CARLETON UNIVERSITY

SCHOOL OF
MATHEMATICS AND STATISTICS

HONOURS PROJECT

TITLE: ERROR ESTIMATION OF ASYMPTOTIC EXPANSION FOR THE T-DISTRIBUTION DENSITY

AUTHOR: MENGJIE WANG

SUPERVISOR: YIQIANG ZHAO

DATE: 2019/05/04

ABSTRACT

This report is on a study of error estimation of the asymptotic expansion for the *t*-distribution. As known to all, the *t*-distribution is asymptotic to the standard normal distribution when the degrees of freedom is large enough. The premise here that the degrees of freedom be large enough is quite vague. We know when the degrees of freedom go to infinity, the two distributions become identical. But usually in practice, we cannot take such large sample population. So there must be an acceptable and applicable level of degrees of freedom in real cases. In other words we want to find how large is large enough. It is an interesting project and worth doing research for many industries. According to Bangjun Ding's report, in 2002, on asymptotic expansions for the *t*-distribution density, he provided an expansion of the *t*-distribution which can be very important in small sample studies. His report only provides an asymptotic expansion for the *t*-distribution to the n^{-2} term. Based on his report, I continue the asymptotic expansion further to include the significant term of order n^{-3} and provide a numerical analysis of the asymptotic expansion, which are my main contributions to the study.

Acknowledgements

I would like to express my deep gratitude to Professor Yiqiang Zhao and Professor Song Cai, my research supervisors, for their valuable and constructive suggestions during the planning and development of this research work. Their willingness to give their time so generously has been very much appreciated.

I would also like to extend my thanks to Professor Bangjun Ding for offering me resources in running the program.

1. INTRODUCTION

First, let us review two important distributions in statistics, the standard normal distribution and the *t*-distribution, and recall how they are defined.

The standard normal distribution is a very important continuous distribution since it is often used to represent random variables whose distributions are unknown and is the theoretical basis for many statistical methods. Suppose that *X* is a normally distributed random variable with mean μ and standard deviation σ. Then the probability density function (pdf) of X is:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \text{ for } -\infty < x < \infty.$$

Now suppose that we want to standardize this variable, that is, convert it into a random variable that has a normal distribution, with $\mu = 0$ and $\sigma^2 = 1$. How would we do that? If we go $X - \mu$, then this quantity would have a mean of zero, and if we divide by σ, then $\frac{\mathrm{X}-\mathrm{\mu}}{\sigma}$ this whole quantity would have a standard deviation of 1. This is a basic liner transformation and we force this quantity to have a mean of zero and a standard deviation of 1. If we let $Z = \frac{X-\mu}{\sigma}$, then this random variable Z has the standard normal distribution. We could write this as Z is distributed normally with a mean of zero and a variance of 1, i.e. $Z \sim N(0,1)$. The standard normal distribution is a bell-shaped curve. Now we conclude that a standard normal distribution is a normal distribution with zero mean and unit variance, given by the probability density function $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

Next before we introduce the *t*-distribution distribution, we would like to introduce the definition of iid and the χ^2 distribution. According to Wikipedia, in probability theory and statistics, a collection of random variables is independent and identically distributed if each random variable has the same probability distribution as the others and all are mutually independent. This property is usually abbreviated as i.i.d. or iid or IID.

If a random variable X has the standard normal distribution, then X^2 has the χ^2 distribution with one degree of freedom. If $X_1, X_2, \dots, X_n$ are independent standard normal random variables, i.e. these random variables are iid, then $X_1^2 + X_2^2 + \cdots + X_n^2$ has a χ^2 distribution with *n* degrees of freedom. We can also write $\chi^2(n) = \sum_{i=1}^{n} X_i^2$, where $X_i \sim N(0,1)$. Note that the degrees of freedom are the number of independent squared standard normal random variables that we are adding up. The mean of the χ^2 distribution with *n* degrees of freedom is *n* and the variance of it is $2n$.

The *t*-distribution is an important continuous probability distribution that is widely used in statistical inference. The *t*-distribution is very closely related to the standard normal distribution. Here is how we can define the *t*-distribution. Suppose:

- X has the standard normal distribution.
- U has the χ^2 distribution with n degrees of freedom.
- X and U are independent random variables.

