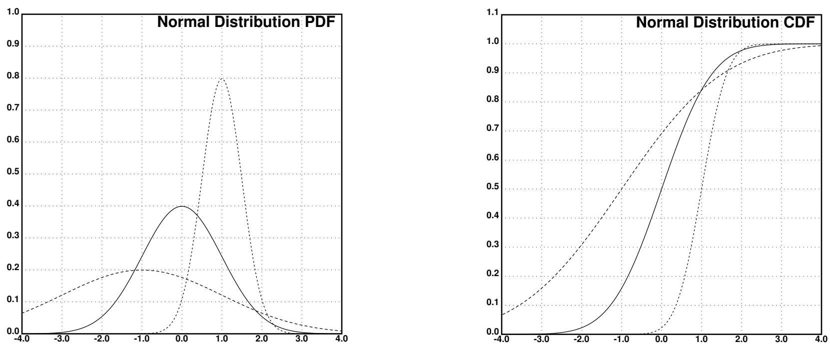


Figure **NOC:** normal distribution probability functions: left: PDF; right: CDF; Both panels: dashed line: $\mu = -1$, $\sigma^2 = 4$, solid line: $\mu = 0$, $\sigma^2 = 1$, dotted line: $\mu = 1$, $\sigma^2 = 0.25$

For example, given an IQ distribution $X \sim N(100, 15^2)$, a raw score of 117 would be correspond to the z-score

$$z = \frac{117 - 100}{\sqrt{15^2}} = \frac{17}{15} = 1.1\bar{3}$$

The *cumulative distribution function* (CDF) of the standard normal distribution can then be used to calculate the *quantile* of the z-score (and hence the quantile of the raw score). See figure NDD for the CDF.

Plotting the *probability density function* of the normal distribution exhibits a characteristic bell-shaped curve called the "bell curve" or "Gauss curve". See figure NOC.

$X \sim N(\mu, \sigma^2)$	
PDF	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
CDF	$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \int_{-\infty}^{x} e^{-\frac{(w-\mu)^2}{2\sigma^2}} dw$
	$F(x) = \frac{1}{2}\left(1 + \frac{erf(x-\mu)}{\sigma\sqrt{2}}\right)$
Statistic	$x \in \mathbf{R}$: raw score
Parameters	$\mu \in \mathbf{R}$: mean
	$\sigma^2 \in \mathbf{R}^+$: variance
μ	μ
σ^2	σ^2
Skewness (γ_1)	0

Figure **NDD:** normal distribution; erf is the Gauss error function

null hypothesis (H_0)
A hypothesis that must be rejected for an *alternative hypothesis* H_A to be accepted. For example, if an *experiment* fails to show that a die is fair (H_0), the alternative hypothesis that the die is loaded (H_A) may be accepted (at a given *level of significance*). The null hypothesis is often the assumption that *data points* adhere to some ideal *model*, like a *probability distribution*. For example, the null hypothesis that a die is fair may be described by a *uniform distribution* of *outcomes* (eyes facing up).

A null hypothesis H_0 is assumed to be true until data gathered in an experiment contradict the hypothesis at a given level of significance.

O

observation
The occurrence of an *outcome* at a specific point of an *experiment*. There are different kinds of observations, like values falling into *categories*, *discrete* values or *continuous* values.

one-tailed (adj)
A *probability distribution* with a single *tail*. More precisely, a distribution whose *probability density function* plots a curve with a single tail. In a *hypothesis test*, a test with one *critical region*; see *z-test for location*.

ordered data set
A *data set* whose elements are ordered, typically in numerically ascending order. For example, the data set $S = \{10, 4, 6, 2, 8\}$ can be ordered by size, giving

$$x_1 = 2,\ \ x_2 = 4,\ \ x_3 = 6,\ \ x_4 = 8,\ \ x_5 = 10$$

where each x_i is an element of S. Ordering a data set is required, for instance, to compute the *median* of a *sample*.

outcome
A possible result of a *trial*. Each trial in a given *experiment* will have an outcome that is contained in its *sample space*. An outcome that is associated with a *probability* is called an *event*.

outlier
A *data point* that is particularly far removed from the *mean* (or other measure of *location*) of a data set. For example, the data set $S = \{1, 2, 7, 8, 9, 11, 12, 17\}$ has the mean $\mu = 8.375$ and *standard deviation* $\sigma \approx 4.9$. The values 1, 2, and 17 are the only ones that are more than one standard deviation removed from the mean, so they can be considered to be outliers. Outliers can severely limit the expressiveness of simple measures of *dispersion*, like the *range*. The proportion of outliers in a *probability distribution* is indicated by its *kurtosis*.

P

p-value (P value, p, P)
The *complement* of the *quantile* of a *test statistic*. Given a *test statistic* $T \sim D$ following a *probability distribution* D, the quantile of the statistic would be $Q = F_D(T)$, where F_D is the *cumulative distribution function* (CDF) of D. That is, if the test statistic in a *hypothesis test* translates to an *improbability* of Q, then the corresponding p-value would be $P = 1 - Q$. So a null hypothesis must be rejected at a *level of significance* of α, if Q falls within a *critical region* of the probability α.

The interpretation of the p-value is the *probability* for the observed (→ *observation*) data to arise by chance while the null hypothesis holds or, in other words, the probability of committing a *type I error*. The p-value is denoted by the symbols p or P, which both have well-known other meanings (→ *probability*). This may cause some confusion.

PDF
→ *probability density function*

Pearson's r
→ *correlation coefficient*

percentile
Any of the 100-quantiles (→ *quantile*). Percentiles are widely used, because they divide a *data set* or a *probability distribution* into 100 equally sized or equally probable "chunks" that can provide an intuitive measure describing some quantifiable property. For example, if your length is in the 75^{th} percentile, then you are taller than 75% of the population (strictly speaking, not shorter than 75% of the population).

permutation
→ k-*permutation*

platykurtic
→ *kurtosis*

point estimator
An *estimator* that estimates a single "best guess" for a statistic.

Compare: *interval estimator*.

Poisson distribution ($X \sim Poi(\lambda)$)
A *discrete probability distribution* modeling the *probability* of an *event* occurring x times per interval, given an average of λ events per interval, where an "interval" may be temporal, spatial, or abstract. The average number of events per interval must be constant. A *random variable* $X \sim Poi(\lambda)$ is Poisson-distributed with an average number of λ events per interval.

For example, the number of water drops falling from a leaky pipe with an average of ten drops per minute follows the Poisson distribution $X \sim Poi(10)$. Then the probability of 15 drops falling from the pipe in one minute is

$$f_X(15) = \frac{10^{15} \cdot e^{-10}}{15!} \approx 0.035$$

and the probability of up to 15 drops falling from the pipe is:

$$F_X(15) = \sum_{i=0}^{i} \frac{10^i \cdot e^{-10}}{i!} \approx 0.92$$

See figure POD for distribution parameters and figure POC for sample plots of the distribution functions.

$X \sim Poi(\lambda)$	
PMF	$f(x) = \frac{\lambda^x \cdot e^{-\lambda}}{x!}$
CDF	$F(x) = \sum_{i=0}^{x} \frac{\lambda^i \cdot e^{-\lambda}}{i!}$
	$F(x) = Q(x+1, \lambda)$
Statistic	$x \in \mathbf{N}_0$: number of events
Parameters	$\lambda \in \mathbf{N}_0$: events per interval
μ	λ
σ^2	λ
Skewness (γ_1)	$\frac{1}{\sqrt{\lambda}}$
Approximations	$N(\lambda, \lambda)$ for $\lambda \geq 15$

Figure **POD:** Poisson distribution; Q is the complement of the regularized incomplete Γ function, a common way for computing the CDF of the distribution

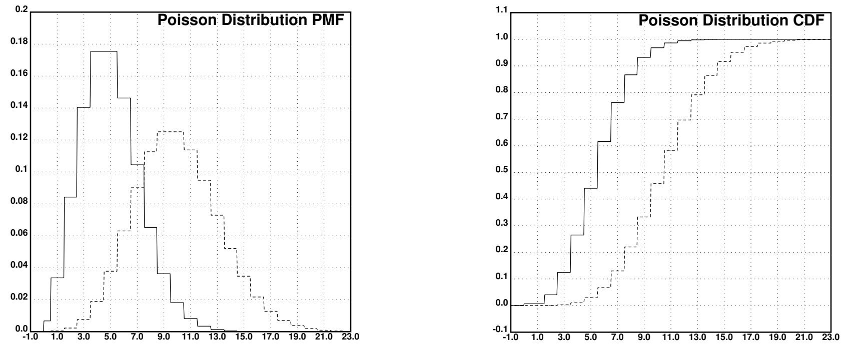


Figure **POC:** Poisson distribution probability functions; left: PMF with $\lambda = 5$ (solid) and $\lambda = 10$ (dashed); right: CDF with same parameters

population

A closed group of *discrete* objects, like all students on a campus, all travels of an elevator car on a specific day, or all attempts to roll 18 eyes with three six-sided dice. Note that a menber of a population does not have to be concrete, nor does the population have to be finite. For instance, any number of attempts can (theoretically) be made to roll 18 eyes. A major goal of *statistics* is to make predictions about a population based on one or multiple *samples* taken from that population.

power of a test

The power of a *hypothesis test* is the *conditional probability* of rejecting a *null hypothesis* given it is invalid:

$$\text{power} = P(H_0 \text{ rejected} \mid H_0 \text{ invalid})$$

It is equal to the complement $1 - \beta$ of the *type II error*.

probability $(p, P(x))$

A measure describing the likelihood of an *event* to occur. Probability is typically expressed as a real number p in the range $[0, 1]$, where 0 indicates (almost) impossibility and 1 means (almost) certainty.

The notation $P(x)$ describes the probability of an *observation* x, and $P(X = x)$ denotes the probability of a *random variable* X to take the value x. For example, the probability of the toss of a fair coin to yield "heads" would be $P(heads) = 0.5$. The notation $P(x_1 \leq X \leq x_2)$ is used express the probability of an observation to fall within an interval of numbers. Here the observation would be X assuming a value between (and including) x_1 and x_2.

probability density function (PDF, f_X, f)
A function describing the relative *probability* of a *continuous random variable* X taking a specific value in its *sample space*. It is denoted by the symbol f_X (or just f, if X is implied), where X is the random variable X or the *probability distribution* of X. Each probability distribution has its own specific PDF.

Relative probability can only compare the likelihood of two points in the sample space of X. For example, the peak of the *standard normal distribution* Z at $x = 0$ would have a relative probability of $f_Z(0) = 0.4$. This does not say anything about the actual probability $P(Z = 0)$, though.

To compute the probability of a continuous random variable falling within a specific interval, the integral of the PDF is used. See *cumulative distribution function* (CDF). The equivalent function for *discrete* probability distributions is the *probability mass function* (PMF).

probability distribution
A statistical *model* that maps *events* to *probabilities*, i.e. it estimates (→ *estimate*) the probability of a given event.

Discrete probability distributions have a *sample space* consisting of individual *outcomes*. Their *probability mass functions* (PMF) deliver the probability of a specific outcome to occur and their *cumulative distribution functions* compute the probability of a range of outcomes to occur. The sample space of discrete distributions may consist of *categories* rather than numeric data.

Continuous probability distributions have a sample space consisting of an interval of real numbers. They use their CDFs to estimate the probability of an outcome falling into a given range. Their counterpart to the PMF is the *probability density function* (PDF), but it is not usually used to compute probabilities.

Common discrete probability distribution include the (discrete) *uniform distribution*, the *geometric distribution*, the *binomial distribution*, and the *Poisson distribution*. Common continuous probability distribution include the *normal distribution*, the

χ^2-distribution (→ *chi-square distribution*), the *t-distribution*, and the *lognormal distribution*.

probability function (f_X, F_X, P)
A function mapping an *outcome* to a *probability*. Mostly used to refer to the functions of *probability distributions*, and among those most frequently used to indicate the *probability density function* (PDF). Sometimes used to describe the probability of a categorical outcome (→ *category*), like $P(Heads)$ to denote a coin showing heads.

probability mass function (PMF, f_X, f)
A function describing the *probability* of a *discrete random variable* X taking a specific value in its *sample space*. It is denoted by the symbol f_X (or just f, if X is implied), where X is the random variable X or the *probability distribution* of X. Each probability distribution has its own specific PMF.

For example, throwing a ball into a bucket may have a probability of success of $p = 0.8$. Then the probability of a single success after a given number of failures is a random variable following a *geometric distribution* $X \sim Geo(0.8)$, whose PMF is

$$f(x) = (1 - p)^{x-1} \cdot p$$

So the probability $P(X = 3)$ of hitting the bucket in the third trial, after two failures is:

$$f_{Geo}(3) \; = \; 0.2^2 \cdot 0.8 \; = \; 0.0032$$

To compute the probability of at least one success, the *cumulative distribution function* (CDF) would be used. The equivalent function for *continuous* probability distributions is the *probability density function* (PDF).

Q

q-quantile
→ *quantile*

quantile
A point dividing a *discrete random variable* or a *data set* into segments of equal size or a *continuous* random variable into intervals of equal *probability*. A q-quantile is any of $q-1$ data points dividing a data set or variable into q segments.

The n^{th} q-quantile ${}_qQ_n$ of a data set $S = \{s_1, \cdots, s_n\}$ is calculated as follows:

$$
{}_qQ_n = \begin{cases} s_1 & \text{if } n = 1 \\ s_{\lfloor n/q \rfloor + 1} & \text{if } \lfloor n/q \rfloor \neq n/q \\ \frac{s_{n/q} + s_{(n/q)+1}}{2} & \text{if } \lfloor n/q \rfloor = n/q \end{cases}
$$

The 2-quantile is also called the *median*, a 4-quantile is called a *quartile*, and a 100-quantile is called a *percentile*. Sometimes the 0^{th} quantile is used to denote the minimum of a random variable or data set and the q^{th} q-quantile is used to refer to its maximum.

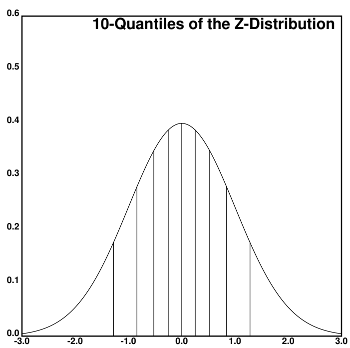


Figure **QNT:** 10-quantiles of the standard normal distribution

The n^{th} q-quantile of a continuous random variable is computed by the *quantile function* of the variable's *probability distribution*. In this case, it is common to talk about fractional quantiles, like the

0.99997^{th} quantile instead of the 99.997^{th} percentile.

Figure QNT shows the 10-quantiles of the standard normal distribution. The areas between the quantiles have the same probability and hence the same size.

quantile function (F_X^{-1}, F^{-1})
The inverse function of the *cumulative distribution function* (CDF) of a *probability distribution*. It is denoted by the symbol F_X^{-1} (or just F^{-1}, if X is implied), where X is the *random variable* X or the *probability distribution* of X. Each distribution has its own specific quantile function. The quantile function $F_X^{-1}(p)$ computes a distribution-specific value x such that $P(x \le X) = p$.

The quantile function is used to compute, for instance, *confidence intervals* and the *level of significance* in *hypothesis tests*. For example, a *two-tailed* confidence interval with a *level of confidence* of $c = 0.9$ would correspond to $F_Z^{-1}(0.95) = 1.64\sigma$ on the *standard normal distribution* (the argument is 0.95, because the distribution is two-tailed).

quartile
Any of the three 4-quantiles (→ *quantile*). The 2^{nd} quartile is also called the *median*. See figure QTS.

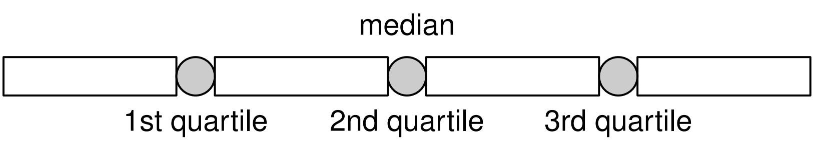


Figure **QTS:** The three quartiles

R

r-squared (r^2)
→ *coefficient of determination*

random sampling
→ *sampling*

random variable (X, Y)
A function that maps an *outcome* to a numerical value. Because some outcomes in the physical world cannot be measured directly, a random variable is used to make them measurable. Most commonly a table is used to map phenomena (outcomes) to values. E.g.:

Outcome	Rain	No rain
Value (X)	1	0
Probability (p)	0.3	0.7

The *probabilities* of the individual values of a random variable always sum up to 1. The probabilities can be omitted, if their distribution is uniform (→ *uniform distribution*). Sometimes a frequency can be used to compute the value of a random variable, like the number of heads when tossing three coins:

Number of heads facing up in 3 coins, H=heads, T=tails								
Outcome	TTT	TTH	THT	THH	HTT	HTH	HHT	HHH
Value (X)	0	1	1	2	1	2	2	3

In many cases outcomes already are measurable. In these cases, the random variable is an identity mapping.

