Toggle the table of contents

Student's *t*-distribution

In probability and statistics, **Student's *t*-distribution** (or simply the ***t*-distribution**) t_ν is a continuous probability distribution that generalizes the standard normal distribution. Like the latter, it is symmetric around zero and bell-shaped.

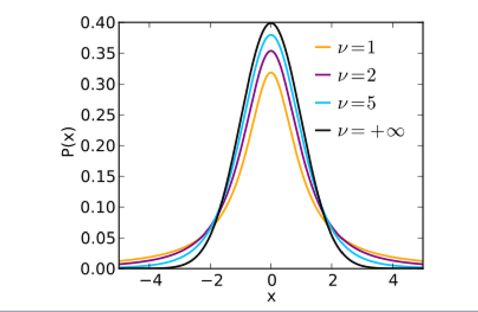
However, t_ν has heavier tails and the amount of probability mass in the tails is controlled by the parameter ν. For $\nu = 1$ the Student's *t* distribution t_ν becomes the standard Cauchy distribution, whereas for $\nu \to \infty$ it becomes the standard normal distribution $N(0,1)$.

The Student's *t*-distribution plays a role in a number of widely used statistical analyses, including Student's *t*-test for assessing the statistical significance of the difference between two sample means, the construction of confidence intervals for the difference between two population means, and in linear regression analysis.

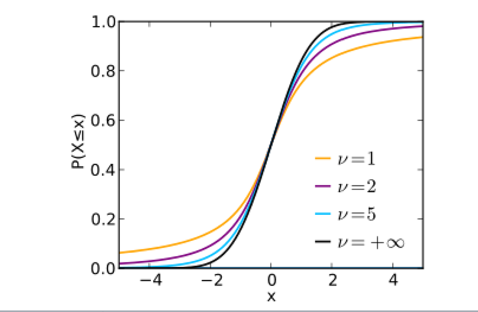
In the form of the **location-scale *t*-distribution** $lst(\mu, \tau^2, \nu)$ it generalizes the normal distribution and also arises in the Bayesian analysis of data from a normal family as a compound distribution when marginalizing over the variance parameter.

Student's *t*

Probability density function



Cumulative distribution function



Parameters	$\nu > 0$ degrees of freedom (real)
Support	$x \in (-\infty, \infty)$
PDF	$\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$
CDF	$\frac{1}{2} + x\Gamma\left(\frac{\nu+1}{2}\right) \times \frac{{}_2F_1\left(\frac{1}{2},\frac{\nu+1}{2};\frac{3}{2};-\frac{x^2}{\nu}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)}$ where ${}_2F_1$ is the hypergeometric function
Mean	0 for $\nu > 1$, otherwise undefined
Median	0
Mode	0
Variance	$\frac{\nu}{\nu-2}$ for $\nu > 2$, ∞ for $1 < \nu \le 2$, otherwise undefined
Skewness	0 for $\nu > 3$, otherwise undefined
Ex. kurtosis	$\frac{6}{\nu-4}$ for $\nu > 4$, ∞ for $2 < \nu \le 4$, otherwise undefined
Entropy	$\frac{\nu+1}{2}\left[\psi\left(\frac{1+\nu}{2}\right) - \psi\left(\frac{\nu}{2}\right)\right] + \ln\left[\sqrt{\nu}B\left(\frac{\nu}{2},\frac{1}{2}\right)\right]$ (nats) ▪ ψ: digamma function, ▪ *B*: beta function
MGF	undefined
CF	$\frac{K_{\nu/2}(\sqrt{\nu}\lvert t\rvert)\cdot(\sqrt{\nu}\lvert t\rvert)^{\nu/2}}{\Gamma(\nu/2)2^{\nu/2-1}}$ for $\nu > 0$ ▪ $K_\nu(x)$: modified Bessel function of the second kind[1]
CVaR (ES)	$\mu + s\left(\frac{\nu + T^{-1}(1-p)^2\tau(T^{-1}(1-p)^2)}{(\nu-1)(1-p)}\right)$ Where T^{-1} is the inverse standardized student-t CDF, and τ is the standardized student-t PDF.[2]

Statistician William Sealy Gosset, known as "Student"

History and etymology

In statistics, the *t*-distribution was first derived as a posterior distribution in 1876 by Helmert[3][4][5] and Lüroth.[6][7][8] The *t*-distribution also appeared in a more general form as Pearson Type IV distribution in Karl Pearson's 1895 paper.[9]

In the English-language literature, the distribution takes its name from William Sealy Gosset's 1908 paper in *Biometrika* under the pseudonym "Student".[10] One version of the origin of the pseudonym is that Gosset's employer preferred staff to use pen names when publishing scientific papers instead of their real name, so he used the name "Student" to hide his identity. Another version is that Guinness did not want their competitors to know that they were using the *t*-test to determine the quality of raw material.[11][12]

Gosset worked at the Guinness Brewery in Dublin, Ireland, and was interested in the problems of small samples – for example, the chemical properties of barley where sample sizes might be as few as 3. Gosset's paper refers to the distribution as the "frequency distribution of standard deviations of samples drawn from a normal population". It became well known through the work of Ronald Fisher, who called the distribution "Student's distribution" and represented the test value with the letter *t*.[13][14]

