Student t test p value table

Find the t-critical value by confidence level & DF for the Student's t distribution >>>Click to use a T-value calculator 2 * Skewness: 0 for ν > 3 * Excess kurtosis: 6ν - 4 for ν > 4 The t-distribution is used to model real-valued random variables that have certain properties, such as having a finite variance. It is also related to other probability distributions, including the normal distribution. In summary, the Student's t-distribution is a mathematical concept used in statistics and probability theory to describe a specific type of probability distribution with certain properties and applications. The Student's t distribution is a statistical model that is used in various analyses, including hypothesis testing and confidence intervals. It has a bell-shaped curve, but its tails are heavier than those of a standard normal distribution. The shape of the t distribution depends on a parameter called "number of degrees of freedom" (ν), which controls the amount of probability mass in the tails. When ν = 1, the t distribution becomes the Cauchy distribution, and when ν approaches infinity, it approaches the standard normal distribution. The name "Student's t distribution" was given to this model by William Gosset, who used a pseudonym for his work at Guinness Brewery in Ireland. The distribution is widely used in statistics, including linear regression analysis and Bayesian analysis. The probability density function of the t distribution is given by a formula that involves the gamma function (Γ) and beta function (B). When ν is an integer, the formula can be simplified to show that the t distribution is related to the normal distribution. Interestingly, the t distribution has a symmetric shape that resembles a bell curve, but it is slightly lower and wider than a standard normal distribution. As the number of degrees of freedom increases, the t distribution approaches the normal distribution with mean 0 and variance 1. The t-distribution, a probability distribution used in statistics, is characterized by a parameter called the normality parameter (ν). The distribution becomes closer to the normal distribution as ν increases. The cumulative distribution function (CDF) of the t-distribution can be expressed using the regularized incomplete beta function or alternative formulas for certain values of ν. The CDF is defined as the integral of the probability density function (PDF) from negative infinity to a given value t. The PDF of the t-distribution has various forms, depending on the value of ν. The text also provides examples and special cases of the t-distribution, including specific formulas for certain values of ν. Additionally, it mentions that certain values of ν give simple forms for the distribution's CDF and PDF. Overall, the text provides a comprehensive overview of the t-distribution, its properties, and its relationships with other distributions, such as the normal distribution. The Student's t-distribution is a probability distribution that is widely used in statistical analysis. It is characterized by a parameter called "degrees of freedom" (ν) which determines its shape and properties. The raw moments of the t-distribution are defined as the expected value of the kth power of the random variable T, denoted as E{Tk}. The formula for these moments depends on whether k is even or odd. For even values of k, the moment can be expressed in terms of special functions such as the gamma function and the error function. The expected value of a t-distribution with ν degrees of freedom is 0 if ν > 1, and its variance is ν(ν-2) if ν > 2. The skewness and excess kurtosis of the distribution also depend on the value of ν. The t-distribution can be defined as the distribution of the random variable T = ZV/√(u), where Z is a standard normal, V has a chi-squared distribution with u degrees of freedom, and Z and V are independent. This definition allows for non-central versions of the t-distribution to be studied. In practice, researchers often use simulated data from a normal distribution to estimate the power of Student's t-tests using a t-distribution. The t-distribution is an important tool in statistical inference, and its properties and behavior have been extensively studied and analyzed. The sample mean X n is given by the formula X n = (1/n) * (X 1 + ... + X n), and the unbiased estimate of the variance s 2 is calculated as s 2 = (1/(n-1)) * Σ (i=1 to n) ((X i - X¯ n)^2). The random variable V = (n-1) * s 2 / σ 2 follows a chi-squared distribution with ν degrees of freedom. σ² {\displaystyle \sigma ^{2}~.} Instead of using the unbiased estimator s² {\displaystyle s^{2}} we can use the maximum likelihood estimator sML² = (1/n) ∑(xi - x̄)² {\displaystyle s_{\mathsf{ML}}^{2}={\frac {1}{n}}\ \sum _{i=1}^{n}(x_{i}-{\bar{x}})^{2}\ } to obtain the statistic tML = (x̄ - μ) / √(sML²/n) {\displaystyle t_{\mathsf{ML}}={\frac {{\bar{x}}-\mu }{\sqrt {s_{\mathsf{ML}}^{2}/n\ }}}.