The *probability* of a *discrete* random variable taking a specific value is expressed using the notation $P(X = x)$ where X is the random variable and x the corresponding value. For instance, given the above table, the probability of getting two heads is $P(X = 2) = \frac{3}{8}$ (3 out of 8 cases map to 2).

When a random variable is *continuous*, each point in its distribution has infinitesimally small probability, so values of continuous variables are expressed as a intervals. For instance, the notation $P(1.5 \le X \le 1.7)$ indicates the probability of the random variable X to lie in the interval $[1.5, 1.7]$.

range
A measure of *dispersion* that is formed by subtracting the smallest element of a *sample* from the largest one:

$$range(x) = max(x) - min(x)$$

Example: the data set $x = \{1, 2, 3, 4, 5, 6\}$ has the maximum $max(x) = 6$ and minimum $min(x) = 1$, so its range is $6 - 1 = 5$.

The range is very sensitive to *outliers*. More robust measures of dispersion include the *interquartile range* and the *standard deviation*.

raw score
See *normal distribution* and z*-score* .

rectangular distribution
→ *uniform distribution*

reducible error
An *error* being introduced by a statistical *model*.

regressand
→ *dependent variable*

regression analysis
The analysis of data (→ *data set*) with the intent of finding a *model* that can be used to estimate (→ *estimator*) *data points* that are not contained in the data set under analysis. The model is then used to predict the value of a *dependent variable* based on the value of an *independent variable*. The model found by regression analysis can never predict data points with certainty due the *irreducible error* and the *bias-variance trade-off*. For a simple and common regression model, see *linear regression*.

regression coefficient
→ *linear regression*

regression line
The graph of a *linear regression* function of the form $y = ax + b$, where a and b are called the *regression coefficients*. The

coefficient a determines the *slope* of the function and b its *intercept*. A regression line is often drawn through a *scatter plot* to visualize the relationship between two *random variables*.

regressor
→ *independent variable*

replacement
The act of putting back (replacing) an element after drawing it from a set. For instance, the 2-combinations (→ k*-combinations*) of the set $S = \{A, B, C\}$ without replacement are

$$AB \quad AC \quad BC$$

because a letter cannot be drawn again after using it in a combination. The 2-combinations with replacement of the same set are

$$AA \quad AB \quad AC \quad BB \quad BC \quad CC$$

residual sum of squares (RSS)
A measure of discrepancy between observed outcomes and outcomes predicted by a statistical *model* (→ *regression analysis*). It measures the part of the variance of a *random variable* or *sample* that is not explained by the given model. A small RSS value indicates a good fit of the model to the data. The RSS is computed as follows:

$$RSS = \sum_{i=1}^{n}(y_i - f(x_i))^2$$

where each y_i is an actual outcome and each $f(x_i)$ is an *estimate* predicted by the model f. A normalized measure of discrepancy is provided by the *coefficient of determination*.

reverse conditional probability (RCB)
The *probability* $P(A|B)$ given the *conditional probabilities* $P(B|A)$, $P(B|\bar{A})$, and the probability of $P(A)$ in general (or their *complements*). Because $P(A|B)$ implies that B already has occurred, it is the proportion of $A \cap B$ occurring and B occurring (with or without A also occurring). So the formula for calculating the RCB is

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \cdot P(B|A)}{P(A) \cdot P(B|A) + P(\bar{A}) \cdot P(B|\bar{A})}$$

This formula is widely known as "Bayes rule" or "Bayes theorem".

Example: There are two boxes. Box A contains 12 black and 8 white marbles, box B contains 4 black and 16 white marbles. When randomly drawing a marble, the probability of drawing a marble from box A is $P(A) = 0.5$. The probability of drawing a black marble from box A is $P(Black|A) = \frac{12}{20} = 0.6$. Analogously, $P(Black|\bar{A}) = \frac{4}{20} = 0.2$. (Each box contains 20 marbles.) Given these probabilities, the RCB of a randomly drawn black marble coming from box A would be:

$$P(A|Black) = \frac{P(A) \cdot P(Black|A)}{P(A) \cdot P(Black|A) + P(\bar{A}) \cdot P(Black|\bar{A})} = \frac{0.5 \cdot 0.6}{0.5 \cdot 0.6 + 0.5 \cdot 0.2} = \frac{0.3}{0.4} = 0.75$$

rule of three sigma

In a *normal distribution*, 68.3% of the *data points* are located within one *standard deviation* (σ, sd) from the mean, 95.4% of the data points within two sd, and 99.7% within three sd. This principle is known as the rule of three sigma. The principle is illustrated in figure 3SG.

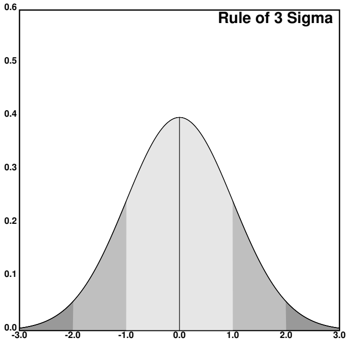


Figure **3SG:** rule of three sigma; first sigma (light gray): p=0.683, second sigma (medium gray, including first): p=0.954, third sigma (dark gray, including 1st and 2nd): p=0.997; vertical line: mean

S

sample
A *data set* created by collecting the *outcomes* of an *experiment*. Often the experiment is to gather quantifiable characteristics of *specimens* randomly chosen from a *population*. If a sample is large enough, it will have properties that are asymptotically equal to those of the population from which the sample was taken. For instance, the sample will have a *mean* and *variance* that is close to that of the population. The size of a sample is typically denoted by the letter n. When population size has to be distinguished from sample size, the letter N is used for the population size. The process of creating a sample is called *sampling*.

sample covariance
→ *covariance*

sample point
A value contained in a *sample*.

sample space
The set S of all possible *outcomes* $x_1, \cdots x_n$ that may occur during an *experiment* or the domain of a *random variable*. Because each outcome x_i must be an element of S,

$$P(S) = \sum_{i=1}^{n} P(x_i) = 1$$

Also,

$$\sum_{x=a}^{b} f(x) = 1$$

where f is the *probability mass function* of a *discrete* random variable and $S = \{a, a+1, \cdots, b\}$ is its domain. Finally,

$$\int_{a}^{b} f(x)dx = 1$$

where f is the *probability density function* of a *continuous* random variable and $S = [a, b]$ is its domain.

For example, the sample space of rolling a six-sided die is

$S = \{1,2,3,4,5,6\}$ and, correspondingly, $\sum_{i=1}^{6} \frac{1}{6} = 1$.

sample variance
→ *variance*

sampling
The process of creating a *sample*. A sample is chosen in such a way that it resembles the distribution of a given measurable characteristic of a *population* as closely as possible. The most common method of sampling is "random sampling", where the *specimens* are chosen from the population randomly, for example by presenting a survey to students on a campus or by taking random numbers from a phone register.

Another approach is "systematic sampling", where the samples are chosen using some fixed pattern, like taking every 13^{th} piece from an assembly line. Another systematic approach is to perform a fixed number n of *experiments* and gather the *outcomes* in a sample. In this case the population is an infinite number of possible *trials*, and the sample is created by collecting the first n outcomes.

When some characteristics of the population are known in advance, "stratified sampling" can help to create a sample that reflects the population better than a pure random sample. For example, if a population is known to consist of 60% females and 40% males, the sample can be chosen from separate "female" and "male" strata, making sure that the sample contains the same proportion of males to females as the population.

sampling distribution
The *probability distribution* of a *population statistic* based on random *samples* represented by *random variables*. For finite, *discrete* populations, the sampling distribution can be approximated by the *frequency distribution* of the statistic computed for all possible samples of the size n. For example, all possible samples obtained by rolling a four-sided die two times (sample size $n = 2$) are shown in figure SD1, next to the corresponding sampling distribution.

X_i	$\bar{x}$	X_i	$\bar{x}$
$\{1,1\}$	1.0	$\{3,1\}$	2.0
$\{1,2\}$	1.5	$\{3,2\}$	2.5
$\{1,3\}$	2.0	$\{3,3\}$	3.0
$\{1,4\}$	2.5	$\{3,4\}$	3.5
$\{2,1\}$	1.5	$\{4,1\}$	2.5
$\{2,2\}$	2.0	$\{4,2\}$	3.0
$\{2,3\}$	2.5	$\{4,3\}$	3.5
$\{2,4\}$	3.0	$\{4,4\}$	4.0

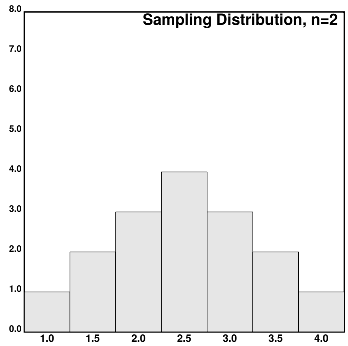


Figure **SD1:** sampling distribution of the mean of rolling a four-sided die $n = 2$ times; left: all possible samples of size $n = 2$ and corresponding means; right: frequency distribution of the sample means

With increasing sample size n, the frequency distribution of the statistic of X becomes asymptotically equal to the sampling distribution of the statistic. In the case of the *sampling distribution of the mean*, the frequency distribution of the mean becomes asymptotically normal (→ *normal distribution*) as n tends toward infinity. This is due to the *central limit theorem*. See figure SD2.

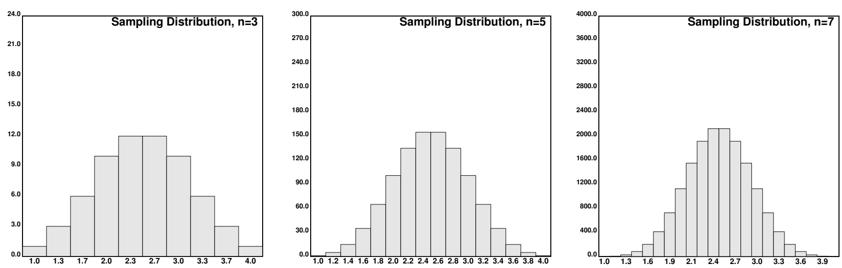


Figure **SD2:** sampling distribution of the mean for rolling a four-sided die; left: $n = 3$; middle: $n = 5$; right: $n = 7$; due to the central limit theorem, the frequency distribution of the mean becomes asymptotically normal

For samples taken from *continuous* populations, the above approach is not feasible, because there is an infinite number of sample points. See *central limit theorem* for a different approach.

Every statistic has a sampling distribution, but not all sampling distributions are normal. For instance, the sampling distribution of the variance has a $\chi^2(n)$ distribution (see *chi-square distribution*).

sampling distribution of the mean ($\bar{X}$)
The *sampling distribution* of the *(sample) mean*:

$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$ where n is the sample size.

I.e., the sampling distribution of the mean is a *normal distribution* with the same mean μ as the *population* from which the samples were drawn. Hence it can be used to create *confidence intervals* describing the probability of the true population mean falling within a certain range, even if a limited sample size and/or number of samples is available. $\bar{X}$ is normally distributed due to the *central limit theorem*.

scatter diagram
→ *scatter plot*

scatter plot
A way of visualizing the *correlation* of *random variables* by plotting the values of one variable as dots offset from the x-axis and the other one as dots offset from the y-axis. See figure SCP.

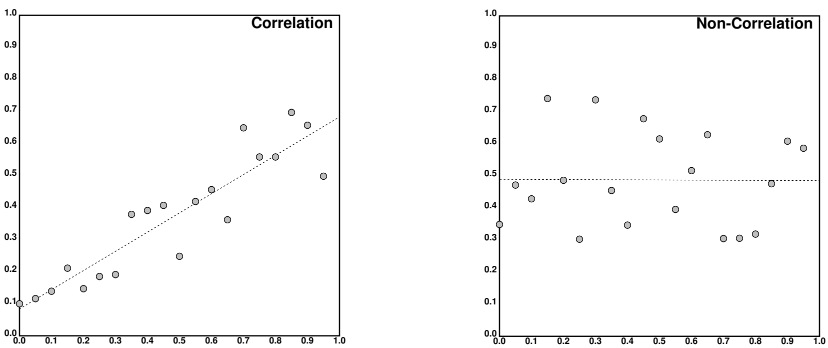


Figure **SCP:** scatter plots; left: correlated variables (correlation coefficient: $r = 0.59$); right: very weakly correlated variables (coefficient: $r = -0.0043$); dotted line = regression line

When the scatter plot shows some shape, like a diagonal line, a skewed triangle, or a parabola, there usually is more than a weak correlation. When the data points appear to be scattered randomly, the correlation is so weak that the variables can be assumed to be uncorrelated.

When drawing a *regression line* through a scatter plot, a horizontal line will indicate a weak correlation or no correlation at all, a line with an ascending *slope* indicates correlation, and a descending slope indicates *anticorrelation*.

score

A *statistic* that is the used as a parameter for a statistical function. See, for example: z*-score*, t*-score*, χ^2 *(chi-square) statistic*.

sd

→ *standard deviation*

significance

The result of a *hypothesis test* has significance (or "is significant") if it is sufficiently improbable to have arisen by chance while a *null hypothesis* holds. The threshold for significance is specified by choosing a *level of significance*. When the *probability* of the result to arise by chance is below the chosen significance level, the null hypothesis has to be rejected.

Given a level of significance of $\alpha = 0.1$ and improbability of a test result of $q = 0.91$, the result would be significant, because $q \geq 1 - \alpha$ or $1 - q \leq \alpha$. The value $P = 1 - q$ is also known as the *p-value* of the experiment.

significance level

→ *level of significance*

single-tailed

→ *one-tailed*

skewness (γ_1)

A measure for the skew of a curve, usually used to describe the *probability density functions* of *probability distributions*. A curve is left-skewed or has negative skewness, if its left *tail* is longer and/or heavier than its right tail. Subsequently, it is right-skewed or has positive skew, if its right tail is longer and/or heavier. There are closed forms for computing the exact skewness of specific probability distributions, but for a rough approximation, the formula $\mu - M$ can be used, where M is the *median* and μ the *mean* of a distribution. The sign of the result corresponds to the skewness of the curve.

The symbol γ_1 is often used to indicate skewness. Symmetric distributions, like the *normal distribution*, have a skewness of $\gamma_1 = 0$. Sample curves can be found in figure SKW.

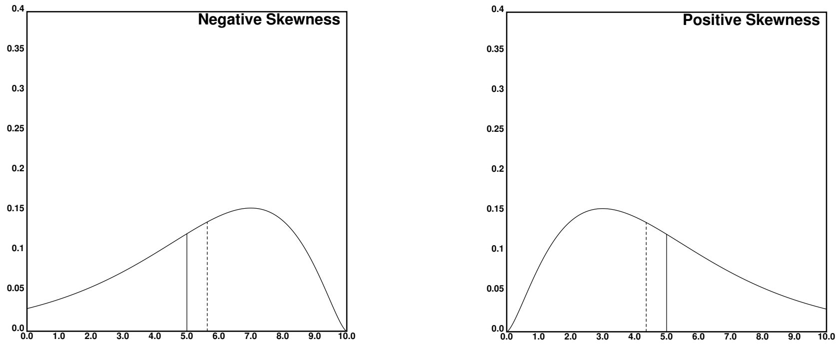


Figure **SKW:** skewness; left: negative skewness, skew to the left; right: positive skewness, skew to the right; in both panels the solid line indicates the mean, the dashed line the median.

slope

The steepness of the graph of a linear function of the form $y = ax + b$. When the coefficient a is zero, the graph will be a horizontal line. When the coefficient is positive, the line will ascend, i.e. the values of the function will grow, and when the coefficient is negative, the line will descend (the function values decrease). The slope of the function $y = ax + b$ is determined by the coefficient a alone while b determines the function's *intercept*. More generally, the slope can also be determined for a non-linear function, by finding its first derivative.

specimen

An individual, inseparable unit taken from a *population*, like a single inhabitant of planet Earth or a single book from a library.

standard deviation (σ, s, sd)

A measure of *dispersion* that is equal to the square root of the *variance*:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{1}{n} \sum (X - \mu)^2}$$

As with the variance, *Bessel's correction* is performed in the sd of a sample; it is defined as

$$s = \sqrt{s^2} = \sqrt{\frac{1}{n-1} \sum (x - \bar{x})^2}$$

The standard deviation plays a central role in the *normal* *distribution*. See also: *rule of three sigma*.

standard error (SE)
The *standard deviation* of the *sampling distribution* of a *statistic*. For example, the standard error of the *mean* is the standard deviation of the *sample mean*, $\bar{x}$:

$$SE_{\bar{x}} = \frac{s}{\sqrt{n}}$$

where s is the standard deviation of the sample and n is the sample size or number of samples. The term "standard error" is also used in statistical modeling, where it has a slightly different meaning; see *standard error of the estimate*.

standard error of the estimate (σ_{est})
A measure of the accuracy of *estimates* that is defined as follows:

$$\sigma_{est} = \sqrt{\frac{RSS}{n}} = \sqrt{\frac{\sum(y-\hat{y})^2}{n}}$$

where n is the number of estimate/observation pairs, $\hat{y}$ is an estimate, y is the corresponding *observation*, and RSS is the *residual sum of squares*. In a *linear regression model*, the standard error of the estimate specifies the *mean* distance between an observation and the *regression line*. When the standard error of a *regression line* is computed from a sample, two *Bessel's corrections* are applied, one for the *slope* and one for the *intercept*, resulting in

$$s_{est} = \sqrt{\frac{\sum(y-\hat{y})^2}{n-2}}$$

The standard error of the estimate, like the *standard deviation*, can be used to specify a *confidence interval* on the estimates. For example, an estimator with $\sigma_{est} \approx 172.59$ that predicts the income of a person will be off by no more than 172.59 currency units in 68.3% of the estimates and no more than $2\sigma_{est} = 345.18$ in 95.4% of the estimates (see *rule of three sigma*).

standard error of the slope coefficient ($SE_{\hat{\beta}}$)
The part of the *standard error* of a *regression line* that is based on the *slope* parameter alone. It is used, for instance, in the z*-test* and the t*-test for regression coefficients*. The standard error of the slope coefficient is computed as follows:

$$SE_{\hat{\beta}} = \frac{\sqrt{\frac{1}{n-2}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}}$$

Like in the *standard error of the estimate*, two *Bessel's corrections* are performed, one for the *slope* and one for the *intercept*.

standard normal distribution (Z-distribution, Z)
A normalized *continuous probability distribution*. A *random variable* $Z \sim N(0, 1)$ is standard normally distributed, if it follows a *normal distribution* with mean $\mu = 0$ and variance $\sigma^2 = 1$. The *cumulative distribution function* (CDF) of the Z-distribution maps z*-scores* to *quantiles*. The function graphs of the Z-distribution are shown in figure ZDC, its *distribution parameters* and functions are listed in figure ZDD.