Definition

Probability density function

Student's ***t*-distribution** has the probability density function (PDF) given by

$$f(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\,\Gamma(\frac{\nu}{2})}\left(1+\frac{t^2}{\nu}\right)^{-(\nu+1)/2},$$

where ν is the number of *degrees of freedom* and Γ is the gamma function. This may also be written as

$$f(t) = \frac{1}{\sqrt{\nu}\,\mathrm{B}(\frac{1}{2},\frac{\nu}{2})}\left(1+\frac{t^2}{\nu}\right)^{-(\nu+1)/2},$$

where B is the Beta function. In particular for integer valued degrees of freedom ν we have:

For $\nu > 1$ even,

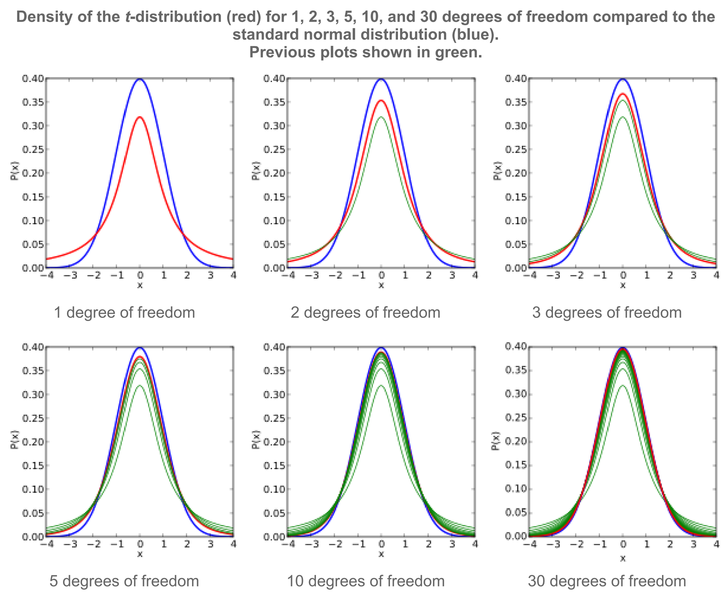
$$\frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\,\Gamma(\frac{\nu}{2})} = \frac{(\nu-1)(\nu-3)\cdots 5\cdot 3}{2\sqrt{\nu}(\nu-2)(\nu-4)\cdots 4\cdot 2}.$$

For $\nu > 1$ odd,

$$\frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\,\Gamma(\frac{\nu}{2})} = \frac{(\nu-1)(\nu-3)\cdots 4\cdot 2}{\pi\sqrt{\nu}(\nu-2)(\nu-4)\cdots 5\cdot 3}.$$

The probability density function is symmetric, and its overall shape resembles the bell shape of a normally distributed variable with mean 0 and variance 1, except that it is a bit lower and wider. As the number of degrees of freedom grows, the *t*-distribution approaches the normal distribution with mean 0 and variance 1. For this reason ν is also known as the normality parameter.[15]

The following images show the density of the *t*-distribution for increasing values of ν. The normal distribution is shown as a blue line for comparison. Note that the *t*-distribution (red line) becomes closer to the normal distribution as ν increases.



Cumulative distribution function

The cumulative distribution function (CDF) can be written in terms of *I*, the regularized incomplete beta function. For *t* > 0,

$$F(t) = \int_{-\infty}^{t} f(u)\,du = 1 - \tfrac{1}{2} I_{x(t)}\left(\tfrac{\nu}{2}, \tfrac{1}{2}\right),$$

where

$$x(t) = \frac{\nu}{t^2+\nu}.$$

Other values would be obtained by symmetry. An alternative formula, valid for $t^2 < \nu$, is

$$\int_{-\infty}^{t} f(u)\,du = \tfrac{1}{2} + t\frac{\Gamma\left(\frac{1}{2}(\nu+1)\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)}\,{}_2F_1\left(\tfrac{1}{2}, \tfrac{1}{2}(\nu+1); \tfrac{3}{2}; -\tfrac{t^2}{\nu}\right),$$

where ${}_2F_1$ is a particular case of the hypergeometric function.

For information on its inverse cumulative distribution function, see quantile function § Student's t-distribution.

Special cases

Certain values of ν give a simple form for Student's t-distribution.

ν	PDF	CDF	notes
1	$\frac{1}{\pi(1+t^2)}$	$\frac{1}{2} + \frac{1}{\pi}\arctan(t)$	See Cauchy distribution
2	$\frac{1}{2\sqrt{2}\left(1+\frac{t^2}{2}\right)^{3/2}}$	$\frac{1}{2} + \frac{t}{2\sqrt{2}\sqrt{1+\frac{t^2}{2}}}$	
3	$\frac{2}{\pi\sqrt{3}\left(1+\frac{t^2}{3}\right)^2}$	$\frac{1}{2} + \frac{1}{\pi}\left[\frac{1}{\sqrt{3}}\frac{t}{1+\frac{t^2}{3}} + \arctan\left(\frac{t}{\sqrt{3}}\right)\right]$	
4	$\frac{3}{8\left(1+\frac{t^2}{4}\right)^{5/2}}$	$\frac{1}{2} + \frac{3}{8}\frac{t}{\sqrt{1+\frac{t^2}{4}}}\left[1 - \frac{1}{12}\frac{t^2}{1+\frac{t^2}{4}}\right]$	
5	$\frac{8}{3\pi\sqrt{5}\left(1+\frac{t^2}{5}\right)^3}$	$\frac{1}{2} + \frac{1}{\pi}\left[\frac{t}{\sqrt{5}\left(1+\frac{t^2}{5}\right)}\left(1+\frac{2}{3\left(1+\frac{t^2}{5}\right)}\right) + \arctan\left(\frac{t}{\sqrt{5}}\right)\right]$	
∞	$\frac{1}{\sqrt{2\pi}}e^{-t^2/2}$	$\frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{t}{\sqrt{2}}\right)\right]$	See *Normal distribution*, *Error function*