Then $\frac{X}{\sqrt{\frac{U}{n}}}$ has the *t*-distribution with *n* degrees of freedom.

So, we often represent this random variable with lower case letter t.

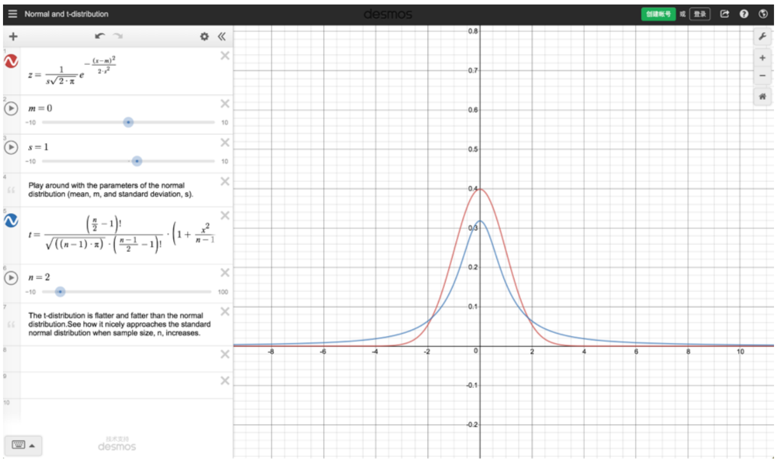
If we are drawing *n* independent observations from a normally distributed population with mean μ and variance σ^2, then $\frac{\bar{X}-\mu}{S/\sqrt{n}}$ has the *t*-distribution with *n*-1 degrees of freedom. Here the sample mean $\bar{X}$ is a random variable representing the mean of the *n* observations, and the sample standard deviation S is a random variable representing the standard deviation of the *n* observations. The implication of this is that the *t*-distribution often arises in statistical inference on means, when we are sampling from a normally distributed population with both mean and variance unknown.

The probability density function (pdf) of the *t*-distribution with *n* degrees of freedom is given by:

$$t(x) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{n\pi}}(1+\frac{x^2}{n})^{-\frac{n+1}{2}}$$

for $-\infty < x < \infty$, where Γ represents the Gamma function. If a random variable has a *t*-distribution, then it can take on any finite value. The mean of the t-distribution is zero and the variance of it is $\frac{n}{n-2}$, n > 2. As we know, the population variance is a constant and the sample variance is a random variable. In many cases the population variance is unknown. So we need to replace the population deviation with the sample deviation. That is to say, when the population variance is known, the sample mean is distributed normally. When the population variance is unknown, the sample mean has the *t*-distribution. This is why the *t*-distribution is very important in statistics.

We can see the similarity from the graphs below (screenshots from online graph generator website). Like the standard normal distribution, the *t*-distribution is symmetric about 0, but the *t*-distribution has heavier tails, with more area in the tails, and it has a lower peak.



Normal and t-distribution

1 $z = \frac{1}{s\sqrt{2\cdot\pi}}e^{-\frac{(x-m)^2}{2\cdot s^2}}$

2 $m = 0$ −10 10

3 $s = 1$ −10 10

4 Play around with the parameters of the normal distribution (mean, m, and standard deviation, s).

5 $t = \frac{\left(\frac{n}{2}-1\right)!}{\sqrt{((n-1)\cdot\pi)}\cdot\left(\frac{n-1}{2}-1\right)!}\cdot\left(1+\frac{x^2}{n-1}\right.$