$Z \sim N(0, 1)$	
PDF	$f(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$
CDF	$F(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{x} e^{-\frac{w^2}{2}} dw$
	$F(x) = \frac{1}{2}\left(1 + \frac{erf(x)}{\sqrt{2}}\right)$
Statistic	$x \in \mathbf{R}$: z-score
Parameters	none
μ	0
σ^2	1
Skewness (γ_1)	0

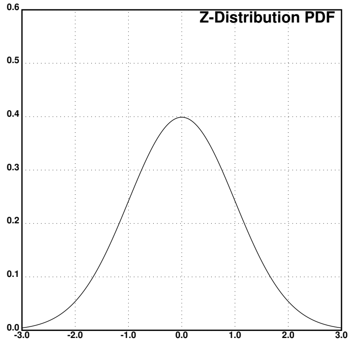
Figure **ZDD:** standard normal distribution; erf is the Gauss error function

Example: if the z-score of a metric, like body length, IQ, etc, is $+2.33\sigma$, then

$F_Z(2.33) \approx 0.99$

so that score is in the 0.99^{th} quantile, i.e. it is greater than (or equal to) 99% of all scores of the population. F_Z is the CDF of the standard normal distribution. See figure ZDD.

F_Z is often denoted using the letter ϕ, i.e. $\phi(x) = F_Z(x)$.



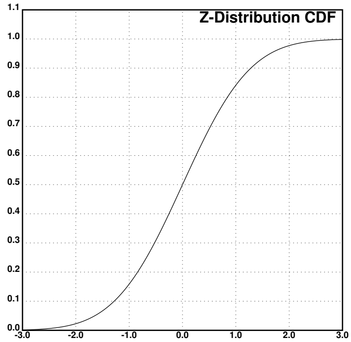


Figure **ZDC:** Z-distribution probability functions: left: PDF; right: CDF

statistic

A measure of some characteristic of a *sample*, *population*, or *random variable*. Also used to denote the function generating the statistic. Examples include the *mean*, the *variance* and *standard deviation*, and the *quantiles*, as well as *scores*, such as the *z-score*, *t-score*, or the χ^2 *statistic* (→ *chi-square statistic*).

statistics

A branch of mathematics that deals with the collection, analysis, interpretation, and visualization of data. Data is collected by *sampling*, which is done in *experiments*. Data analysis is done, for example, by *regression analysis*, *inference*, *hypothesis tests*, and modeling (→ *model*). The border between analysis (→ *hypothesis test*) and interpretation is fluid. Visualization typically uses diagrams to represent data in a way that is intuitively accessible; see, for instance, *probability distribution*, *regression line*, *scatter plot*, and *decision tree*.

stratified sampling

→ *sampling*

Student's t-distribution

→ *t-distribution*

success

→ *Bernoulli trial*

sum of squared error (SSE)
→ *residual sum of squares*

systematic sampling
→ *sampling*

T

t-distribution ($X \sim t(\nu)$)
A *continuous probability distribution* that generalizes the *standard normal distribution* when the *sample* size is small. With growing sample size, the t-distribution becomes asymptotically standard normal. A *random variable* $X \sim t(\nu)$ is t-distributed with ν (nu) *degrees of freedom*. Sometimes the letter T is used to denote t-distributed variables, sometimes the number of degrees of freedom is attached as a subscript: $T_\nu \sim t(\nu)$. The t-distribution is used for various *estimates* and *hypothesis tests*, like location tests and correlation tests; see *t-test for location* and *t-test for regression coefficients*, respectively. The parameters and functions of the t-distribution are shown in figure TDD, sample plots of its probability functions in figure TDC.

$X \sim t(\nu)$	
PDF	$f(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi\nu} \cdot \Gamma(\frac{\nu}{2})} \cdot \left(1 + \frac{x^2}{2}\right)^{-\frac{\nu+1}{2}}$
CDF	$F(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi\nu} \cdot \Gamma(\frac{\nu}{2})} \cdot \int_{-\infty}^{x} \left(1 + \frac{w^2}{2}\right)^{-\frac{\nu+1}{2}} dw$
	$F(x) = 1 - \frac{1}{2} I_{\frac{\nu}{x^2+\nu}}(\frac{\nu}{2}, \frac{1}{2})$
Statistic	$x \in \mathbf{R}$: t statistic
Parameters	$\nu \in \mathbf{R}^+$: degrees of freedom
μ	0
σ^2	$\begin{cases} \frac{\nu}{\nu-2} & for\ \nu > 2 \\ \infty & for\ 1 < \nu \leq 2 \end{cases}$
Skewness (γ_1)	0
Approximations	$N(0, 1)$ for $\nu \geq 30$

Figure **TDD:** t-distribution; $\Gamma(x)$ is the complete gamma function, $I_x(a, b)$ is the regularized incomplete beta function

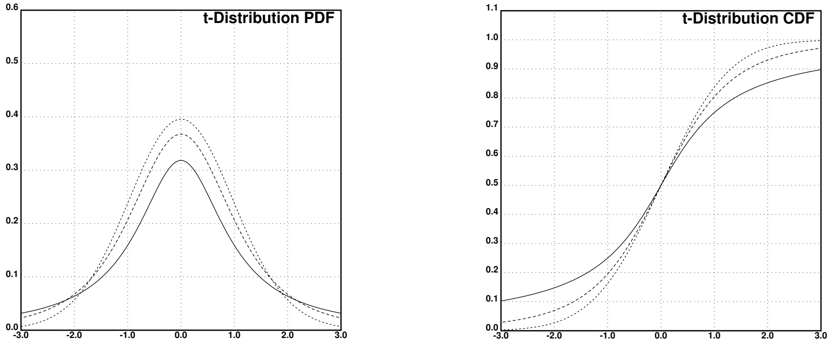


Figure **TDC:** t-distribution probability functions; left PDF with $\nu = 1$ (solid), $\nu = 3$ (dashed), and $\nu = 30$ (dotted); right: CDF with same parameters

An *interval estimator* predicting the *mean* of a *population* given a small sample size of typically less than 30 *specimens* can be created on the basis of the t-distribution. This approach assumes that the population is approximately normally distributed. The estimator is created as follows. The *sampling distribution of the mean* $\bar{X}$ of the given sample is t-distributed with $n-1$ degrees of freedom:

$$\frac{\bar{X}}{\sqrt{s^2/n}} \sim t(n-1)$$

where n is the sample size and s^2 is the sample variance. So the population mean μ falls within the *confidence interval*

$$P(\bar{x} - t \cdot \sqrt{s^2/n} \;\leq\; \bar{X} \;\leq\; \bar{x} + t \cdot \sqrt{s^2/n}) \;=\; c$$

where $\bar{x}$ is the sample mean and $t = F_{t(n-1)}(c + \frac{1-c}{2})$ is the *quantile* of the t-distribution given a probability of c plus half the probability associated with a *critical region* (because the confidence interval is *two-tailed*).

Example: There is a sample of 10 marbles with an average diameter of $\bar{x} = 15mm$ and a variance of $s^2 = 2.5$. Then,

$$\frac{\bar{X}}{\sqrt{2.5/10}} \;=\; \frac{\bar{X}}{0.5} \;\sim\; t(9)$$

and, given a *level of confidence* of $c = 0.9$:

$$F_{t(9)}(0.95) \;\approx\; 1.83$$

and therefore,

$$P(15 - 1.83 \cdot 0.5 \leq \bar{X} \leq 15 + 1.83 \cdot 0.5)$$
$$= P(14.085 \leq \bar{X} \leq 15.915) \approx 0.9$$

I.e. the probability for the population mean μ to fall within the interval $[14.085, 15.915]$ would be $p = 0.9$.

t-score (t)
A t-distributed *test score* (→ *t-distribution*) that expresses the distance from an observed (→ *observation*) *sample mean* $\bar{x}$ to a known (pre-existing or *population*) mean μ_0 given a small ($n < 30$) sample taken from an approximately normal population (→ *normal distribution*). It is computed as follows:

$$t = \frac{\bar{x} - \mu_0}{\sqrt{s^2/n}} = \frac{\bar{x} - \mu_0}{s\sqrt{n}}$$

where n is the size and s^2 the *variance* of the sample. For larger values of n, the t-score is asymptotically equal to the *z-score*.

The score described here is used in a *t-test for location*. The name "t-score" is also used to denote the test score in a *t-test for regression coefficients*.

t-test for location
A *hypothesis test* used to find out whether a *sample mean* $\bar{x}$ differs significantly (→ *significance*) from a known mean μ_0. It uses the *t-distribution* to compute the improbability of the *observation* $\bar{x}$. The t-test for location is used when the sample size n is small, typically $n < 30$, and the population is approximately normally distributed. For larger sample sizes, the z-test can be used instead.

The *score* for the test and the test procedure are the same as for the *z-test for location*, but a t-distribution with ν (nu) *degrees of freedom* is used instead of the *standard normal distribution* to compute the *quantile* of the score. The value of ν is $n - 1$, the sample size minus one.

Example: A new technology is supposed to decrease the fault rate of a machine. The original fault rate is μ_0 = 120 failures per year, and a sample of ten new machines showed an average fault rate of $\bar{x} = 104$ with a variance of $s^2 = 360$. The *null hypothesis* H_0 is

$\bar{x} \geq \mu_0$ (the new technology has no effect) and the chosen *level of significance* is $\alpha = 0.05$. Then the t*-score* is

$$t = \frac{\bar{x} - \mu_0}{\sqrt{s^2/n}} = \frac{104 - 120}{\sqrt{360/10}} = \frac{-16}{6} \approx -2.667$$

and follows a t-distribution $T_9 \sim t(9)$. The quantile associated with the t-score is

$$F_{t(9)}(-2.667) \approx 0.987$$

where $F_{t(9)}$ is the *cumulative distribution function* of the t-distribution with nine degrees of freedom. Due to the chosen null hypothesis, this is a *one-tailed* test and the result falls within the *critical region* $P(T_9 \leq t) = \alpha$, so H_0 can be rejected. See z*-test for location* for further details.

t-test for regression coefficients

A *two-tailed hypothesis test* used to find out whether the *slope* $\hat{\beta}$ of a *regression line* differs significantly ($\rightarrow$ *significance*) from a given slope β_0. It uses the t*-distribution* to compute the *improbability* of the *observation* $\hat{\beta}$ given an assumed slope β_0. The *test score* and procedure of the test are the same as in a z*-test for regression coefficients*, but the test score follows a t-distribution with ν (nu) degrees of freedom instead of a *standard normal distribution*. The number of degrees of freedom is $n-1$, the number of observed *data points* minus one.

For example, given a slope of $\hat{\beta} = 0.4$ and a *standard error of the slope coefficient* of $SE_{\hat{\beta}} = 0.25$, calculated from the data points of 7 estimates, the improbability of observing the given slope under the *null hypothesis* of $\beta_0 = 0$ would be $F_{t(6)}(t)$, where $F_{t(6)}$ is the *cumulative distribution function* of the t-distribution with $\nu = 6$ and

$$t = \frac{\hat{\beta} - \beta_0}{SE_{\hat{\beta}}} = \frac{0.4 - 0}{0.25} = 1.6$$

The improbability of observing the slope is then $F_{t(6)}(1.6) \approx 0.92$. The significance of this result depends on the chosen *level of significance*.

tail

The flat end of a curve, mostly the curve of the *probability density function* of a *probability distribution*. A *one-tailed* curve tends toward zero at one end of the curve and a *two-tailed* one tends toward zero at both, opposite, ends of the curve. See figure TLS for examples. A "heavier" tail (i.e. one that does not tend toward zero quickly) indicates a larger *variance* of the underlying distribution.

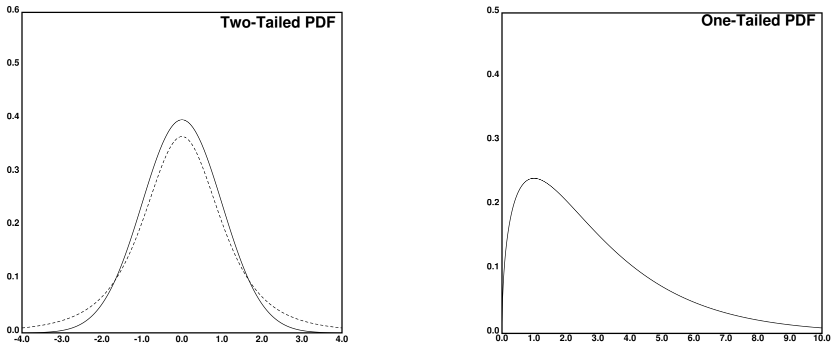


Figure **TLS:** tails of curves; left: two-tailed distributions: standard normal distribution (solid, thin-tailed) and t-distribution with three degrees of freedom (dashed, heavy-tailed); right: one-tailed distribution: χ^2-distribution with three degrees of freedom, tail to the right

test score
→ *test statistic*, *score*

test error
The *error* of a *model* when predicting *data points*. This error can only be approximated because when a model is being fitted, no predictions have been made and so no observed (→ *observation*) data points are available for reference. A common way to approximate the test error is the *cross validation error* (CVE).

test statistic
A *statistic* that is used in *hypothesis tests*.

three sigma
→ *rule of three sigma*

total sum of squares (TSS)
A measure of *dispersion* formed by the sum of squared differences between observed *outcomes* and their *mean*:

$$TSS = \sum_{i=1}^{n} (y - \bar{y})^2$$

Used in the evaluation of statistical *models*. See, for example, *coefficient of determination* (including figure).

training error
A synonym for *error*, mostly the *mean squared error* (MSE). The term "training error" is often used to indicate that an error does not indicate the accuracy of predictions computed by a *model*. Unlike the training error, the *test error* does provide a measure of accuracy of predictions. See *cross validation error*.

trial
An action that may or may not result in a given *observation*. For example, a trial could be rolling a die and the corresponding observation could be "a face with an even number of eyes pointing up". A trial with only two possible *outcomes* is called a *Bernoulli trial*.

two-tailed (adj)
A *probability distribution* with two *tails*. More precisely, a distribution whose *probability density function* plots a curve with two tails at opposite ends of the curve. In a *hypothesis test*, a test with two *critical regions*; see z*-test for location*.

type I error
The rejection of a valid *null hypothesis* in a *hypothesis test*. Also known as a "false positive" or an effect without a cause. The *probability* of a test making a type I error is denoted by the letter α. See *type II error* for an illustration.

type II error
Failure to reject an invalid *null hypothesis* H_0 in a *hypothesis test*. Also known as a "false negative" or a cause without effect. The *probability* of a test making a type II error is denoted by the letter β. See figure ER1 for a table and figure ER2 for a *decision tree* explaining the types of errors. Given diagram ER2, the *conditional probabilities* of the type I and type II error are

$$\alpha = P(H_0 \text{ rejected} \mid H_0 \text{ valid})$$
$$\beta = P(H_0 \text{ accepted} \mid H_0 \text{ invalid})$$

The *complement* of the type II error specifies the *power of a test*.