Moments

For $\nu > 1$, the raw moments of the *t*-distribution are

$$\mathrm{E}(T^k) = \begin{cases} 0 & k \text{ odd}, \quad 0 < k < \nu \\ \frac{1}{\sqrt{\pi}\Gamma\left(\frac{\nu}{2}\right)}\left[\Gamma\left(\frac{k+1}{2}\right)\Gamma\left(\frac{\nu-k}{2}\right)\nu^{\frac{k}{2}}\right] & k \text{ even}, \quad 0 < k < \nu. \end{cases}$$

Moments of order ν or higher do not exist.[16]

The term for $0 < k < \nu$, *k* even, may be simplified using the properties of the gamma function to

$$\mathrm{E}(T^k) = \nu^{\frac{k}{2}} \prod_{i=1}^{k/2} \frac{2i-1}{\nu-2i} \qquad k \text{ even}, \quad 0 < k < \nu.$$

For a *t*-distribution with ν degrees of freedom, the expected value is 0 if $\nu > 1$, and its variance is $\frac{\nu}{\nu-2}$ if $\nu > 2$. The skewness is 0 if $\nu > 3$ and the excess kurtosis is $\frac{6}{\nu-4}$ if $\nu > 4$.

Location-scale *t*-distribution

Location-scale transformation

Student's *t*-distribution generalizes to the three parameter **location-scale *t*-distribution** $lst(\mu, \tau^2, \nu)$ by introducing a location parameter μ and a scale parameter τ. With

$$T \sim t_\nu$$

and location-scale family transformation

$$X = \mu + \tau T$$

we get

$$X \sim lst(\mu, \tau^2, \nu)$$

The resulting distribution is also called the **non-standardized Student's *t*-distribution**.

Density and first two moments

The location-scale *t* distribution has a density defined by:[17]

$$p(x \mid \nu, \mu, \tau) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\pi\nu}\tau}\left(1 + \frac{1}{\nu}\left(\frac{x-\mu}{\tau}\right)^2\right)^{-(\nu+1)/2}$$

Equivalently, the density can be written in terms of τ^2:

$$p(x \mid \nu, \mu, \tau^2) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\pi\nu\tau^2}}\left(1 + \frac{1}{\nu}\frac{(x-\mu)^2}{\tau^2}\right)^{-(\nu+1)/2}$$

Other properties of this version of the distribution are:[17]

$$\begin{aligned} \mathrm{E}(X) &= \mu && \text{for } \nu > 1 \\ \mathrm{var}(X) &= \tau^2\frac{\nu}{\nu-2} && \text{for } \nu > 2 \\ \mathrm{mode}(X) &= \mu \end{aligned}$$

Special cases

- If X follows a location-scale *t*-distribution $X \sim \mathrm{lst}\left(\mu, \tau^2, \nu\right)$ then for $\nu \to \infty$ X is normally distributed $X \sim \mathrm{N}\left(\mu, \tau^2\right)$ with mean μ and variance τ^2.
- The location-scale *t*-distribution $\mathrm{lst}\left(\mu, \tau^2, \nu = 1\right)$ with degree of freedom $\nu = 1$ is equivalent to the Cauchy distribution $\mathrm{Cau}\left(\mu, \tau\right)$.
- The location-scale *t*-distribution $\mathrm{lst}\left(\mu = 0, \tau^2 = 1, \nu\right)$ with $\mu = 0$ and $\tau^2 = 1$ reduces to the Student's *t*-distribution t_ν

How the *t*-distribution arises (characterization)

Sampling distribution of t-statistic

The *t*-distribution arises as the sampling distribution of the *t*-statistic. Below the one-sample *t*-statistic is discussed, for the corresponding two-sample *t*-statistic see Student's t-test.

Unbiased variance estimate

Let $x_1, \ldots, x_n \sim N(\mu, \sigma^2)$ be independent and identically distributed samples from a normal distribution with mean μ and variance σ^2. The sample mean and unbiased sample variance are given by:

$$\bar{x} = \frac{x_1 + \cdots + x_n}{n},$$

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

The resulting (one sample) *t*-statistic is given by

$$t = \frac{\bar{x} - \mu}{\sqrt{s^2/n}} \sim t_{n-1}.$$

and is distributed according to a Student's t-distribution with $n-1$ degrees of freedom.

Thus for inference purposes the t-statistic is a useful "pivotal quantity" in the case when the mean and variance (μ, σ^2) are unknown population parameters, in the sense that the t-statistic has then a probability distribution that depends on neither μ nor σ^2.