6 $n = 2$ −10 100

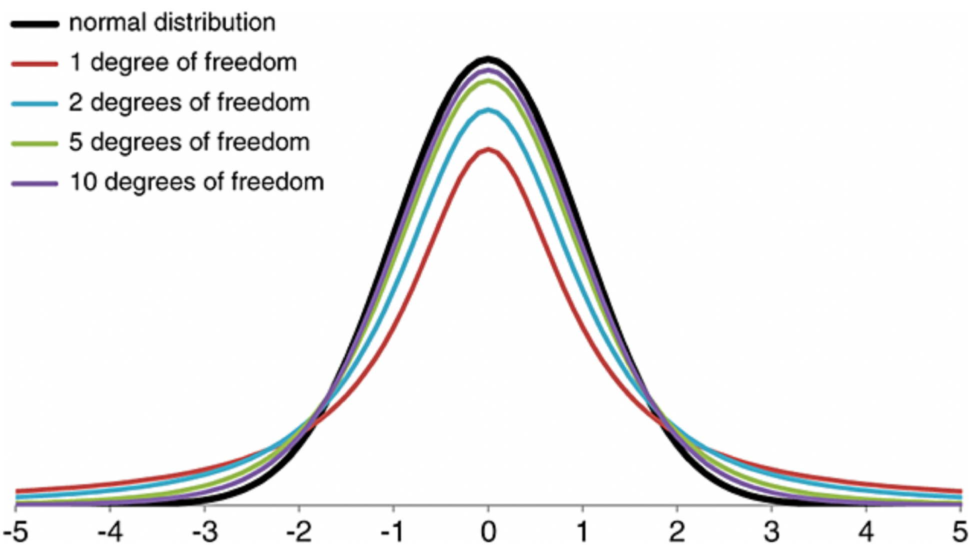
7 The t-distribution is flatter and fatter than the normal distribution.See how it nicely approaches the standard normal distribution when sample size, n, increases.

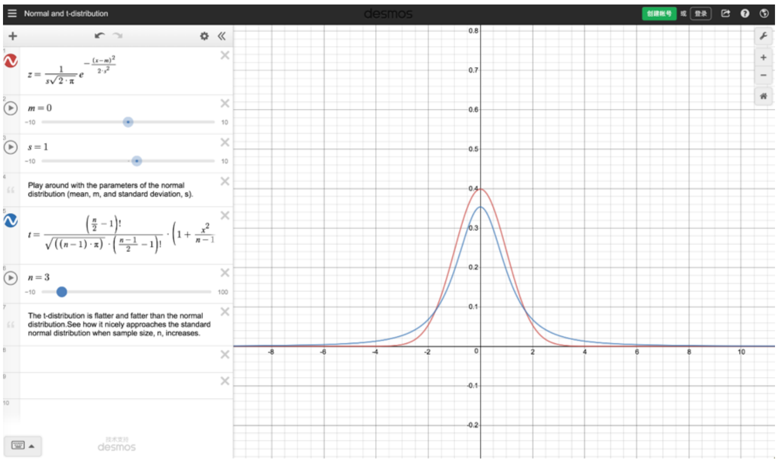
Standard normal distribution in red

The *t*-distribution in blue (1 degree of freedom)

As the degrees of freedom increase, the *t*-distribution tends toward the standard normal distribution (see the trend from the graphs beneath this paragraph). To see this trend, we have provided graphs for the degrees of freedom from *n*=1 to *n*=20. Here, the curve in blue is the *t*-distribution. If we continue to let the degrees of freedoms increase, the blue curve will get closer and closer to the red standard normal curve. The median is at zero and the mean is also zero. Note that the variance is the degrees of freedom over the degrees of freedom minus 2, in the case of 20 degrees of freedom, $\sigma^2 = \frac{20}{18}$. Even though these two curves look very similar, there are some practical differences. The *t*-distribution with 20 degrees of freedom still has quite a bit more area in the tails. So, there are important differences between the t and the standard normal distributions even they look similar when plotted. In practice, we often want to find areas and percentiles of the *t*-distribution, and that requires integrating the probability density function. Unfortunately, there is no closed form solution and that integration must be carried out numerically. Areas and percentiles for the *t*-distribution can be found using software or a t table. (see the comparison between the standard normal distribution and the *t*-distribution with various degrees of freedom from the first graph beneath this paragraph[4])

The t-distribution and its relationship to the normal distribution





$z = \frac{1}{s\sqrt{2 \cdot \pi}} e^{-\frac{(x-m)^2}{2 \cdot s^2}}$

$m = 0$

−10 10

$s = 1$

−10 10

Play around with the parameters of the normal distribution (mean, m, and standard deviation, s).