} This distribution follows the location-scale t-distribution: tML ~ ℓst (0, τ² = n/(n-1), n-1) {\displaystyle t_{\mathsf{ML}}\sim \operatorname {\ell st} (0,\ \tau ^{2}=n/(n-1),\ n-1)~.} The location-scale t-distribution is a result of compounding a Gaussian distribution with mean μ {\displaystyle \mu \ } and unknown variance, with an inverse gamma distribution placed over the variance with parameters a = ν/2 {\displaystyle a={\frac {u}{2}}\ } and b = ντ²/2 {\displaystyle b={\frac {u\ \tau ^{2}\ }{2}}~.} This essentially means that 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). Alternatively, this distribution arises from compounding a Gaussian distribution with a scaled-inverse-chi-squared distribution with parameters ν {\displaystyle u } and τ² {\displaystyle \tau ^{2}~.} The scaled-inverse-chi-squared distribution is equivalent to the inverse gamma distribution but with a different parameterization: ν = 2a, τ² = b/a {\displaystyle u=2\ a,\;{\tau }^{2}={\frac {b}{a}}~.} This characterization of the location-scale t-distribution is useful because 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.[9] Student's t-distribution is the maximum entropy probability distribution for a random variate X having a certain value of E(ln(ν + X²)) {\displaystyle \operatorname{\mathbb{E}}\left\{\ln(u+X^{2})\right\}\ } .[10][clarification needed][better source needed] This follows from the observation that the pdf can be written in exponential family form with ν + X² as sufficient statistic. The function A(t | ν) is the integral of Student's probability density function, f(t), between -t and t, for t ≥ 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 | ν) 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 | ν) is the probability that t would be less than the observed value if the two means were the same. The t-distribution can be defined in terms of its cumulative distribution function (CDF), which can be calculated using the regularized incomplete beta function. The CDF is used to construct p-values for statistical hypothesis testing. The noncentral t-distribution generalizes the standard t-distribution by introducing a noncentrality parameter, resulting in asymmetric distributions. The discrete Student's t-distribution is defined by its probability mass function and arises from constructing a system of discrete distributions similar to those found in continuous distributions. This distribution can be generated using variables from the normal and χ² distributions, or other distributions such as the Irwin-Hall distribution. The t-distribution is an instance of ratio distributions and can be generalized to the three-parameter location-scale t-distribution by introducing a location parameter and a scale parameter. This distribution has a density function defined in terms of the gamma function and the regularized incomplete beta function. Some key points from the text include: * The t-distribution can be defined using its CDF, which can be calculated using the regularized incomplete beta function. * The noncentral t-distribution generalizes the standard t-distribution by introducing a noncentrality parameter. * The discrete Student's t-distribution arises from constructing a system of discrete distributions similar to those found in continuous distributions. * The location-scale t-distribution is a three-parameter distribution that includes a location parameter and a scale parameter. Overall, the text provides an overview of various aspects of the t-distribution, including its definition, properties, and generalizations. The location-scale t distribution can be described by its properties. When u > 1, the expected value of X is μ, and when u > 2, the variance of X is τ^2 * (u / (u - 2)). The mode of X is always μ. If X follows a location-scale t distribution with parameters μ, τ^2, and ν, then as ν approaches infinity, X becomes normally distributed with mean μ and variance τ^2. The location-scale t distribution with degree of freedom ν = 1 is equivalent to the Cauchy distribution. When μ = 0 and τ^2 = 1, the location-scale t distribution reduces to the Student's t distribution with ν degrees of freedom. Student's t distribution arises in statistical estimation problems where the data are observed with additive errors. If the population standard deviation of these errors is unknown, the t distribution is used to account for the extra uncertainty that results from estimating it. In most cases, if the standard deviation were known, a normal distribution would be used instead. Confidence intervals and hypothesis tests often require the quantiles of the sampling distribution