ML variance estimate

Instead of the unbiased estimate s^2 we may also use the maximum likelihood estimate

$$s_{ML}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

yielding the statistic

$$t_{ML} = \frac{\bar{x} - \mu}{\sqrt{s_{ML}^2/n}} = \sqrt{\frac{n}{n-1}}t.$$

This is distributed according to the location-scale t-distribution:

$$t_{ML} \sim lst(0, \tau^2 = n/(n-1), n-1).$$

Compound distribution of normal with inverse gamma distribution

The location-scale t-distribution results from compounding a Gaussian distribution (normal distribution) with mean μ and unknown variance, with an inverse gamma distribution placed over the variance with parameters $a = \nu/2$ and $b = \nu\tau^2/2$. In other words, the random variable X is assumed to have a Gaussian distribution with an unknown variance distributed as inverse gamma, and then the variance is marginalized out (integrated out).

Equivalently, this distribution results from compounding a Gaussian distribution with a scaled-inverse-chi-squared distribution with parameters ν and τ^2. The scaled-inverse-chi-squared distribution is exactly the same distribution as the inverse gamma distribution, but with a different parameterization, i.e. $\nu = 2a,\ \tau^2 = \frac{b}{a}$.

The reason for the usefulness of this characterization is that in Bayesian statistics the inverse gamma distribution is the conjugate prior distribution of the variance of a Gaussian distribution. As a result, the location-scale t-distribution arises naturally in many Bayesian inference problems.[18]

Maximum entropy distribution

Student's t-distribution is the maximum entropy probability distribution for a random variate X for which $\mathrm{E}(\ln(\nu + X^2))$ is fixed.[19]

Further properties

Monte Carlo sampling

There are various approaches to constructing random samples from the Student's t-distribution. The matter depends on whether the samples are required on a stand-alone basis, or are to be constructed by application of a quantile function to uniform samples; e.g., in the multi-dimensional applications basis of copula-dependency. In the case of stand-alone sampling, an extension of the Box–Muller method and its polar form is easily deployed.[20] It has the merit that it applies equally well to all real positive degrees of freedom, ν, while many other candidate methods fail if ν is close to zero.[20]

Integral of Student's probability density function and *p*-value

The function $A(t \mid \nu)$ is the integral of Student's probability density function, $f(t)$ between $-t$ and t, for $t \geq 0$. It thus gives the probability that a value of t less than that calculated from observed data would occur by chance. Therefore, the function $A(t \mid \nu)$ can be used when testing whether the difference between the means of two sets of data is statistically significant, by calculating the corresponding value of t and the probability of its occurrence if the two sets of data were drawn from the same population. This is used in a variety of situations, particularly in t-tests. For the statistic t, with ν degrees of freedom, $A(t \mid \nu)$ is the probability that t would be less than the observed value if the two means were the same (provided that the smaller mean is subtracted from the larger, so that $t \geq 0$). It can be easily calculated from the cumulative distribution function $F_\nu(t)$ of the t-distribution:

$$A(t \mid \nu) = F_\nu(t) - F_\nu(-t) = 1 - I_{\frac{\nu}{\nu+t^2}}\left(\frac{\nu}{2}, \frac{1}{2}\right),$$

where I_x is the regularized incomplete beta function (a, b).

For statistical hypothesis testing this function is used to construct the p-value.

Related distributions

- The noncentral t-distribution generalizes the t-distribution to include a location parameter. Unlike the nonstandardized t-distributions, the noncentral distributions are not symmetric (the median is not the same as the mode).
- The **discrete Student's *t*-distribution** is defined by its probability mass function at r being proportional to:[21]

  $$\prod_{j=1}^{k}\frac{1}{(r+j+a)^2+b^2} \qquad r = \ldots, -1, 0, 1, \ldots.$$

  Here a, b, and k are parameters. This distribution arises from the construction of a system of discrete distributions similar to that of the Pearson distributions for continuous distributions.[22]
- One can generate Student-t samples by taking the ratio of variables from the normal distribution and the square-root of **χ^2-distribution**. If we use instead of the normal distribution, e.g., the Irwin–Hall distribution, we obtain over-all a symmetric 4-parameter distribution, which includes the normal, the uniform, the triangular, the Student-t and the Cauchy distribution. This is also more flexible than some other symmetric generalizations of the normal distribution.
- t-distribution is an instance of ratio distributions.

Uses

In frequentist statistical inference

Student's t-distribution arises in a variety of statistical estimation problems where the goal is to estimate an unknown parameter, such as a mean value, in a setting where the data are observed with additive errors. If (as in nearly all practical statistical work) the population standard deviation of these errors is unknown and has to be estimated from the data, the t-distribution is often used to account for the extra uncertainty that results from this estimation. In most such problems, if the standard deviation of the errors were known, a normal distribution would be used instead of the t-distribution.

Confidence intervals and hypothesis tests are two statistical procedures in which the quantiles of the sampling distribution of a particular statistic (e.g. the standard score) are required. In any situation where this statistic is a linear function of the data, divided by the usual estimate of the standard deviation, the resulting quantity can be rescaled and centered to follow Student's t-distribution. Statistical analyses involving means, weighted means, and regression coefficients all lead to statistics having this form.

Quite often, textbook problems will treat the population standard deviation as if it were known and thereby avoid the need to use the Student's t-distribution. These problems are generally of two kinds: (1) those in which the sample size is so large that one may treat a data-based estimate of the variance as if it were certain, and (2) those that illustrate mathematical reasoning, in which the problem of estimating the standard deviation is temporarily ignored because that is not the point that the author or instructor is then explaining.