	H_0 accepted	H_0 rejected
H_0 valid		Type I Error
H_0 invalid	Type II Error	

Figure **ER1:** type I and II errors

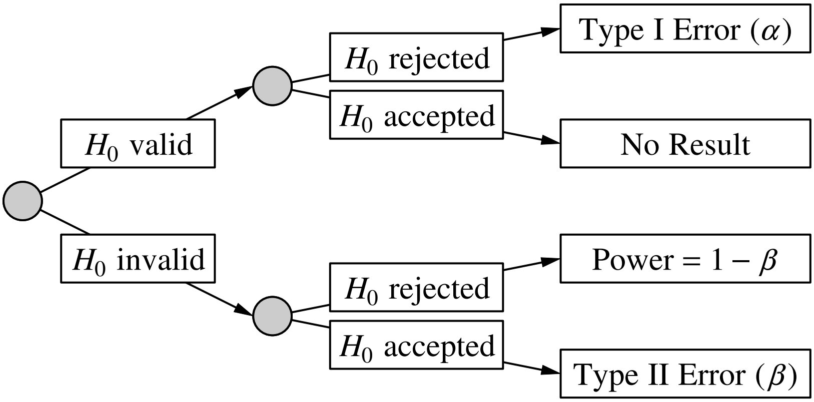


Figure **ER2:** decision tree: type I and II errors

U

unbiased
→ *bias*

uniform distribution
→ *discrete uniform distribution*, *continuous uniform distribution*

union
The union of two *events* A and B is the event of at least one of A and B occurring at a given time. The *probability* of the union is:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

where $P(A \cap B)$ denotes the *intersection* of the events A and B. The intersection has to be subtracted from the sum of $P(A)$ and $P(B)$, because it would be duplicated otherwise. For instance, the probability of drawing a red card or an ace from a standard deck of cards would be

$$\begin{aligned} P(Red \cup Ace) &= P(Red) + P(Ace) - P(Red \cap Ace) \\ &= \frac{1}{2} + \frac{1}{8} - \frac{1}{16} = \frac{9}{16} = 0.5625 \end{aligned}$$

Because there are two aces contributing to $P(Red)$ and two red cards contributing to $P(Ace)$, the red aces would be counted twice, hence $P(Red \cap Ace)$ has to be subtracted once.

When the events A and B are mutually exclusive, i.e. they cannot occur at the same time, then the union of A and B reduces to

$$P(A \cup B) = P(A) + P(B)$$

For instance, the probability of drawing a king or a queen from a deck is $\frac{1}{8} + \frac{1}{8}$, because it is impossible to draw a card that is a king and queen at the same time.

V,W,X,Y

variability
→ *dispersion*

variance (σ^2, s^2, $var(X)$)
A measure of *dispersion* that is equal to the mean squared distance of a set of *data points* from their *mean*:

$$\sigma^2 = \frac{1}{n}\sum(X-\mu)^2$$

Distances are squared so that positive and negative values cannot cancel out each other. For example, the variance of the data set

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

with mean $\mu = 3.5$ equals

$$\frac{(1-3.5)^2+\cdots+(6-3.5)^2}{6} = 2.91\bar{6}$$

The notation σ^2 indicates the *population* variance, while s^2 denotes the *sample* variance. The sample variance is defined as:

$$s^2 = \frac{1}{n-1}\sum(x-\bar{x})^2$$

where the use of the factor of $\frac{1}{n-1}$ instead of $\frac{1}{n}$ is known as *Bessel's correction*. A more intuitive measure of dispersion is the *standard deviation*. *Probability distributions* have closed forms for computing their variance.

Venn diagram
A diagram used to visualize operations on sets of *events*. A Venn diagram consists of a rectangle containing one or more ellipses representing sets. Diagrams containing two or three sets are particularly common. The rectangle represents the *sample space* of the diagram, the space where ellipses overlap indicates the *intersection* of two sets. Shading or color is used to indicate other operations, like the set difference or the *union* of two sets. See

figure VED for a sample Venn diagram.

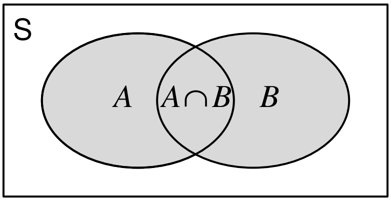


Figure **VED:** Venn diagram; left ellipsis: set A, right ellipsis: set B; area where ellipses overlap: intersection $A \cap B$ of A and B; shaded area: union $A \cup B$ of A and B; white area: complement of $A \cup B$; complete area of the rectangle: sample space S

Z

Z-distribution (Z)
→ *standard normal distribution*

z-score (z)
A normalized *score* that is derived from a (raw, unnormalized) value of a normally distributed *random variable* (→ *normal distribution*). Given a raw score x of a normally distributed random variable $X \sim N(\mu, \sigma^2)$, the z-score z is computed as follows:

$$z = \frac{x - \mu}{\sqrt{\sigma^2}} = \frac{x - \mu}{\sigma}$$

The z-score is used to find the *quantile* of the corresponding raw score. Figure ZST shows the z-scores for some common *levels of significance* and *levels of confidence*.

c	0.9	0.95	0.975	0.99	0.999
α	0.1	0.05	0.025	0.01	0.001
one-tailed z	1.28	1.64	1.96	2.33	3.09
two-tailed z	1.64	1.96	2.24	2.58	3.29

Figure **ZST:** common z-scores for one-tailed and two-tailed intervals and tests; $c =$ level of confidence; $\alpha =$ level of significance

The name "z-score" is also used to denote the *test score* in a *z-test for location* and a *z-test for regression coefficients*.

z-test for location
A *hypothesis test* used to find out whether a *sample mean* $\bar{x}$ differs significantly (→ *significance*) from a known mean μ_0. It uses the *standard normal distribution* to compute the improbability of the *observation* $\bar{x}$. When the sample size is small ($n < 30$), the *t-test for location* should be used instead. The *test statistic* ("z-score") used in the z-test is computed as follows:

$$z = \frac{\bar{x} - \mu_0}{SE} = \frac{\bar{x} - \mu_0}{\sqrt{s^2/n}}$$

where s^2 is the *sample variance*, n is the sample size, and $SE = \sqrt{s^2/n}$ is the *standard error* of the sample.

The type of the z-test depends on its *null hypothesis*, H_0. When H_0 states $\bar{x} = \mu_0$, then the test is *two-tailed*, otherwise, when H_0 states either $\bar{x} \leq \mu_0$ or $\bar{x} \geq \mu_0$, it is one-tailed. The corresponding *critical regions* are shown in table ZTC.

H_0	Critical Region	Tail
$\bar{x} = \mu_0$	$P(Z \leq -\lvert z\rvert) = \frac{\alpha}{2}$ and $P(Z \geq \lvert z\rvert) = \frac{\alpha}{2}$	Both
$\bar{x} \leq \mu_0$	$P(Z \geq z) = \alpha$	Right
$\bar{x} \geq \mu_0$	$P(Z \leq z) = \alpha$	Left

Figure **ZTC:** critical regions of the z-test; α is the desired level of significance

The z-score (in σ; see *standard deviation*) indicates the improbability of the observed mean given the null hypothesis. It can be converted to a *quantile* using the *cumulative distribution function* F_Z of the standard normal distribution.

Example: A training program is supposed to boost the IQ of its participants. The average IQ is $\mu_0 = 100$, the mean IQ taken from a sample of 50 participants is $\bar{x} = 105$ with a variance of $s^2 = 450$, and the chosen *level of significance* is $\alpha = 0.05$. The test score is then

$$z = \frac{105 - 100}{\sqrt{450/50}} = \frac{5}{3} \approx 1.667$$

Because in this case the null hypothesis H_0 would be that the IQ has not increased (i.e. $\bar{x} \leq \mu_0$), the corresponding critical region would start at 0.95. The quantile of 1.667 is $F_Z(1.667) \approx 0.952$, so the result is (barely) within the critical region.

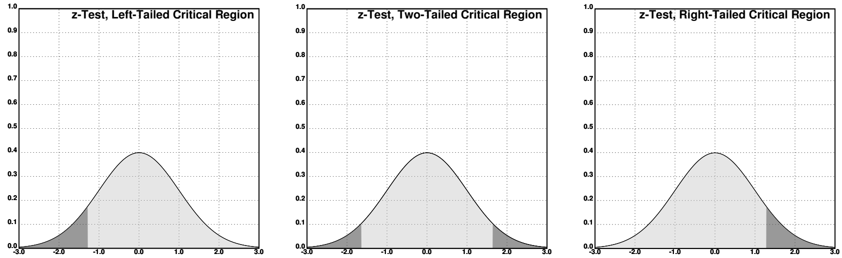


Figure **CRZ:** critical regions of the z-test Left: critical region in the left tail; middle: one critical region in each tail; right: critical region in the right tail

z-test for regression coefficients

A *two-tailed hypothesis test* used to find out whether the *slope* $\hat{\beta}$ of a *regression line* differs significantly ($\rightarrow$ *significance*) from a given slope β_0. It uses the *standard normal distribution* to compute the *improbability* of the *observation* $\hat{\beta}$ given an assumed slope β_0.

Often β_0 is assumed to be zero. In this case the *null hypothesis* H_0 of the test would be that the variables of the underlying *linear regression* are independent. The z*-score* of the z-test for regression coefficients is

$$z = \frac{\hat{\beta} - \beta_0}{SE_{\hat{\beta}}}$$

where $SE_{\hat{\beta}}$ is the *standard error of the slope coefficient* $\hat{\beta}$. When the number of *data points* for computing the standard error is smaller than 30, a t*-test for regression coefficients* should be used instead of a z-test.

The test itself follows the same procedure as a two-tailed z*-test for location*. For example, given a slope of $\hat{\beta} = 0.4$ and a standard error of $SE_{\hat{\beta}} = 0.25$, the *improbability* of observing the given slope under the *null hypothesis* of $\beta_0 = 0$ would be $F_Z(z)$, where F_Z is the *cumulative distribution function* of the standard normal distribution and

$$z = \frac{0.4 - 0}{0.25} = 1.6$$

by inserting values into the above formula. The improbability of observing the slope is then $F_Z(1.6) \approx 0.95$. The significance of this result depends on the chosen *level of significance*.

Appendix: Mathematical Notation

iff	Short for "if, and only if" or "exactly if"
$\{a, b, c\}$	Set containing the members a, b, c
$x \in \{a, b, c\}$	x is an element of the set $\{a, b, c\}$
$x \in \mathbf{N}_0$	x is an element of the set of natural numbers, including zero: $\{0, 1, 2, \cdots\}$
$x \in \mathbf{N}^+$	x is an element of the set of positive natural numbers $\{1, 2, \cdots\}$
$x \in \mathbf{Z}$	x is an element of the set of integer numbers $\{\cdots, -2, -1, 0, 1, 2, \cdots\}$
$x \in \mathbf{R}$	x is an element of the set of real numbers
$-\infty$, ∞	Negative infinity, infinity
$X \sim D(x, \cdots)$	X follows the distribution D with parameters $x, \cdots$
$[a, b]$	Interval covering the range from a to b, including both a and b
(a, b)	Interval covering the range from a to b, excluding both a and b
xy, $x \cdot y$	Product of x and y
$max(X)$	Greatest value (maximum) of X
$min(X)$	Least value (minimum) of X
$n!$	n factorial, i.e. $1 \cdot 2 \cdot \ \cdots \ \cdot n$
$\lfloor x \rfloor$	Floor of x, i.e. the greatest $y \in \mathbf{Z}$ that is not greater than x
$\lvert x \rvert$	Magnitude of x, i.e. the distance between x and 0, or x with any negative sign removed
$\ln(x)$	Natural logarithm y of x, where $x = e^y$
$\bar{A}$, A'	Complement of A

$P(X)$	Probability of the event X
$P(X = x)$	Probability of the random variable X taking the value x
$P(X \leq x)$	Probability of the random variable X taking a value not greater than x
$P(B \vert A)$	(Conditional) probability of B given A
$A \cap B$	Intersection or conjunction (AND) of A and B
$A \cup B$	Union or disjunction (OR) of A and B
$\bigcup_{i=1}^{n} x_i$	$x_1 \cup \cdots \cup x_n$
$\sum_{i=1}^{n} x_i$	$x_1 + \cdots + x_n$
$\prod_{i=1}^{n} x_i$	$x_1 \cdot \; \cdots \; \cdot x_n$
$\int_a^b f(x)dx$	Integral of $f(x)$, the area under the curve $f(x)$ from a to b
$\lim_{n \to m} x$	Limit of x as n goes toward m
${}_nP_k$	k-permutations of a set of n elements
${}_nC_k$	k-combinations of a set of n elements ("n choose k")
${}_qQ_n$	The n^{th} q-quantile
$\binom{n}{k}$	binomial coefficient, ${}_nC_k$
$f_X(x), f(x)$	Probability mass function (discrete) or probability density function (continuous) of X
$F_X(x), F(x)$	Cumulative distribution function of X
$F_X^{-1}(x)$ $F^{-1}(x)$	Quantile function of X
μ_X, μ	Mean of X

$E(X)$	Expectation of X
$\bar{x}$	Mean of a sample $x_1, \cdots$
s_x^2	Sample variance of sample x
$s_{x,y}$	Covariance of samples x and y
$\hat{\mu}$	Point estimator for the population mean
$\hat{\sigma}^2$	Point estimator for the population variance
$\bar{x}$	Sample mean
$\bar{X}$	Sampling distribution of the (sample) mean
σ_X, σ	Standard deviation of X
σ_{est}	Standard error of the estimate
σ_X^2, σ^2 $var(X)$	Variance of X
$\phi(x)$	Cumulative distribution function of the standard normal distribution
$\phi^{-1}(x)$	Quantile function of the standard normal distribution
$erf(x)$	Gauss error function
$\Gamma(x)$	(Complete) gamma function
$\gamma(a, x)$	Lower incomplete gamma function
$\Gamma(a, x)$	Upper incomplete gamma function
$P(a, x)$	Regularized incomplete gamma function
$Q(a, x)$	(Complement of the) regularized incomplete gamma function
$\mathrm{B}(a, b)$	(Complete) beta function
$\mathrm{B}_x(a, b)$	Incomplete beta function
$I_x(a, b)$	Regularized incomplete beta function

Appendix: Special Functions

Gauss Error Function

$$erf(x) = \frac{1}{\sqrt{\pi}} \int_{-x}^{x} e^{-w^2} dw$$

Gamma Function

$\Gamma(x) = (x-1)!$ for $x \in \mathbf{N}^+$

$$\Gamma(x) = \int_{0}^{\infty} e^{-w} w^{x-1} dw$$

Upper incomplete Γ function

$$\Gamma(a, x) = \frac{1}{\Gamma(a)} \int_{x}^{\infty} e^{-w} w^{a-1} dw$$

Lower incomplete Γ function

$$\gamma(a, x) = \frac{1}{\Gamma(a)} \int_{0}^{x} e^{-w} w^{a-1} dw$$

Regularized incomplete Γ function

$$P(a, x) = \frac{\gamma(a, x)}{\Gamma(a)}$$

$Q(a, x) = 1 - P(a, x)$

Relationships

$\Gamma(a) = \Gamma(a, 0)$

$\Gamma(a) = \gamma(a, x) + \Gamma(a, x)$

$$P(a, x) = \begin{cases} \dfrac{\gamma(a, x)}{\Gamma(a)} & \text{if } x < a + 1 \\ 1 - \dfrac{\Gamma(a, x)}{\Gamma(a)} & \text{if } x \geq a + 1 \end{cases}$$

Beta Function

$$\mathrm{B}(a, b) = \int_0^1 w^{a-1} \cdot (1-w)^{b-1} dw$$

Incomplete B function

$$\mathrm{B}_x(a, b) = \int_0^x w^{a-1} \cdot (1-w)^{b-1} dw$$

Regularized incomplete B function

$$I_x(a, b) = \frac{1}{\mathrm{B}(a, b)} \cdot \int_0^x w^{a-1} \cdot (1-w)^{b-1} dw$$