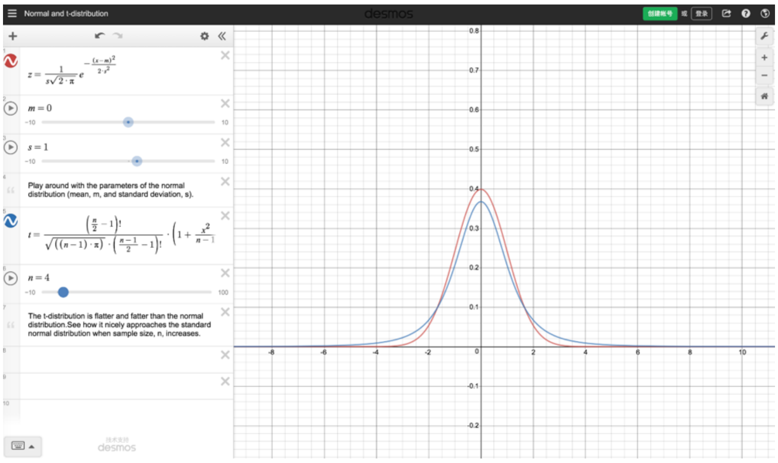
$t = \frac{\left(\frac{n}{2} - 1\right)!}{\sqrt{((n-1) \cdot \pi)} \cdot \left(\frac{n-1}{2} - 1\right)!} \cdot \left(1 + \frac{x^2}{n-1}\right.$

$n = 3$

−10 100

The t-distribution is flatter and fatter than the normal distribution.See how it nicely approaches the standard normal distribution when sample size, n, increases.

The *t*-distribution with 2 degrees of freedom



Normal and t-distribution

$z=\frac{1}{s\sqrt{2\cdot\pi}}e^{-\frac{(x-m)^2}{2\cdot s^2}}$

$m=0$

$s=1$

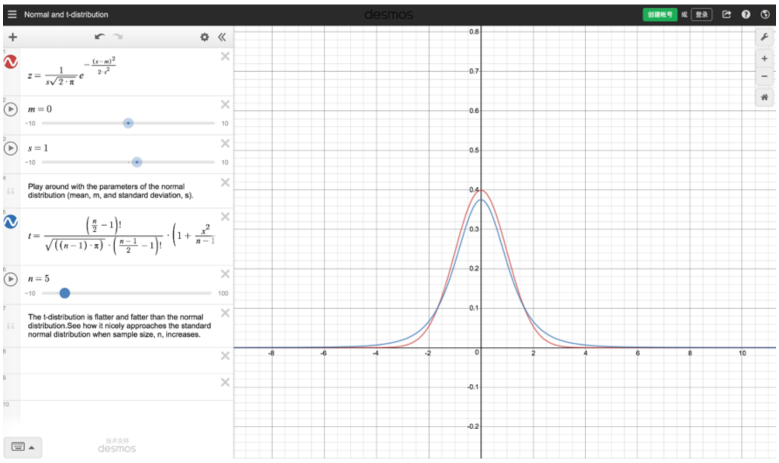
Play around with the parameters of the normal distribution (mean, m, and standard deviation, s).

$t=\frac{\left(\frac{n}{2}-1\right)!}{\sqrt{((n-1)\cdot\pi)}\cdot\left(\frac{n-1}{2}-1\right)!}\cdot\left(1+\frac{x^2}{n-1}\right.$

$n=4$

The t-distribution is flatter and fatter than the normal distribution.See how it nicely approaches the standard normal distribution when sample size, n, increases.

The *t*-distribution with 3 degrees of freedom



Normal and t-distribution

$z=\frac{1}{s\sqrt{2\cdot\pi}}e^{-\frac{(x-m)^2}{2\cdot s^2}}$

$m=0$

$s=1$

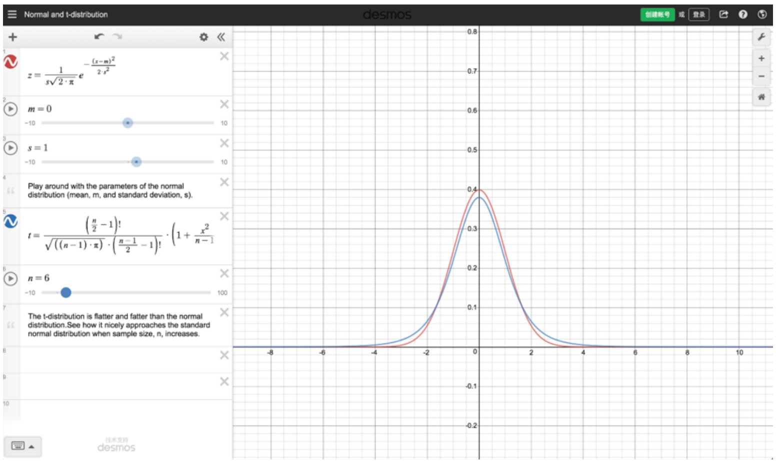
Play around with the parameters of the normal distribution (mean, m, and standard deviation, s).