Hypothesis testing

A number of statistics can be shown to have t-distributions for samples of moderate size under null hypotheses that are of interest, so that the t-distribution forms the basis for significance tests. For example, the distribution of Spearman's rank correlation coefficient ρ, in the null case (zero correlation) is well approximated by the t distribution for sample sizes above about 20.

Confidence intervals

Suppose the number A is so chosen that

$$\Pr(-A < T < A) = 0.9,$$

when T has a t-distribution with $n-1$ degrees of freedom. By symmetry, this is the same as saying that A satisfies

$$\Pr(T < A) = 0.95,$$

so A is the "95th percentile" of this probability distribution, or $A = t_{(0.05, n-1)}$. Then

$$\Pr\left(-A < \frac{\overline{X}_n - \mu}{S_n/\sqrt{n}} < A\right) = 0.9,$$

and this is equivalent to

$$\Pr\left(\overline{X}_n - A\frac{S_n}{\sqrt{n}} < \mu < \overline{X}_n + A\frac{S_n}{\sqrt{n}}\right) = 0.9.$$

Therefore, the interval whose endpoints are

$$\overline{X}_n \pm A\frac{S_n}{\sqrt{n}}$$

is a 90% confidence interval for μ. Therefore, if we find the mean of a set of observations that we can reasonably expect to have a normal distribution, we can use the t-distribution to examine whether the confidence limits on that mean include some theoretically predicted value – such as the value predicted on a null hypothesis.

It is this result that is used in the Student's t-tests: since the difference between the means of samples from two normal distributions is itself distributed normally, the t-distribution can be used to examine whether that difference can reasonably be supposed to be zero.

If the data are normally distributed, the one-sided $(1 - \alpha)$-upper confidence limit (UCL) of the mean, can be calculated using the following equation:

$$\mathrm{UCL}_{1-\alpha} = \overline{X}_n + t_{\alpha,n-1}\frac{S_n}{\sqrt{n}}.$$

The resulting UCL will be the greatest average value that will occur for a given confidence interval and population size. In other words, $\overline{X}_n$ being the mean of the set of observations, the probability that the mean of the distribution is inferior to $\mathrm{UCL}_{1-\alpha}$ is equal to the confidence level $1 - \alpha$.

Prediction intervals

The t-distribution can be used to construct a prediction interval for an unobserved sample from a normal distribution with unknown mean and variance.

In Bayesian statistics

The Student's t-distribution, especially in its three-parameter (location-scale) version, arises frequently in Bayesian statistics as a result of its connection with the normal distribution. Whenever the variance of a normally distributed random variable is unknown and a conjugate prior placed over it that follows an inverse gamma distribution, the resulting marginal distribution of the variable will follow a Student's t-distribution. Equivalent constructions with the same results involve a conjugate scaled-inverse-chi-squared distribution over the variance, or a conjugate gamma distribution over the precision. If an improper prior proportional to σ^{-2} is placed over the variance, the t-distribution also arises. This is the case regardless of whether the mean of the normally distributed variable is known, is unknown distributed according to a conjugate normally distributed prior, or is unknown distributed according to an improper constant prior.

Related situations that also produce a t-distribution are:

- The marginal posterior distribution of the unknown mean of a normally distributed variable, with unknown prior mean and variance following the above model.
- The prior predictive distribution and posterior predictive distribution of a new normally distributed data point when a series of independent identically distributed normally distributed data points have been observed, with prior mean and variance as in the above model.

Robust parametric modeling

The t-distribution is often used as an alternative to the normal distribution as a model for data, which often has heavier tails than the normal distribution allows for; see e.g. Lange et al.[23] The classical approach was to identify outliers (e.g., using Grubbs's test) and exclude or downweight them in some way. However, it is not always easy to identify outliers (especially in high dimensions), and the t-distribution is a natural choice of model for such data and provides a parametric approach to robust statistics.

A Bayesian account can be found in Gelman et al.[24] The degrees of freedom parameter controls the kurtosis of the distribution and is correlated with the scale parameter. The likelihood can have multiple local maxima and, as such, it is often necessary to fix the degrees of freedom at a fairly low value and estimate the other parameters taking this as given. Some authors report that values between 3 and 9 are often good choices. Venables and Ripley suggest that a value of 5 is often a good choice.

Student's *t*-process

For practical regression and prediction needs, Student's t-processes were introduced, that are generalisations of the Student t-distributions for functions. A Student's t-process is constructed from the Student t-distributions like a Gaussian process is constructed from the Gaussian distributions. For a Gaussian process, all sets of values have a multidimensional Gaussian distribution. Analogously, $X(t)$ is a Student t-process on an interval $I = [a, b]$ if the correspondent values of the process $X(t_1), \ldots, X(t_n)$ $(t_i \in I)$ have a joint multivariate Student t-distribution.[25] These processes are used for regression, prediction, Bayesian optimization and related problems. For multivariate regression and multi-output prediction, the multivariate Student t-processes are introduced and used.[26]

Table of selected values

The following table lists values for t-distributions with ν degrees of freedom for a range of one-sided or two-sided critical regions. The first column is ν, the percentages along the top are confidence levels, and the numbers in the body of the table are the $t_{\alpha,n-1}$ factors described in the section on confidence intervals.