Relationships

$$\mathrm{B}(a, b) = \mathrm{B}_1(a, b) = I_1(a, b) = \frac{\Gamma(a) + \Gamma(b)}{\Gamma(a+b)}$$

$$I_x(a, b) = \frac{\mathrm{B}_x(a, b)}{\mathrm{B}(a, b)}$$

Appendix: Probability Tables

PDF of the Standard Normal Distribution

$\pm\sigma$	±0.00	±0.01	±0.02	±0.03	±0.04	±0.05	±0.06	±0.07	±0.08	±0.09
0.0	.3989	.3989	.3989	.3988	.3986	.3984	.3982	.3980	.3977	.3973
0.1	.3970	.3965	.3961	.3956	.3951	.3945	.3939	.3932	.3925	.3918
0.2	.3910	.3902	.3894	.3885	.3876	.3867	.3857	.3847	.3836	.3825
0.3	.3814	.3802	.3790	.3778	.3765	.3752	.3739	.3725	.3712	.3697
0.4	.3683	.3668	.3653	.3637	.3621	.3605	.3589	.3572	.3555	.3538
0.5	.3521	.3503	.3485	.3467	.3448	.3429	.3410	.3391	.3372	.3352
0.6	.3332	.3312	.3292	.3271	.3251	.3230	.3209	.3187	.3166	.3144
0.7	.3123	.3101	.3079	.3056	.3034	.3011	.2989	.2966	.2943	.2920
0.8	.2897	.2874	.2850	.2827	.2803	.2780	.2756	.2732	.2709	.2685
0.9	.2661	.2637	.2613	.2589	.2565	.2541	.2516	.2492	.2468	.2444
1.0	.2420	.2396	.2371	.2347	.2323	.2299	.2275	.2251	.2227	.2203
1.1	.2179	.2155	.2131	.2107	.2083	.2059	.2036	.2012	.1989	.1965
1.2	.1942	.1919	.1895	.1872	.1849	.1826	.1804	.1781	.1758	.1736
1.3	.1714	.1691	.1669	.1647	.1626	.1604	.1582	.1561	.1539	.1518
1.4	.1497	.1476	.1456	.1435	.1415	.1394	.1374	.1354	.1334	.1315
1.5	.1295	.1276	.1257	.1238	.1219	.1200	.1182	.1163	.1145	.1127
1.6	.1109	.1092	.1074	.1057	.1040	.1023	.1006	.0989	.0973	.0957
1.7	.0940	.0925	.0909	.0893	.0878	.0863	.0848	.0833	.0818	.0804
1.8	.0790	.0775	.0761	.0748	.0734	.0721	.0707	.0694	.0681	.0669
1.9	.0656	.0644	.0632	.0620	.0608	.0596	.0584	.0573	.0562	.0551
2.0	.0540	.0529	.0519	.0508	.0498	.0488	.0478	.0468	.0459	.0449
2.1	.0440	.0431	.0422	.0413	.0404	.0396	.0387	.0379	.0371	.0363
2.2	.0355	.0347	.0339	.0332	.0325	.0317	.0310	.0303	.0297	.0290
2.3	.0283	.0277	.0270	.0264	.0258	.0252	.0246	.0241	.0235	.0229
2.4	.0224	.0219	.0213	.0208	.0203	.0198	.0194	.0189	.0184	.0180
2.5	.0175	.0171	.0167	.0163	.0158	.0154	.0151	.0147	.0143	.0139
2.6	.0136	.0132	.0129	.0126	.0122	.0119	.0116	.0113	.0110	.0107
2.7	.0104	.0101	.0099	.0096	.0093	.0091	.0088	.0086	.0084	.0081
2.8	.0079	.0077	.0075	.0073	.0071	.0069	.0067	.0065	.0063	.0061
2.9	.0060	.0058	.0056	.0055	.0053	.0051	.0050	.0048	.0047	.0046
3.0	.0044	.0043	.0042	.0040	.0039	.0038	.0037	.0036	.0035	.0034

CDF of the Standard Normal Distribution

$-\sigma$	–0.09	–0.08	–0.07	–0.06	–0.05	–0.04	–0.03	–0.02	–0.01	–0.00
3.0	.0010	.0010	.0011	.0011	.0011	.0012	.0012	.0013	.0013	.0013
2.9	.0014	.0014	.0015	.0015	.0016	.0016	.0017	.0018	.0018	.0019
2.8	.0019	.0020	.0021	.0021	.0022	.0023	.0023	.0024	.0025	.0026
2.7	.0026	.0027	.0028	.0029	.0030	.0031	.0032	.0033	.0034	.0035
2.6	.0036	.0037	.0038	.0039	.0040	.0041	.0043	.0044	.0045	.0047
2.5	.0048	.0049	.0051	.0052	.0054	.0055	.0057	.0059	.0060	.0062
2.4	.0064	.0066	.0068	.0069	.0071	.0073	.0075	.0078	.0080	.0082
2.3	.0084	.0087	.0089	.0091	.0094	.0096	.0099	.0102	.0104	.0107
2.2	.0110	.0113	.0116	.0119	.0122	.0125	.0129	.0132	.0136	.0139
2.1	.0143	.0146	.0150	.0154	.0158	.0162	.0166	.0170	.0174	.0179
2.0	.0183	.0188	.0192	.0197	.0202	.0207	.0212	.0217	.0222	.0228
1.9	.0233	.0239	.0244	.0250	.0256	.0262	.0268	.0274	.0281	.0287
1.8	.0294	.0301	.0307	.0314	.0322	.0329	.0336	.0344	.0351	.0359
1.7	.0367	.0375	.0384	.0392	.0401	.0409	.0418	.0427	.0436	.0446
1.6	.0455	.0465	.0475	.0485	.0495	.0505	.0516	.0526	.0537	.0548
1.5	.0559	.0571	.0582	.0594	.0606	.0618	.0630	.0643	.0655	.0668
1.4	.0681	.0694	.0708	.0721	.0735	.0749	.0764	.0778	.0793	.0808
1.3	.0823	.0838	.0853	.0869	.0885	.0901	.0918	.0934	.0951	.0968
1.2	.0985	.1003	.1020	.1038	.1056	.1075	.1093	.1112	.1131	.1151
1.1	.1170	.1190	.1210	.1230	.1251	.1271	.1292	.1314	.1335	.1357
1.0	.1379	.1401	.1423	.1446	.1469	.1492	.1515	.1539	.1562	.1587
0.9	.1611	.1635	.1660	.1685	.1711	.1736	.1762	.1788	.1814	.1841
0.8	.1867	.1894	.1922	.1949	.1977	.2005	.2033	.2061	.2090	.2119
0.7	.2148	.2177	.2207	.2236	.2266	.2297	.2327	.2358	.2389	.2420
0.6	.2451	.2483	.2514	.2546	.2578	.2611	.2643	.2676	.2709	.2743
0.5	.2776	.2810	.2843	.2877	.2912	.2946	.2981	.3015	.3050	.3085
0.4	.3121	.3156	.3192	.3228	.3264	.3300	.3336	.3372	.3409	.3446
0.3	.3483	.3520	.3557	.3594	.3632	.3669	.3707	.3745	.3783	.3821
0.2	.3859	.3897	.3936	.3974	.4013	.4052	.4090	.4129	.4168	.4207
0.1	.4247	.4286	.4325	.4364	.4404	.4443	.4483	.4522	.4562	.4602
0.0	.4641	.4681	.4721	.4761	.4801	.4840	.4880	.4920	.4960	.5000

CDF of the Standard Normal Distribution

σ	+0.00	+0.01	+0.02	+0.03	+0.04	+0.05	+0.06	+0.07	+0.08	+0.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7703	.7734	.7764	.7793	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990

Quantiles of the χ^2_ν-Distribution

q / ν	1	2	3	4	5	6	7	8	9	10
.0010	0	.0020	.0243	.0908	.2102	.3811	.5985	.8571	1.152	1.479
.0025	0	.0050	.0449	.1449	.3075	.5266	.7945	1.104	1.450	1.827
.0050	0	.0100	.0717	.2070	.4117	.6757	.9893	1.344	1.735	2.156
.0100	.0002	.0201	.1148	.2971	.5543	.8721	1.239	1.646	2.088	2.558
.0150	.0004	.0302	.1516	.3682	.6618	1.016	1.418	1.860	2.335	2.837
.0200	.0006	.0404	.1848	.4294	.7519	1.134	1.564	2.032	2.532	3.059
.0250	.0010	.0506	.2158	.4844	.8312	1.237	1.690	2.180	2.700	3.247
.0300	.0014	.0609	.2451	.5351	.9031	1.330	1.802	2.310	2.848	3.412
.0400	.0025	.0816	.3001	.6271	1.031	1.492	1.997	2.537	3.105	3.697
.0500	.0039	.1026	.3518	.7107	1.145	1.635	2.167	2.733	3.325	3.940
.1000	.0158	.2107	.5844	1.064	1.610	2.204	2.833	3.490	4.168	4.865
.1500	.0358	.3250	.7978	1.366	1.994	2.661	3.358	4.078	4.817	5.570
.2000	.0642	.4463	1.005	1.649	2.343	3.070	3.822	4.594	5.380	6.179
.2500	.1015	.5754	1.213	1.923	2.675	3.455	4.255	5.071	5.899	6.737
.3000	.1485	.7133	1.424	2.195	3.000	3.828	4.671	5.527	6.393	7.267
.3500	.1935	.8310	1.597	2.415	3.260	4.123	5.000	5.886	6.780	7.681
.4000	.2750	1.022	1.869	2.753	3.656	4.570	5.493	6.423	7.357	8.296
.6000	.7083	1.833	2.946	4.045	5.132	6.211	7.283	8.351	9.414	10.47
.6500	.8735	2.100	3.283	4.438	5.573	6.695	7.806	8.909	10.10	11.10
.7000	1.074	2.408	3.665	4.878	6.064	7.231	8.383	9.525	10.66	11.78
.7500	1.323	2.773	4.108	5.385	6.626	7.841	9.037	10.22	11.39	12.55
.8000	1.642	3.219	4.642	5.989	7.289	8.558	9.803	11.03	12.24	13.44
.8500	2.072	3.794	5.317	6.745	8.115	9.446	1.750	12.03	13.29	14.53
.9000	2.706	4.605	6.251	7.779	9.236	10.64	12.02	13.36	14.68	15.99
.9500	3.841	5.991	7.815	9.488	11.07	12.59	14.07	15.51	16.92	18.31
.9600	4.218	6.438	8.311	10.30	11.64	13.20	14.70	16.17	17.61	19.02
.9700	4.709	7.013	8.947	10.71	12.37	13.97	15.51	17.01	18.48	19.92
.9750	5.024	7.378	9.348	11.14	12.83	14.45	16.01	17.53	19.02	20.48
.9800	5.412	7.824	9.837	11.67	13.39	15.03	16.62	18.17	19.68	21.16
.9850	5.916	8.399	10.47	12.34	14.10	15.78	17.40	18.97	20.51	22.02
.9900	6.635	9.210	11.34	13.28	15.09	16.81	18.48	20.09	21.67	23.21
.9950	7.879	10.60	12.84	14.86	16.75	18.55	20.28	21.95	23.59	25.19
.9975	9.141	11.98	14.32	16.42	18.39	20.25	22.04	23.77	25.46	27.11
.9990	10.83	13.82	16.27	18.47	20.52	22.46	24.32	26.12	27.88	29.59

Quantiles of the χ^2_ν-Distribution

q / ν	11	12	13	14	15	16	17	18	19	20
.0010	1.834	2.214	2.617	3.041	3.483	3.942	4.416	4.905	5.407	5.921
.0025	2.232	2.661	3.112	3.582	4.070	4.573	5.092	5.623	6.167	6.723
.0050	2.603	3.074	3.565	4.075	4.601	5.142	5.697	6.265	6.844	7.434
.0100	3.053	3.571	4.107	4.660	5.229	5.812	6.408	7.015	7.633	8.260
.0150	3.363	3.910	4.476	5.057	5.653	6.263	6.884	7.516	8.159	8.811
.0200	3.609	4.178	4.765	5.368	5.985	6.614	7.255	7.906	8.567	9.237
.0250	3.816	4.404	5.009	5.629	6.262	6.908	7.564	8.231	8.907	9.591
.0300	3.997	4.601	5.221	5.856	6.503	7.163	7.832	8.512	9.200	9.897
.0400	4.309	4.939	5.584	6.243	6.914	7.596	8.288	8.989	9.698	10.42
.0500	4.575	5.226	5.892	6.571	7.261	7.962	8.672	9.391	10.12	10.85
.1000	5.578	6.304	7.042	7.790	8.547	9.312	10.09	10.86	11.65	12.44
.1500	6.336	7.114	7.901	8.696	9.499	10.31	11.12	11.95	12.77	13.60
.2000	6.989	7.807	8.634	9.467	10.31	11.15	12.00	12.86	13.72	14.58
.2500	7.584	8.438	9.299	10.17	11.04	11.91	12.79	13.68	14.56	15.45
.3000	8.148	9.034	9.926	10.82	11.72	12.62	13.53	14.44	15.35	16.27
.3500	8.587	9.497	10.41	11.33	12.25	13.17	14.10	15.03	15.96	16.89
.4000	9.237	10.18	11.13	12.08	13.03	13.98	14.94	15.89	16.85	17.81
.6000	11.53	12.58	13.64	14.69	15.73	16.78	17.82	18.87	19.91	20.95
.6500	12.18	13.27	14.35	15.42	16.49	17.56	18.63	19.70	20.76	21.83
.7000	12.90	14.01	15.12	16.22	17.32	18.42	19.51	20.60	21.69	22.77
.7500	13.70	14.85	15.98	17.12	18.25	19.37	20.49	21.60	22.72	23.83
.8000	14.63	15.81	16.98	18.15	19.31	20.47	21.61	22.76	23.90	25.04
.8500	15.77	16.99	18.20	19.41	20.60	21.79	22.98	24.16	25.33	26.50
.9000	17.27	18.55	19.81	21.06	22.31	23.54	24.77	25.99	27.20	28.41
.9500	19.68	21.03	22.36	23.68	25.00	26.30	27.59	28.87	30.14	31.41
.9600	20.41	21.79	23.14	24.49	25.82	27.14	28.44	29.75	31.04	32.32
.9700	21.34	22.74	24.12	25.49	26.85	28.19	29.52	30.84	32.16	33.46
.9750	21.92	23.34	24.74	26.12	27.49	28.85	30.19	31.53	32.85	34.17
.9800	22.62	24.05	25.47	26.87	28.26	29.63	31.00	32.35	33.69	35.02
.9850	23.50	24.96	26.40	27.83	29.23	30.63	32.01	33.38	34.74	36.09
.9900	24.72	26.22	27.69	29.14	30.58	32.00	33.41	34.81	36.19	37.57
.9950	26.76	28.30	29.82	31.32	32.80	34.27	35.72	37.16	38.58	40.00
.9975	28.73	30.32	31.88	33.43	34.95	36.46	37.95	39.42	40.88	42.34
.9990	31.26	32.91	34.53	36.12	37.70	39.25	40.79	42.31	43.82	45.31

Quantiles of the χ^2_ν-Distribution

q / ν	21	22	23	24	25	26	27	28	29	30
.0010	6.447	6.983	7.529	8.085	8.649	9.222	9.803	10.39	10.99	11.59
.0025	7.289	7.865	8.450	9.044	9.646	10.26	10.87	11.50	12.13	12.76
.0050	8.034	8.643	9.260	9.886	10.52	11.16	11.81	12.46	13.12	13.79
.0100	8.897	9.543	10.20	10.86	11.52	12.20	12.88	13.56	14.26	14.95
.0150	9.471	10.14	10.81	11.50	12.19	12.88	13.58	14.29	15.00	15.72
.0200	9.915	10.60	11.29	11.99	12.70	13.41	14.13	14.85	15.57	16.31
.0250	10.28	10.98	11.69	12.40	13.12	13.84	14.57	15.31	16.05	16.79
.0300	10.60	11.31	12.03	12.75	13.48	14.22	14.96	15.70	16.45	17.21
.0400	11.14	11.87	12.61	13.35	14.10	14.85	15.61	16.37	17.14	17.91
.0500	11.59	12.34	13.09	13.85	14.61	15.38	16.15	16.93	17.71	18.49
.1000	13.24	14.04	14.85	15.66	16.47	17.29	18.11	18.94	19.77	20.60
.1500	14.44	15.28	16.12	16.97	17.82	18.67	19.53	20.39	21.25	22.11
.2000	15.44	16.31	17.19	18.06	18.94	19.82	20.70	21.59	22.48	23.36
.2500	16.34	17.24	18.14	19.04	19.94	20.84	21.75	22.66	23.57	24.48
.3000	17.18	18.10	19.02	19.94	20.87	21.79	22.72	23.65	24.58	25.51
.3500	17.83	18.76	19.70	20.64	21.58	22.52	23.46	24.40	25.35	26.29
.4000	18.77	19.73	20.69	21.65	22.62	23.58	24.54	25.51	26.48	27.44
.6000	21.99	23.03	24.07	25.11	26.14	27.18	28.21	29.25	30.28	31.32
.6500	22.89	23.95	25.01	26.06	27.12	28.17	29.23	30.28	31.33	32.38
.7000	23.86	24.94	26.02	27.10	28.17	29.25	30.32	31.39	32.46	33.53
.7500	24.93	26.04	27.14	28.24	29.34	30.43	31.53	32.62	33.71	34.80
.8000	26.17	27.30	28.43	29.55	30.68	31.79	32.91	34.03	35.14	36.25
.8500	27.66	28.82	29.98	31.13	32.28	33.43	34.57	35.71	36.85	37.99
.9000	29.62	30.81	32.01	33.20	34.38	35.56	36.74	37.92	39.09	40.26
.9500	32.67	33.92	35.17	36.42	37.65	38.89	40.11	41.34	42.56	43.77
.9600	33.60	34.87	36.13	37.39	38.64	39.89	41.13	42.37	43.60	44.83
.9700	34.76	36.05	37.33	38.61	39.88	41.15	42.41	43.66	44.91	46.16
.9750	35.48	36.78	38.08	39.36	40.65	41.92	43.19	44.46	45.72	46.98
.9800	36.34	37.66	38.97	40.27	41.57	42.86	44.14	45.42	46.69	47.96
.9850	37.43	38.77	40.09	41.41	42.73	44.03	45.33	46.63	47.91	49.20
.9900	38.93	40.29	41.64	42.98	44.31	45.64	46.96	48.28	49.59	50.89
.9950	41.40	42.80	44.18	45.56	46.93	48.29	49.64	50.99	52.34	53.67
.9975	43.78	45.20	46.62	48.03	49.44	50.83	52.22	53.59	54.97	56.33
.9990	46.80	48.27	49.73	51.18	52.62	54.05	55.48	56.89	58.30	59.70