$t=\frac{\left(\frac{n}{2}-1\right)!}{\sqrt{((n-1)\cdot\pi)}\cdot\left(\frac{n-1}{2}-1\right)!}\cdot\left(1+\frac{x^2}{n-1}\right.$

$n=5$

The t-distribution is flatter and fatter than the normal distribution.See how it nicely approaches the standard normal distribution when sample size, n, increases.

The *t*-distribution with 4 degrees of freedom



Normal and t-distribution

$z=\frac{1}{s\sqrt{2\cdot\pi}}e^{-\frac{(x-m)^2}{2\cdot s^2}}$

$m=0$

$s=1$

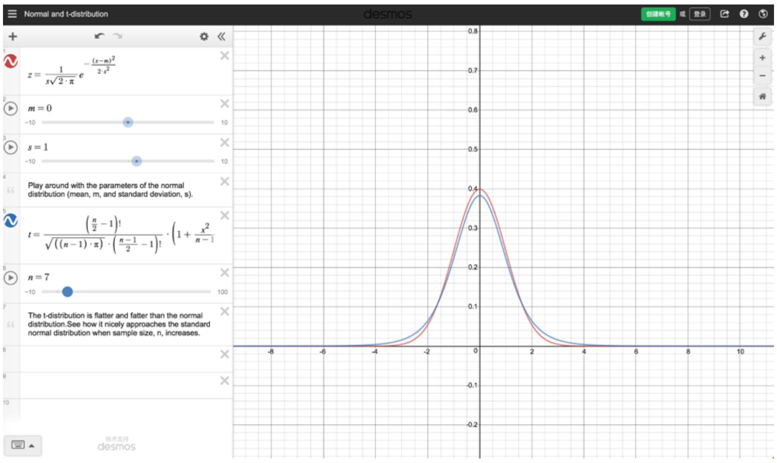
Play around with the parameters of the normal distribution (mean, m, and standard deviation, s).

$t=\frac{\left(\frac{n}{2}-1\right)!}{\sqrt{((n-1)\cdot\pi)}\cdot\left(\frac{n-1}{2}-1\right)!}\cdot\left(1+\frac{x^2}{n-1}\right.$

$n=6$

The t-distribution is flatter and fatter than the normal distribution.See how it nicely approaches the standard normal distribution when sample size, n, increases.

The *t*-distribution with 5 degrees of freedom



$z=\frac{1}{s\sqrt{2\cdot\pi}}e^{-\frac{(x-m)^2}{2\cdot s^2}}$

$m=0$ −10 10

$s=1$ −10 10

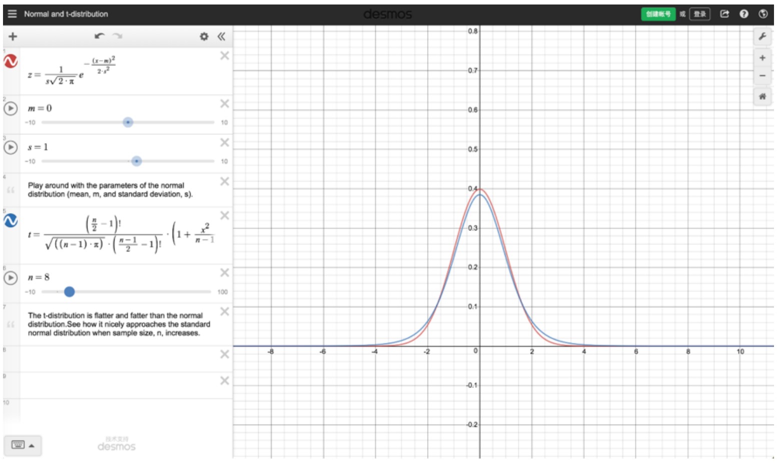
Play around with the parameters of the normal distribution (mean, m, and standard deviation, s).