The last row with infinite ν gives critical points for a normal distribution since a t-distribution with infinitely many degrees of freedom is a normal distribution. (See Related distributions above).

One-sided	**75%**	**80%**	**85%**	**90%**	**95%**	**97.5%**	**99%**	**99.5%**	**99.75%**	**99.9%**	**99.95%**
Two-sided	**50%**	**60%**	**70%**	**80%**	**90%**	**95%**	**98%**	**99%**	**99.5%**	**99.8%**	**99.9%**
1	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657	127.321	318.309	636.619
2	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	14.089	22.327	31.599
3	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	7.453	10.215	12.924
4	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767
24	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
50	0.679	0.849	1.047	1.299	1.676	2.009	2.403	2.678	2.937	3.261	3.496
60	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
80	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	2.887	3.195	3.416
100	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	2.871	3.174	3.390
120	0.677	0.845	1.041	1.289	1.658	1.980	2.358	2.617	2.860	3.160	3.373
∞	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291
One-sided	**75%**	**80%**	**85%**	**90%**	**95%**	**97.5%**	**99%**	**99.5%**	**99.75%**	**99.9%**	**99.95%**
Two-sided	**50%**	**60%**	**70%**	**80%**	**90%**	**95%**	**98%**	**99%**	**99.5%**	**99.8%**	**99.9%**

Calculating the confidence interval

Let's say we have a sample with size 11, sample mean 10, and sample variance 2. For 90% confidence with 10 degrees of freedom, the one-sided t-value from the table is 1.372. Then with confidence interval calculated from

$$\overline{X}_n \pm t_{\alpha,\nu}\frac{S_n}{\sqrt{n}},$$

we determine that with 90% confidence we have a true mean lying below

$$10 + 1.372\frac{\sqrt{2}}{\sqrt{11}} = 10.585.$$

In other words, 90% of the times that an upper threshold is calculated by this method from particular samples, this upper threshold exceeds the true mean.

And with 90% confidence we have a true mean lying above

$$10 - 1.372\frac{\sqrt{2}}{\sqrt{11}} = 9.414.$$

In other words, 90% of the times that a lower threshold is calculated by this method from particular samples, this lower threshold lies below the true mean.

So that at 80% confidence (calculated from 100% − 2 × (1 − 90%) = 80%), we have a true mean lying within the interval

$$\left(10 - 1.372\frac{\sqrt{2}}{\sqrt{11}}, 10 + 1.372\frac{\sqrt{2}}{\sqrt{11}}\right) = (9.414, 10.585).$$

Saying that 80% of the times that upper and lower thresholds are calculated by this method from a given sample, the true mean is both below the upper threshold and above the lower threshold is not the same as saying that there is an 80% probability that the true mean lies between a particular pair of upper and lower thresholds that have been calculated by this method; see confidence interval and prosecutor's fallacy.

Nowadays, statistical software, such as the R programming language, and functions available in many spreadsheet programs compute values of the t-distribution and its inverse without tables.

See also

- *F*-distribution
- Folded-*t* and half-*t* distributions
- Hotelling's *T*-squared distribution
- Multivariate Student distribution
- Standard normal table (*Z*-distribution table)
- *t*-statistic
- Tau distribution, for internally studentized residuals
- Wilks' lambda distribution
- Wishart distribution
- Modified half-normal distribution[27] with the pdf on $(0,\infty)$ is given as $$f(x) = \frac{2\beta^{\frac{\alpha}{2}} x^{\alpha-1}\exp(-\beta x^2 + \gamma x)}{\Psi\left(\frac{\alpha}{2}, \frac{\gamma}{\sqrt{\beta}}\right)}, \text{ where } \Psi(\alpha, z) = {}_1\Psi_1\left(\begin{matrix}\left(\alpha, \frac{1}{2}\right)\\(1,0)\end{matrix}; z\right)$$ denotes the Fox–Wright Psi function.