of a statistic. When 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 with this form. Textbook problems often treat the population standard deviation as known to avoid using the Student's t distribution. These problems are typically of two kinds: those with large sample sizes where a data-based estimate of the variance can be treated as certain, and mathematical reasoning problems that temporarily ignore the problem of estimating the standard deviation. Given that the author or instructor is not focusing on this point, it's important to note that the t-distribution serves as a foundation for significance tests. In particular, Spearman's rank correlation coefficient ρ can be approximated by the t-distribution when sample sizes are moderate and null hypotheses are of interest. It's also worth noting that if we choose A such that P(-A < T < A) = 0.9, where T has a t-distribution with n-1 degrees of freedom, then A is equivalent to the "95th percentile" of this probability distribution or A = t(0.05, n-1). This allows us to construct a 90% confidence interval for μ, which can be used to examine whether a theoretically predicted value is included within reasonable limits. Specifically, if we find the mean of a set of observations that are normally distributed, we can use the t-distribution to determine whether the confidence limits on that mean include some theoretically predicted value. The Student's t-test also relies on this result, as it examines whether the difference between the means of two normal distributions is zero. If the data are normally distributed, we can calculate a one-sided (1-α) upper confidence limit (UCL) using the equation: UCL1-α = X¯n + tα, n-1 Sn/n. This resulting UCL represents the greatest average value that will occur for a given confidence interval and population size. The t distribution arises in various statistical models, particularly when a conjugate prior is placed over the variance of a normally distributed random variable. This can include inverse gamma or scaled-inverse-chi-squared distributions, as well as gamma distributions with precision. Additionally, an improper prior proportional to 1/σ² will also yield a t distribution, regardless of whether the mean is known or unknown. The multivariate Student t processes are introduced and used for regression and multi-output prediction. A table listing critical values for t distributions with varying degrees of freedom (ν) is provided, along with confidence levels (α) and corresponding t α , n − 1 factors. The table covers one-sided and two-sided critical regions for different degrees of freedom, including a row with infinite ν that corresponds to a normal distribution. The table shows the critical values for various degrees of freedom, including: * One-sided critical regions: + 75%, 80%, 85%, 90%, 95%, 97.5%, 99%, 99.5%, 99.75%, 99.9%, and 99.95% * Two-sided critical regions: + 50%, 60%, 70%, 80%, 90%, 95%, 98%, 99%, 99.5%, 99.8%, 99.9%, and 1 The table provides the corresponding t α , n − 1 factors for each degree of freedom, which can be used to calculate confidence intervals or test hypotheses in regression models. Note that the original text appears to be a technical specification or documentation for statistical software or programming language, but the paraphrased version aims to provide a general understanding of the content without losing its technical details. The confidence interval for the sample mean was calculated using the one-sided t value from the table. For a 90% confidence level with 10 degrees of freedom, the one-sided t value is 1.372. This value is used to determine the upper and lower bounds of the confidence interval. With 90% confidence, we can say that there is a 90% probability that the true mean lies below 10 + (1.372 * √2 / √11) = 10.585. Similarly, with 90% confidence, we can also say that there is a 90% probability that the true mean lies above 10 - (1.372 * √2 / √11) = 9.414. At an 80% confidence level, which is calculated from 100% - 2 × (1 - 90%) = 80%, we can say that there is a 80% probability that the true mean lies within the interval (9.414, 10.585). It's worth noting that this does not necessarily mean 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. Statistical software such as R programming language and many spreadsheet programs can compute values of the t distribution and its inverse without tables. There are various approaches to constructing random samples from the Student's t distribution, depending 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. The t distribution was first derived as a posterior distribution in 1876 by Helmert and Lüroth, and is an example of Stigler's Law of Eponymy. The Student's t-distribution appeared in a more general form as the Pearson type IV distribution in Karl