$t=\frac{\left(\frac{n}{2}-1\right)!}{\sqrt{((n-1)\cdot\pi)}\cdot\left(\frac{n-1}{2}-1\right)!}\cdot\left(1+\frac{x^2}{n-1}\right.$

$n=7$ −10 100

The t-distribution is flatter and fatter than the normal distribution.See how it nicely approaches the standard normal distribution when sample size, n, increases.

The t-distribution with 6 degrees of freedom



$z=\frac{1}{s\sqrt{2\cdot\pi}}e^{-\frac{(x-m)^2}{2\cdot s^2}}$

$m=0$ −10 10

$s=1$ −10 10

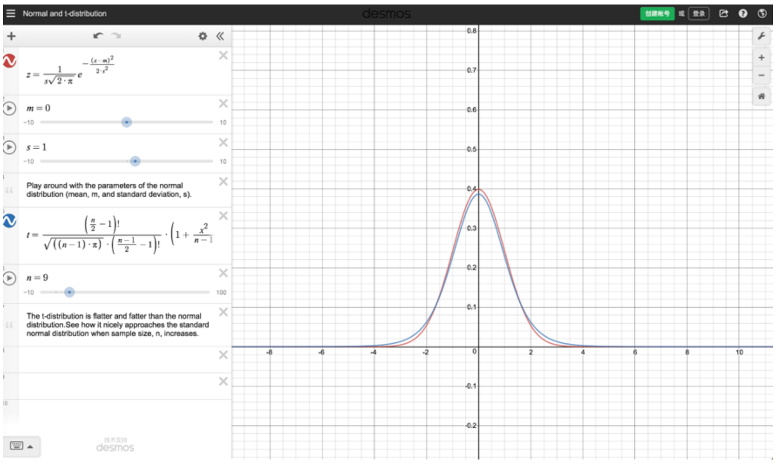
Play around with the parameters of the normal distribution (mean, m, and standard deviation, s).

$t=\frac{\left(\frac{n}{2}-1\right)!}{\sqrt{((n-1)\cdot\pi)}\cdot\left(\frac{n-1}{2}-1\right)!}\cdot\left(1+\frac{x^2}{n-1}\right.$

$n=8$ −10 100

The t-distribution is flatter and fatter than the normal distribution.See how it nicely approaches the standard normal distribution when sample size, n, increases.

The t-distribution with 7 degrees of freedom



$z=\frac{1}{s\sqrt{2\cdot\pi}}e^{-\frac{(x-m)^2}{2\cdot s^2}}$

$m=0$ −10 10

$s=1$ −10 10

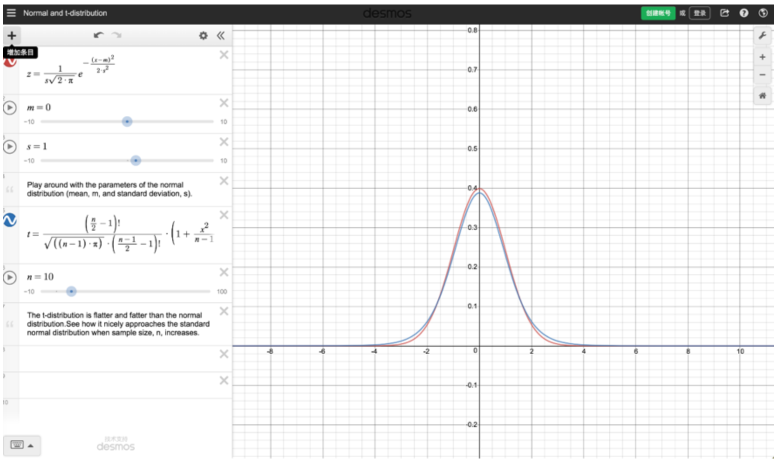
Play around with the parameters of the normal distribution (mean, m, and standard deviation, s).

$t=\frac{\left(\frac{n}{2}-1\right)!}{\sqrt{((n-1)\cdot\pi)}\cdot\left(\frac{n-1}{2}-1\right)!}\cdot\left(1+\frac{x^2}{n-1}\right.$

$n=9$ −10 100

The t-distribution is flatter and fatter than the normal distribution.See how it nicely approaches the standard normal distribution when sample size, n, increases.