Mathematics portal

Notes

1. Hurst, Simon. "The Characteristic Function of the Student *t* Distribution" (https://web.archive.org/web/20100218072259/http://wwwmaths.anu.edu.au/research.reports/srr/95/044/). *Financial Mathematics Research Report No. FMRR006-95, Statistics Research Report No. SRR044-95*. Archived from the original (http://wwwmaths.anu.edu.au/research.reports/srr/95/044/) on February 18, 2010.
2. Norton, Matthew; Khokhlov, Valentyn; Uryasev, Stan (2019). "Calculating CVaR and bPOE for common probability distributions with application to portfolio optimization and density estimation" (http://uryasev.ams.stonybrook.edu/wp-content/uploads/2019/10/Norton2019_CVaR_bPOE.pdf) (PDF). *Annals of Operations Research*. Springer. **299** (1–2): 1281–1315. doi:10.1007/s10479-019-03373-1 (https://doi.org/10.1007%2Fs10479-019-03373-1). Retrieved 2023-02-27.
3. Helmert FR (1875). "Über die Berechnung des wahrscheinlichen Fehlers aus einer endlichen Anzahl wahrer Beobachtungsfehler". *Z. Math. U. Physik*. **20**: 300–3.
4. Helmert FR (1876). "Über die Wahrscheinlichkeit der Potenzsummen der Beobachtungsfehler und uber einige damit in Zusammenhang stehende Fragen". *Z. Math. Phys*. **21**: 192–218.
5. Helmert FR (1876). "Die Genauigkeit der Formel von Peters zur Berechnung des wahrscheinlichen Beobachtungsfehlers directer Beobachtungen gleicher Genauigkeit" (https://zenodo.org/record/1424695) [The accuracy of Peters' formula for calculating the probable observation error of direct observations of the same accuracy]. *Astron. Nachr.* (in German). **88** (8–9): 113–132. Bibcode:1876AN.....88..113H (https://ui.adsabs.harvard.edu/abs/1876AN.....88..113H). doi:10.1002/asna.18760880802 (https://doi.org/10.1002%2Fasna.18760880802).
6. Lüroth J (1876). "Vergleichung von zwei Werten des wahrscheinlichen Fehlers" (https://zenodo.org/record/1424693). *Astron. Nachr.* **87** (14): 209–20. Bibcode:1876AN.....87..209L (https://ui.adsabs.harvard.edu/abs/1876AN.....87..209L). doi:10.1002/asna.18760871402 (https://doi.org/10.1002%2Fasna.18760871402).
7. Pfanzagl J, Sheynin O (1996). "Studies in the history of probability and statistics. XLIV. A forerunner of the *t*-distribution". *Biometrika*. **83** (4): 891–898. doi:10.1093/biomet/83.4.891 (https://doi.org/10.1093%2Fbiomet%2F83.4.891). MR 1766040 (https://mathscinet.ams.org/mathscinet-getitem?mr=1766040).
8. Sheynin O (1995). "Helmert's work in the theory of errors". *Arch. Hist. Exact Sci.* **49** (1): 73–104. doi:10.1007/BF00374700 (https://doi.org/10.1007%2FBF00374700). S2CID 121241599 (https://api.semanticscholar.org/CorpusID:121241599).
9. Pearson, K. (1895-01-01). "Contributions to the Mathematical Theory of Evolution. II. Skew Variation in Homogeneous Material" (https://doi.org/10.1098%2Frsta.1895.0010). *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. **186**: 343–414 (374). Bibcode:1895RSPTA.186..343P (https://ui.adsabs.harvard.edu/abs/1895RSPTA.186..343P). doi:10.1098/rsta.1895.0010 (https://doi.org/10.1098%2Frsta.1895.0010). ISSN 1364-503X (https://www.worldcat.org/issn/1364-503X).
10. "Student" [William Sealy Gosset] (1908). "The probable error of a mean" (http://www.york.ac.uk/depts/maths/histstat/student.pdf) (PDF). *Biometrika*. **6** (1): 1–25. doi:10.1093/biomet/6.1.1 (https://doi.org/10.1093%2Fbiomet%2F6.1.1). hdl:10338.dmlcz/143545 (https://hdl.handle.net/10338.dmlcz%2F143545). JSTOR 2331554 (https://www.jstor.org/stable/2331554).
11. Wendl MC (2016). "Pseudonymous fame". *Science*. **351** (6280): 1406. Bibcode:2016Sci...351.1406W (https://ui.adsabs.harvard.edu/abs/2016Sci...351.1406W). doi:10.1126/science.351.6280.1406 (https://doi.org/10.1126%2Fscience.351.6280.1406). PMID 27013722 (https://pubmed.ncbi.nlm.nih.gov/27013722).
12. Mortimer RG (2005). *Mathematics for physical chemistry* (https://archive.org/details/mathematicsforph00mort_321) (3rd ed.). Burlington, MA: Elsevier. pp. 326 (https://archive.org/details/mathematicsforph00mort_321/page/n339). ISBN 9780080492889. OCLC 156200058 (https://www.worldcat.org/oclc/156200058).
13. Fisher RA (1925). "Applications of 'Student's' distribution" (https://web.archive.org/web/20160305130235/http://www.sothis.ro/user/content/4ef6e90670749a86-student_distribution_1925.pdf) (PDF). *Metron*. **5**: 90–104. Archived from the original (http://www.sothis.ro/user/content/4ef6e90670749a86-student_distribution_1925.pdf) (PDF) on 5 March 2016.
14. Walpole RE, Myers R, Myers S, et al. (2006). *Probability & Statistics for Engineers & Scientists* (7th ed.). New Delhi: Pearson. p. 237. ISBN 9788177584042. OCLC 818811849 (https://www.worldcat.org/oclc/818811849).
15. Kruschke JK (2015). *Doing Bayesian Data Analysis* (2nd ed.). Academic Press. ISBN 9780124058880. OCLC 959632184 (https://www.worldcat.org/oclc/959632184).
16. Casella G, Berger RL (1990). *Statistical Inference*. Duxbury Resource Center. p. 56. ISBN 9780534119584.
17. Jackman, S. (2009). *Bayesian Analysis for the Social Sciences* (https://archive.org/details/bayesianmodeling00jack). Wiley Series in Probability and Statistics. Wiley. p. 507 (https://archive.org/details/bayesianmodeling00jack/page/n542). doi:10.1002/9780470686621 (https://doi.org/10.1002%2F9780470686621). ISBN 9780470011546.
18. Gelman AB, Carlin JS, Rubin DB, et al. (1997). *Bayesian Data Analysis* (2nd ed.). Boca Raton: Chapman & Hall. p. 68. ISBN 9780412039911.
19. Park SY, Bera AK (2009). "Maximum entropy autoregressive conditional heteroskedasticity model". *J. Econom.* **150** (2): 219–230. doi:10.1016/j.jeconom.2008.12.014 (https://doi.org/10.1016%2Fj.jeconom.2008.12.014).
20. Bailey RW (1994). "Polar Generation of Random Variates with the *t*-Distribution". *Math. Comput.* **62** (206): 779–781. Bibcode:1994MaCom..62..779B (https://ui.adsabs.harvard.edu/abs/1994MaCom..62..779B). doi:10.2307/2153537 (https://doi.org/10.2307%2F2153537). JSTOR 2153537 (https://www.jstor.org/stable/2153537).
21. Ord JK (1972). *Families of Frequency Distributions*. London: Griffin. ISBN 9780852641378. See Table 5.1.
22. Ord JK (1972). "Chapter 5". *Families of frequency distributions*. London: Griffin. ISBN 9780852641378.
23. Lange KL, Little RJ, Taylor JM (1989). "Robust Statistical Modeling Using the *t* Distribution" (https://cloudfront.escholarship.org/dist/prd/content/qt27s1d3h7/qt27s1d3h7.pdf) (PDF). *J. Am. Stat. Assoc.* **84** (408): 881–896. doi:10.1080/01621459.1989.10478852 (https://doi.org/10.1080%2F01621459.1989.10478852). JSTOR 2290063 (https://www.jstor.org/stable/2290063).
24. Gelman AB, Carlin JB, Stern HS, et al. (2014). "Computationally efficient Markov chain simulation". *Bayesian Data Analysis*. Boca Raton, Florida: CRC Press. p. 293. ISBN 9781439898208.
25. Shah, Amar; Wilson, Andrew Gordon; Ghahramani, Zoubin (2014). "Student-*t* processes as alternatives to Gaussian processes" (http://proceedings.mlr.press/v33/shah14.pdf) (PDF). *JMLR*. **33** (Proceedings of the 17th International Conference on Artificial Intelligence and Statistics (AISTATS) 2014, Reykjavik, Iceland): 877–885. arXiv:1402.4306 (https://arxiv.org/abs/1402.4306).
26. Chen, Zexun; Wang, Bo; Gorban, Alexander N. (2019). "Multivariate Gaussian and Student-*t* process regression for multi-output prediction" (https://doi.org/10.1007%2Fs00521-019-04687-8). *Neural Computing and Applications*. **32** (8): 3005–3028. arXiv:1703.04455 (https://arxiv.org/abs/1703.04455). doi:10.1007/s00521-019-04687-8 (https://doi.org/10.1007%2Fs00521-019-04687-8).
27. Sun, Jingchao; Kong, Maiying; Pal, Subhadip (22 June 2021). "The Modified-Half-Normal distribution: Properties and an efficient sampling scheme" (https://www.tandfonline.com/doi/abs/10.1080/03610926.2021.1934700?journalCode=lsta20). *Communications in Statistics - Theory and Methods*. **52** (5): 1591–1613. doi:10.1080/03610926.2021.1934700 (https://doi.org/10.1080%2F03610926.2021.1934700). ISSN 0361-0926 (https://www.worldcat.org/issn/0361-0926). S2CID 237919587 (https://api.semanticscholar.org/CorpusID:237919587).