Quantiles of the χ^2_ν-Distribution

q / ν	31	32	33	34	35	36	37	38	39	40
.0010	12.20	12.81	13.43	14.06	14.69	15.32	15.97	16.61	17.26	17.92
.0025	13.41	14.06	14.71	15.37	16.03	16.70	17.37	18.05	18.73	19.42
.0050	14.46	15.13	15.82	16.50	17.19	17.89	18.59	19.29	20.00	20.71
.0100	15.66	16.36	17.07	17.79	18.51	19.23	19.96	20.69	21.43	22.16
.0150	16.44	17.17	17.90	18.63	19.37	20.11	20.86	21.61	22.36	23.11
.0200	17.04	17.78	18.53	19.28	20.03	20.78	21.54	22.30	23.07	23.84
.0250	17.54	18.29	19.05	19.81	20.57	21.34	22.11	22.88	23.65	24.43
.0300	17.97	18.73	19.49	20.26	21.03	21.81	22.59	23.37	24.16	24.94
.0400	18.68	19.46	20.24	21.03	21.82	22.61	23.40	24.20	25.00	25.80
.0500	19.28	20.07	20.87	21.66	22.47	23.27	24.07	24.88	25.70	26.51
.1000	21.43	22.27	23.11	23.95	24.80	25.64	26.49	27.34	28.20	29.05
.1500	22.98	23.84	24.71	25.59	26.46	27.34	28.21	29.09	29.97	30.86
.2000	24.26	25.15	26.04	26.94	27.84	28.73	29.64	30.54	31.44	32.34
.2500	25.39	26.30	27.22	28.14	29.05	29.97	30.89	31.81	32.74	33.66
.3000	26.44	27.37	28.31	29.24	30.18	31.12	32.05	32.99	33.93	34.87
.3500	27.24	28.19	29.14	30.09	31.04	31.99	32.94	33.89	34.84	35.79
.4000	28.41	29.38	30.34	31.31	32.28	33.25	34.22	35.19	36.16	37.13
.6000	32.35	33.38	34.41	35.44	36.47	37.50	38.53	39.56	40.59	41.62
.6500	33.43	34.48	35.53	36.58	37.62	38.67	39.71	40.76	41.80	42.85
.7000	34.60	35.66	36.73	37.80	38.86	39.92	40.98	42.05	43.11	44.16
.7500	35.89	36.97	38.06	39.14	40.22	41.30	42.38	43.46	44.54	45.62
.8000	37.36	38.47	39.57	40.68	41.78	42.88	43.98	45.08	46.17	47.27
.8500	39.12	40.26	41.39	42.51	43.64	44.76	45.89	47.01	48.13	49.24
.9000	41.42	42.58	43.75	44.90	46.06	47.21	48.36	49.51	50.66	51.81
.9500	44.99	46.19	47.40	48.60	49.80	51.00	52.19	53.38	54.57	55.76
.9600	46.06	47.28	48.50	49.72	50.93	52.14	53.34	54.55	55.75	56.95
.9700	47.40	48.64	49.88	51.11	52.34	53.56	54.78	56.00	57.22	58.43
.9750	48.23	49.48	50.73	51.97	53.20	54.44	55.67	56.90	58.12	59.34
.9800	49.23	50.49	51.74	53.00	54.24	55.49	56.73	57.97	59.20	60.44
.9850	50.48	51.75	53.02	54.29	55.55	56.81	58.07	59.32	60.57	61.81
.9900	52.19	53.49	54.78	56.06	57.34	58.62	59.89	61.16	62.43	63.69
.9950	55.00	56.33	57.65	58.96	60.27	61.58	62.88	64.18	65.48	66.77
.9975	57.69	59.05	60.39	61.74	63.08	64.41	65.74	67.06	68.38	69.70
.9990	61.10	62.49	63.87	65.25	66.62	67.99	69.35	70.70	72.05	73.40

Quantiles of the χ^2_ν-Distribution

q / ν	41	42	43	44	45	46	47	48	49	50
.0010	18.58	19.24	19.91	20.58	21.25	21.93	22.61	23.29	23.98	24.67
.0025	20.11	20.80	21.50	22.20	22.90	23.61	24.32	25.03	25.74	26.46
.0050	21.42	22.14	22.86	23.58	24.31	25.04	25.77	26.51	27.25	27.99
.0100	22.91	23.65	24.40	25.15	25.90	26.66	27.42	28.18	28.94	29.71
.0150	23.87	24.63	25.40	26.16	26.93	27.71	28.48	29.26	30.04	30.82
.0200	24.61	25.38	26.16	26.94	27.72	28.50	29.29	30.08	30.87	31.66
.0250	25.21	26.00	26.79	27.57	28.37	29.16	29.96	30.75	31.55	32.36
.0300	25.73	26.53	27.32	28.12	28.92	29.72	30.53	31.33	32.14	32.95
.0400	26.60	27.41	28.22	29.03	29.84	30.66	31.48	32.30	33.12	33.94
.0500	27.33	28.14	28.96	29.79	30.61	31.44	32.27	33.10	33.93	34.76
.1000	29.91	30.77	31.63	32.49	33.35	34.22	35.08	35.95	36.82	37.69
.1500	31.74	32.63	33.51	34.40	35.29	36.18	37.07	37.96	38.86	39.75
.2000	33.25	34.16	35.07	35.97	36.88	37.80	38.71	39.62	40.53	41.45
.2500	34.58	35.51	36.44	37.36	38.29	39.22	40.15	41.08	42.01	42.94
.3000	35.81	36.75	37.70	38.64	39.58	40.53	41.47	42.42	43.37	44.31
.3500	36.75	37.70	38.66	39.61	40.57	41.52	42.48	43.44	44.40	45.36
.4000	38.11	39.08	40.05	41.02	41.99	42.97	43.94	44.92	45.89	46.86
.6000	42.65	43.68	44.71	45.73	46.76	47.79	48.81	49.84	50.87	51.89
.6500	43.89	44.93	45.98	47.02	48.06	49.10	50.14	51.18	52.22	53.26
.7000	45.22	46.28	47.34	48.40	49.45	50.51	51.56	52.62	53.67	54.72
.7500	46.69	47.77	48.84	49.91	50.99	52.06	53.13	54.20	55.27	56.33
.8000	48.36	49.46	50.55	51.64	52.73	53.82	54.91	55.99	57.08	58.16
.8500	50.36	51.47	52.59	53.70	54.81	55.92	57.03	58.14	59.24	60.35
.9000	52.95	54.09	55.23	56.37	57.51	58.64	59.77	60.91	62.04	63.17
.9500	56.94	58.12	59.30	60.48	61.66	62.83	64.00	65.17	66.34	67.50
.9600	58.14	59.33	60.53	61.71	62.90	64.09	65.27	66.45	67.63	68.80
.9700	59.64	60.85	62.05	63.25	64.45	65.65	66.85	68.04	69.23	70.42
.9750	60.56	61.78	62.99	64.20	65.41	66.62	67.82	69.02	70.22	71.42
.9800	61.67	62.89	64.12	65.34	66.56	67.77	68.99	70.20	71.41	72.61
.9850	63.05	64.29	65.53	66.76	67.99	69.22	70.45	71.67	72.89	74.11
.9900	64.95	66.21	67.46	68.71	69.96	71.20	72.44	73.68	74.92	76.15
.9950	68.05	69.34	70.62	71.89	73.17	74.44	75.70	76.97	78.23	79.49
.9975	71.01	72.32	73.62	74.93	76.22	77.52	78.81	80.10	81.38	82.66
.9990	74.74	76.08	77.42	78.75	80.08	81.40	82.72	84.04	85.35	86.66

Quantiles of the χ^2_ν-Distribution

q / ν	51	52	53	54	55	56	57	58	59	60
.0010	25.37	26.07	26.77	27.47	28.17	28.88	29.59	30.30	31.02	31.74
.0025	27.19	27.91	28.64	29.37	30.10	30.83	31.57	32.31	33.05	33.79
.0050	28.73	29.48	30.23	30.98	31.73	32.49	33.25	34.01	34.77	35.53
.0100	30.48	31.25	32.02	32.79	33.57	34.35	35.13	35.91	36.70	37.48
.0150	31.60	32.39	33.18	33.97	34.76	35.55	36.35	37.14	37.94	38.74
.0200	32.46	33.26	34.06	34.86	35.66	36.46	37.27	38.08	38.89	39.70
.0250	33.16	33.97	34.78	35.59	36.40	37.21	38.03	38.84	39.66	40.48
.0300	33.76	34.58	35.39	36.21	37.03	37.85	38.67	39.50	40.32	41.15
.0400	34.77	35.59	36.42	37.25	38.08	38.92	39.75	40.59	41.43	42.27
.0500	35.60	36.44	37.28	38.12	38.96	39.80	40.65	41.49	42.34	43.19
.1000	38.56	39.43	40.31	41.18	42.06	42.94	43.82	44.70	45.58	46.46
.1500	40.65	41.55	42.45	43.34	44.24	45.15	46.05	46.95	47.85	48.76
.2000	42.36	43.28	44.20	45.12	46.04	46.96	47.88	48.80	49.72	50.64
.2500	43.87	44.81	45.74	46.68	47.61	48.55	49.48	50.42	51.36	52.29
.3000	45.26	46.21	47.16	48.11	49.06	50.01	50.96	51.91	52.86	53.81
.3500	46.31	47.27	48.23	49.19	50.15	51.11	52.07	53.04	54.00	54.96
.4000	47.84	48.81	49.79	50.76	51.74	52.72	53.69	54.67	55.64	56.62
.6000	52.92	53.94	54.97	55.99	57.02	58.04	59.06	60.09	61.11	62.13
.6500	54.30	55.33	56.37	57.41	58.45	59.48	60.52	61.56	62.59	63.63
.7000	55.77	56.83	57.88	58.93	59.98	61.03	62.08	63.13	64.18	65.23
.7500	57.40	58.47	59.53	60.60	61.67	62.73	63.79	64.86	65.92	66.98
.8000	59.25	60.33	61.41	62.50	63.58	64.66	65.74	66.82	67.89	68.97
.8500	61.45	62.55	63.65	64.76	65.85	66.95	68.05	69.15	70.25	71.34
.9000	64.30	65.42	66.55	67.67	68.80	69.92	71.04	72.16	73.28	74.40
.9500	68.67	69.83	70.99	72.15	73.31	74.47	75.62	76.78	77.93	79.08
.9600	69.98	71.15	72.32	73.49	74.66	75.83	76.99	78.16	79.32	80.48
.9700	71.61	72.80	73.98	75.16	76.34	77.52	78.70	79.88	81.05	82.23
.9750	72.62	73.81	75.00	76.19	77.38	78.57	79.75	80.94	82.12	83.30
.9800	73.82	75.02	76.22	77.42	78.62	79.81	81.01	82.20	83.39	84.58
.9850	75.33	76.54	77.75	78.96	80.17	81.38	82.58	83.79	84.99	86.19
.9900	77.39	78.62	79.84	81.07	82.29	83.51	84.73	85.95	87.17	88.38
.9950	80.75	82.00	83.25	84.50	85.75	86.99	88.24	89.48	90.72	91.95
.9975	83.94	85.22	86.49	87.77	89.03	90.30	91.57	92.83	94.09	95.34
.9990	87.97	89.27	90.57	91.87	93.17	94.46	95.75	97.04	98.32	99.61

Quantiles of the χ_ν^2-Distribution

q / ν	61	62	63	64	65	66	67	68	69	70
.0010	32.46	33.18	33.91	34.63	35.36	36.09	36.83	37.56	38.30	39.04
.0025	34.54	35.28	36.03	36.78	37.54	38.29	39.05	39.81	40.57	41.33
.0050	36.30	37.07	37.84	38.61	39.38	40.16	40.93	41.71	42.49	43.28
.0100	38.27	39.06	39.86	40.65	41.44	42.24	43.04	43.84	44.64	45.44
.0150	39.55	40.35	41.16	41.96	42.77	43.58	44.39	45.21	46.02	46.84
.0200	40.51	41.33	42.14	42.96	43.78	44.60	45.42	46.24	47.07	47.89
.0250	41.30	42.13	42.95	43.78	44.60	45.43	46.26	47.09	47.92	48.76
.0300	41.98	42.81	43.64	44.47	45.31	46.14	46.98	47.82	48.66	49.50
.0400	43.11	43.95	44.79	45.63	46.48	47.33	48.17	49.02	49.87	50.72
.0500	44.04	44.89	45.74	46.59	47.45	48.31	49.16	50.02	50.88	51.74
.1000	47.34	48.23	49.11	50.00	50.88	51.77	52.66	53.55	54.44	55.33
.1500	49.66	50.57	51.48	52.38	53.29	54.20	55.11	56.02	56.93	57.84
.2000	51.56	52.49	53.41	54.34	55.26	56.19	57.11	58.04	58.97	59.90
.2500	53.23	54.17	55.11	56.05	56.99	57.93	58.87	59.81	60.76	61.70
.3000	54.76	55.71	56.67	57.62	58.57	59.53	60.48	61.44	62.39	63.35
.3500	55.92	56.88	57.85	58.81	59.77	60.74	61.70	62.66	63.63	64.59
.4000	57.60	58.57	59.55	60.53	61.51	62.48	63.46	64.44	65.42	66.40
.6000	63.16	64.18	65.20	66.23	67.25	68.27	69.29	70.32	71.34	72.36
.6500	64.66	65.70	66.73	67.77	68.80	69.84	70.87	71.90	72.94	73.97
.7000	66.27	67.32	68.37	69.42	70.46	71.51	72.55	73.60	74.64	75.69
.7500	68.04	69.10	70.16	71.23	72.28	73.34	74.40	75.46	76.52	77.58
.8000	70.05	71.13	72.20	73.28	74.35	75.42	76.50	77.57	78.64	79.71
.8500	72.44	73.53	74.62	75.72	76.81	77.90	78.99	80.08	81.17	82.26
.9000	75.51	76.63	77.75	78.86	79.97	81.09	82.20	83.31	84.42	85.53
.9500	80.23	81.38	82.53	83.68	84.82	85.97	87.11	88.25	89.39	90.53
.9600	81.64	82.80	83.96	85.11	86.27	87.42	88.57	89.73	90.88	92.02
.9700	83.40	84.57	85.74	86.90	88.07	89.23	90.40	91.56	92.72	93.88
.9750	84.48	85.65	86.83	88.00	89.18	90.35	91.52	92.69	93.86	95.02
.9800	85.77	86.95	88.14	89.32	90.50	91.68	92.86	94.04	95.21	96.39
.9850	87.39	88.58	89.78	90.97	92.16	93.35	94.54	95.73	96.91	98.10
.9900	89.59	90.80	92.01	93.22	94.42	95.63	96.83	98.03	99.23	100.4
.9950	93.19	94.42	95.65	96.88	98.11	99.33	100.6	101.8	103.0	104.2
.9975	96.60	97.85	99.10	100.4	101.6	102.8	104.1	105.3	106.6	107.8
.9990	100.9	102.2	103.4	104.7	106.0	107.3	108.5	109.8	111.1	112.3