Pearson's 1895 paper, but it got its name from William Sealy Gosset's 1908 paper where he used the pseudonym "Student". The origin of the pseudonym is unclear, with some saying it was to hide his identity and others that Guinness didn't want competitors to know they were using the t-test. Gosset worked at Guinness and studied small samples like barley quality where sample sizes could be as low as 3. Ronald Fisher popularized the distribution by calling it "Student's distribution" and representing the test value with the letter t. The t-distribution, also known as Student's t-distribution, is used in statistics to analyze data when the standard normal distribution does not accurately model the observed values. The distribution was first introduced by William Sealy Gosset in 1908 and was later developed further by Karl Pearson. Several key works have contributed to our understanding of the t-distribution: * Fisher's (1925) work on "Applications of 'Student's' distribution" laid the foundation for its use in statistical analysis. * Gelman et al. (1997) discussed Bayesian data analysis, which often employs the t-distribution. * Park and Bera (2009) developed a maximum entropy autoregressive conditional heteroskedasticity model using the t-distribution. The t-distribution has been used in various fields, including: * Statistical inference: Helmert's work on error distributions in the late 19th century laid the groundwork for later developments of the t-distribution. * Bayesian analysis: Gelman et al. (2014) discussed efficient Markov chain simulation methods using the t-distribution. * Machine learning: Shah et al. (2014) proposed student t processes as alternatives to Gaussian processes. Some notable researchers have made significant contributions to our understanding of the t-distribution, including: * Gosset, who first introduced the distribution * Pearson, who developed the concept further * Helmert, who worked on error distributions and their applications Overall, the t-distribution is a crucial tool in statistical analysis, particularly when dealing with non-normal data or estimating parameters. The concept of Student's t-distribution has a rich history. In the late 19th century, researchers such as Helmert and Pearson made significant contributions to the theory of errors, which laid the groundwork for the development of the t-distribution. The term "Student" was actually a pseudonym used by William Sealy Gosset, who published a paper on the probable error of a mean in 1908. The t-distribution was further developed and refined over the years by various researchers, including Pfanzagl and Sheynin, who studied Helmert's work in the theory of errors. The distribution gained widespread acceptance and became a fundamental concept in statistics. In recent times, the t-distribution has been used extensively in statistical analysis and modeling. Researchers such as Wendl and Sun et al. have made significant contributions to the understanding and application of the t-distribution. The t-distribution table, also known as the T Table or Student's T-Table, is a widely used reference for calculating critical values and confidence intervals. The table contains values for one-tailed and two-tailed distributions up to 1000 degrees of freedom (df) and a confidence level of 99.9%. Overall, the t-distribution has become an essential tool in statistical analysis, and its history reflects the contributions of many researchers over the years. To use a T-table, several prerequisites must be met beforehand. #### The number of tails: - Determine whether the t-test is one-tailed or two-tailed to select the appropriate row in the table for alpha levels. - One-tail rows have values such as 0.50, 0.25, 0.20, and so on. - Two-tails rows include values like 1.00, 0.50, 0.40, 0.30, etc. #### Degrees of freedom: - The degrees of freedom (df) indicate the number of independent values that can vary without violating any constraints. - If mentioned in the problem statement, use it directly. Otherwise, calculate df by subtracting one from the sample size (n – 1). #### Alpha level: - Also known as the significance level, this is the probability of rejecting the null hypothesis when it's true. - Common alpha levels for t-tests include 0.01, 0.05, and 0.10. For an one-tailed test with a sample size of 23, the degrees of freedom are calculated as n - 1 = 22. In a research study using one-tailed test with alpha level of 0.10 and sample size of 12, the critical value should be compared to the t-value obtained as the test statistic, where the degrees of freedom (df) is 11. The T distribution is used for small samples, whereas other distributions like Z and Chi Squared are used in different contexts. A T Table is preferred over a Z Table for small sample sizes, typically N