References

- Senn, S.; Richardson, W. (1994). "The first *t*-test". *Statistics in Medicine*. **13** (8): 785–803. doi:10.1002/sim.4780130802 (https://doi.org/10.1002%2Fsim.4780130802). PMID 8047737 (https://pubmed.ncbi.nlm.nih.gov/8047737).
- Hogg RV, Craig AT (1978). *Introduction to Mathematical Statistics* (4th ed.). New York: Macmillan. ASIN B010WFO0SA (https://www.amazon.com/dp/B010WFO0SA).
- Venables, W. N.; Ripley, B. D. (2002). *Modern Applied Statistics with S* (Fourth ed.). Springer.
- Gelman, Andrew; John B. Carlin; Hal S. Stern; Donald B. Rubin (2003). *Bayesian Data Analysis* (http://www.stat.columbia.edu/~gelman/book/) (Second ed.). CRC/Chapman & Hall. ISBN 1-58488-388-X.

External links

- "Student distribution" (https://www.encyclopediaofmath.org/index.php?title=Student_distribution), *Encyclopedia of Mathematics*, EMS Press, 2001 [1994]
- Earliest Known Uses of Some of the Words of Mathematics (S) (http://jeff560.tripod.com/s.html) *(Remarks on the history of the term "Student's distribution")*
- Rouaud, M. (2013), *Probability, Statistics and Estimation* (http://www.incertitudes.fr/book.pdf) (PDF) (short ed.) First Students on page 112.
- Student's t-Distribution (https://flexbooks.ck12.org/cbook/ck-12-interactive-algebra-2-for-ccss/section/6.6/related/lesson/students-t-distribution-adv-pst/), Archived (https://web.archive.org/web/20210410192326/https://flexbooks.ck12.org/cbook/ck-12-interactive-algebra-2-for-ccss/section/6.6/related/lesson/students-t-distribution-adv-pst/) 2021-04-10 at the Wayback Machine