Quantiles of the χ^2_ν-Distribution

q / ν	71	72	73	74	75	76	77	78	79	80
.0010	39.78	40.52	41.26	42.01	42.76	43.51	44.26	45.01	45.76	46.52
.0025	42.10	42.86	43.63	44.40	45.17	45.94	46.71	47.49	48.27	49.04
.0050	44.06	44.84	45.63	46.42	47.21	48.00	48.79	49.58	50.38	51.17
.0100	46.25	47.05	47.86	48.67	49.48	50.29	51.10	51.91	52.72	53.54
.0150	47.65	48.47	49.29	50.11	50.93	51.76	52.58	53.41	54.23	55.06
.0200	48.72	49.55	50.38	51.21	52.04	52.87	53.71	54.54	55.38	56.21
.0250	49.59	50.43	51.26	52.10	52.94	53.78	54.62	55.47	56.31	57.15
.0300	50.34	51.18	52.02	52.87	53.71	54.56	55.41	56.26	57.10	57.96
.0400	51.58	52.43	53.28	54.14	55.00	55.85	56.71	57.57	58.43	59.29
.0500	52.60	53.46	54.33	55.19	56.05	56.92	57.79	58.65	59.52	60.39
.1000	56.22	57.11	58.01	58.90	59.79	60.69	61.59	62.48	63.38	64.28
.1500	58.76	59.67	60.58	61.50	62.41	63.33	64.24	65.16	66.08	66.99
.2000	60.83	61.76	62.69	63.62	64.55	65.48	66.41	67.34	68.27	69.21
.2500	62.64	63.58	64.53	65.47	66.42	67.36	68.31	69.25	70.20	71.14
.3000	64.30	65.26	66.21	67.17	68.13	69.08	70.04	71.00	71.96	72.92
.3500	65.56	66.52	67.49	68.46	69.42	70.39	71.35	72.32	73.29	74.26
.4000	67.37	68.35	69.33	70.31	71.29	72.27	73.25	74.23	75.21	76.19
.6000	73.38	74.40	75.42	76.44	77.46	78.48	79.50	80.52	81.55	82.57
.6500	75.00	76.03	77.06	78.10	79.13	80.16	81.19	82.22	83.25	84.28
.7000	76.73	77.78	78.82	79.86	80.91	81.95	82.99	84.04	85.08	86.12
.7500	78.63	79.69	80.75	81.80	82.86	83.91	84.97	86.02	87.08	88.13
.8000	80.79	81.86	82.93	84.00	85.07	86.13	87.20	88.27	89.34	90.41
.8500	83.34	84.43	85.52	86.60	87.69	88.77	89.86	90.94	92.02	93.11
.9000	86.64	87.74	88.85	89.96	91.06	92.17	93.27	94.37	95.48	96.58
.9500	91.67	92.81	93.95	95.08	96.22	97.35	98.48	99.62	100.7	101.9
.9600	93.17	94.32	95.46	96.61	97.75	98.90	100.0	101.2	102.3	103.5
.9700	95.04	96.20	97.35	98.51	99.66	100.8	102.0	103.1	104.3	105.4
.9750	96.19	97.35	98.52	99.68	100.8	102.0	103.2	104.3	105.5	106.6
.9800	97.56	98.73	99.90	101.1	102.2	103.4	104.6	105.7	106.9	108.1
.9850	99.28	100.5	101.6	102.8	104.0	105.2	106.4	107.5	108.7	109.9
.9900	101.6	102.8	104.0	105.2	106.4	107.6	108.8	110.0	111.1	112.3
.9950	105.4	106.6	107.9	109.1	110.3	111.5	112.7	113.9	115.1	116.3
.9975	109.0	110.3	111.5	112.7	114.0	115.2	116.4	117.7	118.9	120.1
.9990	113.6	114.8	116.1	117.3	118.6	119.9	121.1	122.3	123.6	124.8

Quantiles of the χ_ν^2-Distribution

q / ν	81	82	83	84	85	86	87	88	89	90
.0010	47.28	48.04	48.80	49.56	50.32	51.08	51.85	52.62	53.39	54.16
.0025	49.82	50.60	51.38	52.17	52.95	53.74	54.52	55.31	56.10	56.89
.0050	51.97	52.77	53.57	54.37	55.17	55.97	56.78	57.58	58.39	59.20
.0100	54.36	55.17	55.99	56.81	57.63	58.46	59.28	60.10	60.93	61.75
.0150	55.89	56.72	57.55	58.38	59.22	60.05	60.88	61.72	62.56	63.39
.0200	57.05	57.89	58.73	59.57	60.41	61.25	62.10	62.94	63.79	64.63
.0250	58.00	58.84	59.69	60.54	61.39	62.24	63.09	63.94	64.79	65.65
.0300	58.81	59.66	60.51	61.37	62.22	63.08	63.93	64.79	65.65	66.51
.0400	60.15	61.01	61.88	62.74	63.61	64.47	65.34	66.21	67.07	67.94
.0500	61.26	62.13	63.00	63.88	64.75	65.62	66.50	67.37	68.25	69.13
.1000	65.18	66.08	66.98	67.88	68.78	69.68	70.58	71.48	72.39	73.29
.1500	67.91	68.83	69.75	70.67	71.59	72.51	73.43	74.35	75.27	76.20
.2000	70.14	71.07	72.01	72.94	73.88	74.81	75.75	76.69	77.62	78.56
.2500	72.09	73.04	73.99	74.93	75.88	76.83	77.78	78.73	79.68	80.63
.3000	73.87	74.83	75.79	76.75	77.71	78.67	79.63	80.59	81.55	82.51
.3500	75.22	76.19	77.16	78.13	79.09	80.06	81.03	82.00	82.97	83.94
.4000	77.17	78.15	79.13	80.11	81.09	82.07	83.05	84.03	85.01	85.99
.6000	83.59	84.61	85.63	86.65	87.67	88.69	89.70	90.72	91.74	92.76
.6500	85.31	86.34	87.37	88.40	89.43	90.46	91.49	92.52	93.55	94.58
.7000	87.16	88.20	89.24	90.28	91.32	92.37	93.41	94.44	95.48	96.52
.7500	89.18	90.24	91.29	92.34	93.39	94.45	95.50	96.55	97.60	98.65
.8000	91.47	92.54	93.60	94.67	95.73	96.80	97.86	98.93	99.99	101.1
.8500	94.19	95.27	96.35	97.43	98.51	99.59	100.7	101.7	102.8	103.9
.9000	97.68	98.78	99.88	101.0	102.1	103.2	104.3	105.4	106.5	107.6
.9500	103.0	104.1	105.3	106.4	107.5	108.6	109.8	110.9	112.0	113.1
.9600	104.6	105.7	106.9	108.0	109.1	110.3	111.4	112.5	113.7	114.8
.9700	106.6	107.7	108.9	110.0	111.2	112.3	113.4	114.6	115.7	116.9
.9750	107.8	108.9	110.1	111.2	112.4	113.5	114.7	115.8	117.0	118.1
.9800	109.2	110.4	111.6	112.7	113.9	115.0	116.2	117.3	118.5	119.6
.9850	111.0	112.2	113.4	114.6	115.7	116.9	118.1	119.2	120.4	121.5
.9900	113.5	114.7	115.9	117.1	118.2	119.4	120.6	121.8	122.9	124.1
.9950	117.5	118.7	119.9	121.1	122.3	123.5	124.7	125.9	127.1	128.3
.9975	121.3	122.5	123.8	125.0	126.2	127.4	128.6	129.8	131.0	132.3
.9990	126.1	127.3	128.6	129.8	131.0	132.3	133.5	134.7	136.0	137.2

Quantiles of the χ^2_ν-Distribution

q / ν	91	92	93	94	95	96	97	98	99	100
.0010	54.93	55.70	56.47	57.25	58.02	58.80	59.58	60.36	61.14	61.92
.0025	57.68	58.48	59.27	60.07	60.86	61.66	62.46	63.26	64.06	64.86
.0050	60.00	60.81	61.63	62.44	63.25	64.06	64.88	65.69	66.51	67.33
.0100	62.58	63.41	64.24	65.07	65.90	66.73	67.56	68.40	69.23	70.07
.0150	64.23	65.07	65.91	66.75	67.60	68.44	69.28	70.13	70.97	71.82
.0200	65.48	66.33	67.18	68.03	68.88	69.73	70.58	71.43	72.29	73.14
.0250	66.50	67.36	68.21	69.07	69.92	70.78	71.64	72.50	73.36	74.22
.0300	67.37	68.23	69.09	69.95	70.82	71.68	72.54	73.41	74.28	75.14
.0400	68.81	69.68	70.55	71.43	72.30	73.17	74.05	74.92	75.80	76.67
.0500	70.00	70.88	71.76	72.64	73.52	74.40	75.28	76.16	77.05	77.93
.1000	74.20	75.10	76.01	76.91	77.82	78.73	79.63	80.54	81.45	82.36
.1500	77.12	78.04	78.96	79.89	80.81	81.74	82.66	83.59	84.51	85.44
.2000	79.50	80.43	81.37	82.31	83.25	84.19	85.13	86.07	87.01	87.95
.2500	81.57	82.52	83.47	84.42	85.38	86.33	87.28	88.23	89.18	90.13
.3000	83.47	84.43	85.39	86.36	87.32	88.28	89.24	90.20	91.17	92.13
.3500	84.91	85.88	86.84	87.81	88.78	89.75	90.72	91.70	92.67	93.64
.4000	86.97	87.95	88.94	89.92	90.90	91.88	92.86	93.84	94.83	95.81
.6000	93.78	94.80	95.82	96.84	97.86	98.87	99.89	100.9	101.9	102.9
.6500	95.61	96.64	97.67	98.70	99.72	100.8	101.8	102.8	103.8	104.9
.7000	97.56	98.60	99.64	100.7	101.7	102.8	103.8	104.8	105.9	106.9
.7500	99.70	100.8	101.8	102.8	103.9	104.9	106.0	107.0	108.1	109.1
.8000	102.1	103.2	104.2	105.3	106.4	107.4	108.5	109.5	110.6	111.7
.8500	105.0	106.1	107.1	108.2	109.3	110.4	111.4	112.5	113.6	114.7
.9000	108.7	109.8	110.9	111.9	113.0	114.1	115.2	116.3	117.4	118.5
.9500	114.3	115.4	116.5	117.6	118.8	119.9	121.0	122.1	123.2	124.3
.9600	115.9	117.1	118.2	119.3	120.5	121.6	122.7	123.8	125.0	126.1
.9700	118.0	119.1	120.3	121.4	122.6	123.7	124.8	126.0	127.1	128.2
.9750	119.3	120.4	121.6	122.7	123.9	125.0	126.1	127.3	128.4	129.6
.9800	120.8	122.0	123.1	124.3	125.4	126.6	127.7	128.8	130.0	131.1
.9850	122.7	123.9	125.0	126.2	127.3	128.5	129.7	130.8	132.0	133.1
.9900	125.3	126.5	127.6	128.8	130.0	131.1	132.3	133.5	134.6	135.8
.9950	129.5	130.7	131.9	133.1	134.2	135.4	136.6	137.8	139.0	140.2
.9975	133.5	134.7	135.9	137.1	138.3	139.5	140.7	141.9	143.1	144.3
.9990	138.4	139.7	140.9	142.1	143.3	144.6	145.8	147.0	148.2	149.4

Quantiles of the t_ν-Distribution

q / ν	1	2	3	4	5	6	7	8	9	10
.5500	.1584	.1421	.1366	.1338	.1322	.1311	.1303	.1297	.1293	.1289
.6000	.3249	.2887	.2767	.2707	.2672	.2648	.2632	.2619	.2610	.2602
.6500	.5095	.4447	.4242	.4142	.4082	.4043	.4015	.3995	.3979	.3966
.7000	.7265	.6172	.5844	.5686	.5594	.5534	.5491	.5459	.5435	.5415
.7500	1.000	.8165	.7649	.7407	.7267	.7176	.7111	.7064	.7027	.6998
.8000	1.376	1.061	.9785	.9410	.9195	.9057	.8960	.8889	.8834	.8791
.8500	1.963	1.386	1.250	1.190	1.156	1.134	1.119	1.108	1.100	1.093
.8750	2.414	1.604	1.423	1.344	1.301	1.273	1.254	1.240	1.230	1.221
.9000	3.078	1.886	1.638	1.533	1.476	1.440	1.415	1.397	1.383	1.372
.9250	4.165	2.282	1.924	1.778	1.699	1.650	1.617	1.592	1.574	1.559
.9500	6.314	2.920	2.353	2.132	2.015	1.943	1.895	1.860	1.833	1.812
.9600	7.916	3.320	2.605	2.333	2.191	2.104	2.046	2.004	1.973	1.948
.9700	10.58	3.896	2.951	2.601	2.422	2.313	2.241	2.189	2.150	2.120
.9750	12.71	4.303	3.182	2.776	2.571	2.447	2.365	2.306	2.262	2.228
.9800	15.89	4.849	3.482	2.999	2.757	2.612	2.517	2.449	2.398	2.359
.9850	21.20	5.643	3.896	3.298	3.003	2.829	2.715	2.634	2.574	2.527
.9900	31.82	6.965	4.541	3.747	3.365	3.143	2.998	2.896	2.821	2.764
.9950	63.66	9.925	5.841	4.604	4.032	3.707	3.499	3.355	3.250	3.169
.9975	127.3	14.09	7.453	5.598	4.773	4.317	4.029	3.833	3.690	3.581
.9980	159.2	15.76	8.053	5.951	5.030	4.524	4.207	3.991	3.835	3.716
.9990	318.3	22.33	10.21	7.173	5.893	5.208	4.785	4.501	4.297	4.144

Quantiles of the t_ν-Distribution

q / ν	11	12	13	14	15	16	17	18	19	20
.5500	.1286	.1283	.1281	.1280	.1278	.1277	.1276	.1274	.1274	.1273
.6000	.2596	.2590	.2586	.2582	.2579	.2576	.2573	.2571	.2569	.2567
.6500	.3956	.3947	.3940	.3933	.3928	.3923	.3919	.3915	.3912	.3909
.7000	.5399	.5386	.5375	.5365	.5357	.5350	.5344	.5338	.5333	.5329
.7500	.6974	.6955	.6938	.6924	.6912	.6901	.6892	.6884	.6876	.6869
.8000	.8755	.8726	.8701	.8681	.8662	.8647	.8633	.8620	.8609	.8600
.8500	1.088	1.083	1.079	1.076	1.074	1.071	1.069	1.067	1.066	1.064
.8750	1.214	1.209	1.204	1.200	1.197	1.194	1.191	1.189	1.187	1.185
.9000	1.363	1.356	1.350	1.345	1.341	1.337	1.333	1.330	1.328	1.325
.9250	1.548	1.538	1.530	1.523	1.517	1.512	1.508	1.504	1.500	1.497
.9500	1.796	1.782	1.771	1.761	1.753	1.746	1.740	1.734	1.729	1.725
.9600	1.928	1.912	1.899	1.887	1.878	1.869	1.862	1.855	1.850	1.844
.9700	2.096	2.076	2.060	2.046	2.034	2.024	2.015	2.007	2.000	1.994
.9750	2.201	2.179	2.160	2.145	2.131	2.120	2.110	2.101	2.093	2.086
.9800	2.328	2.303	2.282	2.264	2.249	2.235	2.224	2.214	2.205	2.197
.9850	2.491	2.461	2.436	2.415	2.397	2.382	2.368	2.356	2.346	2.336
.9900	2.718	2.681	2.650	2.624	2.602	2.583	2.567	2.552	2.539	2.528
.9950	3.106	3.055	3.012	2.977	2.947	2.921	2.898	2.878	2.861	2.845
.9975	3.497	3.428	3.372	3.326	3.286	3.252	3.222	3.197	3.174	3.153
.9980	3.624	3.550	3.489	3.438	3.395	3.358	3.326	3.298	3.273	3.251
.9990	4.025	3.930	3.852	3.787	3.733	3.686	3.646	3.610	3.579	3.552

Quantiles of the t_ν-Distribution

q / ν	21	22	23	24	25	26	27	28	29	30
.5500	.1272	.1271	.1271	.1270	.1269	.1269	.1268	.1268	.1268	.1267
.6000	.2566	.2564	.2563	.2562	.2561	.2559	.2559	.2558	.2557	.2556
.6500	.3906	.3904	.3902	.3900	.3898	.3896	.3894	.3893	.3892	.3890
.7000	.5325	.5321	.5317	.5314	.5312	.5309	.5306	.5304	.5302	.5300
.7500	.6863	.6858	.6853	.6848	.6844	.6840	.6837	.6833	.6830	.6828
.8000	.8591	.8583	.8575	.8569	.8562	.8556	.8551	.8546	.8542	.8537
.8500	1.063	1.061	1.060	1.059	1.058	1.057	1.057	1.056	1.055	1.055
.8750	1.183	1.182	1.180	1.179	1.178	1.177	1.176	1.175	1.174	1.173
.9000	1.323	1.321	1.319	1.318	1.316	1.315	1.314	1.313	1.311	1.310
.9250	1.494	1.492	1.489	1.487	1.485	1.483	1.482	1.480	1.479	1.477
.9500	1.721	1.717	1.714	1.711	1.708	1.706	1.703	1.701	1.699	1.697
.9600	1.840	1.835	1.832	1.828	1.825	1.822	1.819	1.817	1.814	1.812
.9700	1.988	1.983	1.978	1.974	1.970	1.967	1.963	1.960	1.957	1.955
.9750	2.080	2.074	2.069	2.064	2.060	2.056	2.052	2.048	2.045	2.042
.9800	2.189	2.183	2.177	2.172	2.167	2.162	2.158	2.154	2.150	2.147
.9850	2.328	2.320	2.313	2.307	2.301	2.296	2.291	2.286	2.282	2.278
.9900	2.518	2.508	2.500	2.492	2.485	2.479	2.473	2.467	2.462	2.457
.9950	2.831	2.819	2.807	2.797	2.787	2.779	2.771	2.763	2.756	2.750
.9975	3.135	3.119	3.104	3.091	3.078	3.067	3.057	3.047	3.038	3.030
.9980	3.231	3.214	3.198	3.183	3.170	3.158	3.147	3.136	3.127	3.118
.9990	3.527	3.505	3.485	3.467	3.450	3.435	3.421	3.408	3.396	3.385

Quantiles of the t_ν-Distribution

q / ν	31	32	33	34	35	36	37	38	39	40
.5500	.1267	.1267	.1266	.1266	.1266	.1266	.1265	.1265	.1265	.1265
.6000	.2555	.2555	.2554	.2553	.2553	.2552	.2552	.2551	.2551	.2550
.6500	.3889	.3888	.3887	.3886	.3885	.3884	.3883	.3882	.3882	.3881
.7000	.5298	.5297	.5295	.5294	.5292	.5291	.5289	.5288	.5287	.5286
.7500	.6825	.6822	.6820	.6818	.6816	.6814	.6812	.6810	.6808	.6807
.8000	.8533	.8530	.8526	.8523	.8520	.8517	.8514	.8512	.8509	.8507
.8500	1.054	1.054	1.053	1.052	1.052	1.052	1.051	1.051	1.050	1.050
.8750	1.172	1.172	1.171	1.170	1.170	1.169	1.169	1.168	1.168	1.167
.9000	1.309	1.309	1.308	1.307	1.306	1.306	1.305	1.304	1.304	1.303
.9250	1.476	1.475	1.474	1.473	1.472	1.471	1.470	1.469	1.468	1.468
.9500	1.696	1.694	1.692	1.691	1.690	1.688	1.687	1.686	1.685	1.684
.9600	1.810	1.808	1.806	1.805	1.803	1.802	1.800	1.799	1.798	1.796
.9700	1.952	1.950	1.948	1.946	1.944	1.942	1.940	1.939	1.937	1.936
.9750	2.040	2.037	2.035	2.032	2.030	2.028	2.026	2.024	2.023	2.021
.9800	2.144	2.141	2.138	2.136	2.133	2.131	2.129	2.127	2.125	2.123
.9850	2.275	2.271	2.268	2.265	2.262	2.260	2.257	2.255	2.252	2.250
.9900	2.453	2.449	2.445	2.441	2.438	2.435	2.431	2.429	2.426	2.423
.9950	2.744	2.738	2.733	2.728	2.724	2.719	2.715	2.712	2.708	2.704
.9975	3.022	3.015	3.008	3.002	2.996	2.990	2.985	2.980	2.976	2.971
.9980	3.109	3.102	3.094	3.088	3.081	3.075	3.070	3.064	3.059	3.055
.9990	3.375	3.365	3.356	3.348	3.340	3.333	3.326	3.319	3.313	3.307

Quantiles of the t_ν-Distribution

q / ν	41	42	43	44	45	46	47	48	49	50
.5500	.1264	.1264	.1264	.1264	.1264	.1264	.1263	.1263	.1263	.1263
.6000	.2550	.2550	.2549	.2549	.2549	.2548	.2548	.2547	.2547	.2547
.6500	.3880	.3880	.3879	.3878	.3878	.3877	.3877	.3876	.3876	.3875
.7000	.5285	.5284	.5283	.5282	.5281	.5281	.5280	.5279	.5278	.5278
.7500	.6805	.6804	.6802	.6801	.6800	.6799	.6797	.6796	.6795	.6794
.8000	.8505	.8503	.8501	.8499	.8497	.8495	.8493	.8491	.8490	.8489
.8500	1.050	1.049	1.049	1.049	1.048	1.048	1.048	1.048	1.047	1.047
.8750	1.167	1.166	1.166	1.166	1.165	1.165	1.165	1.164	1.164	1.164
.9000	1.303	1.302	1.302	1.301	1.301	1.300	1.300	1.299	1.299	1.299
.9250	1.467	1.466	1.466	1.465	1.465	1.464	1.463	1.463	1.462	1.462
.9500	1.683	1.682	1.681	1.680	1.679	1.679	1.678	1.677	1.677	1.676
.9600	1.795	1.794	1.793	1.792	1.791	1.790	1.789	1.789	1.788	1.787
.9700	1.934	1.933	1.932	1.931	1.929	1.928	1.927	1.926	1.925	1.924
.9750	2.020	2.018	2.017	2.015	2.014	2.013	2.012	2.011	2.010	2.009
.9800	2.121	2.120	2.118	2.116	2.115	2.114	2.112	2.111	2.110	2.109
.9850	2.248	2.246	2.244	2.243	2.241	2.239	2.238	2.237	2.235	2.234
.9900	2.421	2.418	2.416	2.414	2.412	2.410	2.408	2.407	2.405	2.403
.9950	2.701	2.698	2.695	2.692	2.690	2.687	2.685	2.682	2.680	2.678
.9975	2.967	2.963	2.959	2.956	2.952	2.949	2.946	2.943	2.940	2.937
.9980	3.050	3.046	3.042	3.038	3.034	3.030	3.027	3.024	3.021	3.018
.9990	3.301	3.296	3.291	3.286	3.281	3.277	3.273	3.269	3.265	3.261

Quantiles of the t_ν-Distribution

q / ν	51	52	53	54	55	56	57	58	59	60
.5500	.1263	.1263	.1263	.1263	.1262	.1262	.1262	.1262	.1262	.1262
.6000	.2547	.2546	.2546	.2546	.2546	.2546	.2545	.2545	.2545	.2545
.6500	.3875	.3875	.3874	.3874	.3873	.3873	.3873	.3872	.3872	.3872
.7000	.5277	.5276	.5276	.5275	.5275	.5274	.5273	.5273	.5272	.5272
.7500	.6793	.6792	.6791	.6791	.6790	.6789	.6788	.6787	.6787	.6786
.8000	.8487	.8486	.8484	.8483	.8482	.8481	.8480	.8479	.8478	.8476
.8500	1.047	1.047	1.047	1.046	1.046	1.046	1.046	1.046	1.046	1.045
.8750	1.164	1.163	1.163	1.163	1.163	1.162	1.162	1.162	1.162	1.162
.9000	1.298	1.298	1.298	1.297	1.297	1.297	1.297	1.296	1.296	1.296
.9250	1.462	1.461	1.461	1.460	1.460	1.460	1.459	1.459	1.459	1.458
.9500	1.675	1.675	1.674	1.674	1.673	1.673	1.672	1.672	1.671	1.671
.9600	1.786	1.786	1.785	1.784	1.784	1.783	1.782	1.782	1.781	1.781
.9700	1.924	1.923	1.922	1.921	1.920	1.920	1.919	1.918	1.918	1.917
.9750	2.008	2.007	2.006	2.005	2.004	2.003	2.002	2.002	2.001	2.000
.9800	2.108	2.107	2.106	2.105	2.104	2.103	2.102	2.101	2.100	2.099
.9850	2.233	2.231	2.230	2.229	2.228	2.227	2.226	2.225	2.224	2.223
.9900	2.402	2.400	2.399	2.397	2.396	2.395	2.394	2.392	2.391	2.390
.9950	2.676	2.674	2.672	2.670	2.668	2.667	2.665	2.663	2.662	2.660
.9975	2.934	2.932	2.929	2.927	2.925	2.923	2.920	2.918	2.916	2.915
.9980	3.015	3.012	3.009	3.007	3.005	3.002	3.000	2.998	2.996	2.994
.9990	3.258	3.255	3.251	3.248	3.245	3.242	3.239	3.237	3.234	3.232

Quantiles of the t_ν-Distribution

q / ν	61	62	63	64	65	66	67	68	69	70
.5500	.1262	.1262	.1262	.1262	.1262	.1261	.1261	.1261	.1261	.1261
.6000	.2545	.2544	.2544	.2544	.2544	.2544	.2544	.2543	.2543	.2543
.6500	.3871	.3871	.3871	.3871	.3870	.3870	.3870	.3870	.3869	.3869
.7000	.5272	.5271	.5271	.5270	.5270	.5269	.5269	.5269	.5268	.5268
.7500	.6785	.6785	.6784	.6783	.6783	.6782	.6782	.6781	.6781	.6780
.8000	.8476	.8475	.8474	.8473	.8472	.8471	.8470	.8469	.8469	.8468
.8500	1.045	1.045	1.045	1.045	1.045	1.045	1.045	1.044	1.044	1.044
.8750	1.161	1.161	1.161	1.161	1.161	1.161	1.160	1.160	1.160	1.160
.9000	1.295	1.295	1.295	1.295	1.295	1.294	1.294	1.294	1.294	1.294
.9250	1.458	1.458	1.457	1.457	1.457	1.456	1.456	1.456	1.456	1.456
.9500	1.670	1.670	1.669	1.669	1.669	1.668	1.668	1.668	1.667	1.667
.9600	1.780	1.780	1.779	1.779	1.778	1.778	1.778	1.777	1.777	1.776
.9700	1.916	1.916	1.915	1.915	1.914	1.914	1.913	1.913	1.912	1.912
.9750	2.000	1.999	1.998	1.998	1.997	1.997	1.996	1.995	1.995	1.994
.9800	2.099	2.098	2.097	2.096	2.096	2.095	2.095	2.094	2.093	2.093
.9850	2.222	2.221	2.220	2.220	2.219	2.218	2.217	2.217	2.216	2.215
.9900	2.389	2.388	2.387	2.386	2.385	2.384	2.383	2.382	2.382	2.381
.9950	2.659	2.657	2.656	2.655	2.654	2.652	2.651	2.650	2.649	2.648
.9975	2.913	2.911	2.909	2.908	2.906	2.904	2.903	2.902	2.900	2.899
.9980	2.992	2.990	2.988	2.986	2.984	2.983	2.981	2.980	2.978	2.977
.9990	3.229	3.227	3.225	3.223	3.220	3.218	3.216	3.214	3.213	3.211

Quantiles of the t_ν-Distribution

q / ν	71	72	73	74	75	76	77	78	79	80
.5500	.1261	.1261	.1261	.1261	.1261	.1261	.1261	.1261	.1261	.1261
.6000	.2543	.2543	.2543	.2543	.2542	.2542	.2542	.2542	.2542	.2542
.6500	.3869	.3869	.3868	.3868	.3868	.3868	.3868	.3867	.3867	.3867
.7000	.5267	.5267	.5267	.5267	.5266	.5266	.5266	.5265	.5265	.5265
.7500	.6779	.6779	.6779	.6778	.6778	.6777	.6777	.6776	.6776	.6776
.8000	.8466	.8466	.8466	.8465	.8464	.8464	.8463	.8463	.8462	.8461
.8500	1.044	1.044	1.044	1.044	1.044	1.044	1.043	1.043	1.043	1.043
.8750	1.160	1.160	1.160	1.159	1.159	1.159	1.159	1.159	1.159	1.159
.9000	1.293	1.293	1.293	1.293	1.293	1.293	1.293	1.292	1.292	1.292
.9250	1.455	1.455	1.455	1.454	1.454	1.454	1.454	1.454	1.453	1.453
.9500	1.666	1.666	1.666	1.666	1.665	1.665	1.665	1.665	1.664	1.664
.9600	1.776	1.776	1.775	1.775	1.775	1.774	1.774	1.774	1.773	1.773
.9700	1.911	1.911	1.910	1.910	1.910	1.909	1.909	1.909	1.908	1.908
.9750	1.994	1.993	1.993	1.993	1.992	1.992	1.991	1.991	1.990	1.990
.9800	2.092	2.092	2.091	2.091	2.090	2.090	2.089	2.089	2.088	2.088
.9850	2.215	2.214	2.213	2.213	2.212	2.212	2.211	2.211	2.210	2.209
.9900	2.380	2.379	2.379	2.378	2.377	2.376	2.376	2.375	2.374	2.374
.9950	2.647	2.646	2.645	2.644	2.643	2.642	2.641	2.640	2.640	2.639
.9975	2.897	2.896	2.895	2.894	2.892	2.891	2.890	2.889	2.888	2.887
.9980	2.975	2.974	2.972	2.971	2.970	2.969	2.967	2.966	2.965	2.964
.9990	3.209	3.207	3.206	3.204	3.202	3.201	3.199	3.198	3.197	3.195

Quantiles of the t_ν-Distribution

q / ν	81	82	83	84	85	86	87	88	89	90
.5500	.1261	.1261	.1260	.1260	.1260	.1260	.1260	.1260	.1260	.1260
.6000	.2542	.2542	.2542	.2542	.2541	.2541	.2541	.2541	.2541	.2541
.6500	.3867	.3867	.3867	.3866	.3866	.3866	.3866	.3866	.3866	.3866
.7000	.5265	.5264	.5264	.5264	.5264	.5263	.5263	.5263	.5263	.5263
.7500	.6775	.6775	.6775	.6774	.6774	.6774	.6773	.6773	.6773	.6772
.8000	.8461	.8460	.8460	.8459	.8459	.8458	.8458	.8457	.8457	.8456
.8500	1.043	1.043	1.043	1.043	1.043	1.043	1.043	1.043	1.043	1.042
.8750	1.159	1.159	1.158	1.158	1.158	1.158	1.158	1.158	1.158	1.158
.9000	1.292	1.292	1.292	1.292	1.292	1.291	1.291	1.291	1.291	1.291
.9250	1.453	1.453	1.453	1.453	1.453	1.452	1.452	1.452	1.452	1.452
.9500	1.664	1.664	1.663	1.663	1.663	1.663	1.663	1.662	1.662	1.662
.9600	1.773	1.773	1.772	1.772	1.772	1.772	1.771	1.771	1.771	1.771
.9700	1.908	1.907	1.907	1.907	1.906	1.906	1.906	1.905	1.905	1.905
.9750	1.990	1.989	1.989	1.989	1.988	1.988	1.988	1.987	1.987	1.987
.9800	2.087	2.087	2.087	2.086	2.086	2.085	2.085	2.085	2.084	2.084
.9850	2.209	2.209	2.208	2.208	2.207	2.207	2.206	2.206	2.205	2.205
.9900	2.373	2.373	2.372	2.372	2.371	2.370	2.370	2.369	2.369	2.368
.9950	2.638	2.637	2.636	2.636	2.635	2.634	2.634	2.633	2.632	2.632
.9975	2.886	2.885	2.884	2.883	2.882	2.881	2.880	2.880	2.879	2.878
.9980	2.963	2.962	2.961	2.960	2.959	2.958	2.957	2.956	2.955	2.954
.9990	3.194	3.193	3.191	3.190	3.189	3.188	3.187	3.185	3.184	3.183

Quantiles of the t_ν-Distribution

q / ν	91	92	93	94	95	96	97	98	99	100
.5500	.1260	.1260	.1260	.1260	.1260	.1260	.1260	.1260	.1260	.1260
.6000	.2541	.2541	.2541	.2541	.2541	.2541	.2540	.2540	.2540	.2540
.6500	.3865	.3865	.3865	.3865	.3865	.3865	.3864	.3865	.3864	.3864
.7000	.5262	.5262	.5262	.5262	.5262	.5262	.5261	.5261	.5261	.5261
.7500	.6772	.6772	.6771	.6771	.6771	.6771	.6770	.6770	.6770	.6769
.8000	.8456	.8455	.8455	.8455	.8454	.8454	.8452	.8453	.8453	.8452
.8500	1.042	1.042	1.042	1.042	1.042	1.042	1.042	1.042	1.042	1.042
.8750	1.158	1.158	1.158	1.157	1.157	1.157	1.157	1.157	1.157	1.157
.9000	1.291	1.291	1.291	1.291	1.290	1.290	1.290	1.290	1.290	1.290
.9250	1.452	1.452	1.451	1.451	1.451	1.451	1.451	1.451	1.451	1.451
.9500	1.662	1.662	1.661	1.661	1.661	1.661	1.660	1.660	1.660	1.660
.9600	1.770	1.770	1.770	1.770	1.770	1.769	1.769	1.769	1.769	1.769
.9700	1.905	1.904	1.904	1.904	1.904	1.903	1.903	1.903	1.903	1.902
.9750	1.986	1.986	1.986	1.986	1.985	1.985	1.985	1.984	1.984	1.984
.9800	2.084	2.083	2.083	2.083	2.082	2.082	2.082	2.081	2.081	2.081
.9850	2.205	2.204	2.204	2.204	2.203	2.203	2.203	2.202	2.202	2.201
.9900	2.368	2.368	2.367	2.367	2.366	2.366	2.365	2.365	2.365	2.364
.9950	2.631	2.630	2.630	2.629	2.629	2.628	2.628	2.627	2.626	2.626
.9975	2.877	2.876	2.876	2.875	2.874	2.873	2.873	2.872	2.871	2.871
.9980	2.953	2.952	2.952	2.951	2.950	2.949	2.949	2.948	2.947	2.946
.9990	3.182	3.181	3.180	3.179	3.178	3.177	3.176	3.175	3.175	3.174

Appendix: Greek Alphabet

A	α	alpha	I	ι	iota	P	ρ	rho
B	β	beta	K	κ	kappa	Σ	σ	sigma
Γ	γ	gamma	Λ	λ	lambda	T	τ	tau
Δ	δ	delta	M	μ	mu	Υ	υ	upsilon
E	ε	epsilon	N	ν	nu	Φ	ϕ	phi
Z	ζ	zeta	Ξ	ξ	xi	X	χ	chi
H	η	eta	O	o	omicron	Ψ	ψ	psi
Θ	θ	theta	Π	π	pi	Ω	ω	omega

Appendix: Missing Entries

Currently none! Yay! :)