The t-distribution with 8 degrees of freedom



$z = \frac{1}{s\sqrt{2\cdot\pi}} e^{-\frac{(x-m)^2}{2\cdot s^2}}$

$m = 0$

$s = 1$

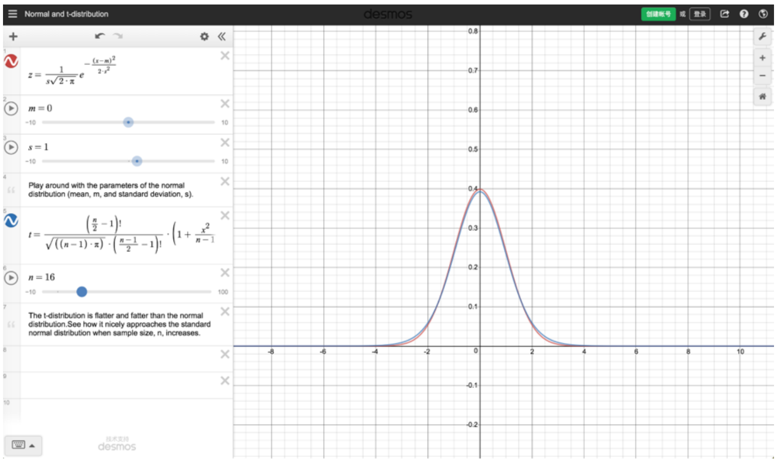
Play around with the parameters of the normal distribution (mean, m, and standard deviation, s).

$t = \frac{\left(\frac{n}{2}-1\right)!}{\sqrt{((n-1)\cdot\pi)}\cdot\left(\frac{n-1}{2}-1\right)!}\cdot\left(1+\frac{x^2}{n-1}\right.$

$n = 10$

The t-distribution is flatter and fatter than the normal distribution.See how it nicely approaches the standard normal distribution when sample size, n, increases.

The *t*-distribution with 9 degrees of freedom



$z = \frac{1}{s\sqrt{2\cdot\pi}} e^{-\frac{(x-m)^2}{2\cdot s^2}}$

$m = 0$

$s = 1$

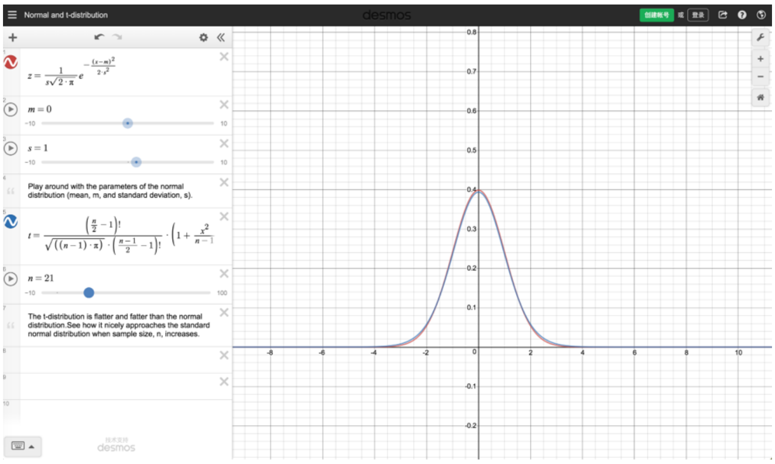
Play around with the parameters of the normal distribution (mean, m, and standard deviation, s).

$t = \frac{\left(\frac{n}{2}-1\right)!}{\sqrt{((n-1)\cdot\pi)}\cdot\left(\frac{n-1}{2}-1\right)!}\cdot\left(1+\frac{x^2}{n-1}\right.$

$n = 16$

The t-distribution is flatter and fatter than the normal distribution.See how it nicely approaches the standard normal distribution when sample size, n, increases.

The *t*-distribution with 15 degrees of freedom



$z = \frac{1}{s\sqrt{2\cdot\pi}} e^{-\frac{(x-m)^2}{2\cdot s^2}}$

$m = 0$

$s = 1$

Play around with the parameters of the normal distribution (mean, m, and standard deviation, s).

$t = \frac{\left(\frac{n}{2}-1\right)!}{\sqrt{((n-1)\cdot\pi)}\cdot\left(\frac{n-1}{2}-1\right)!}\cdot\left(1+\frac{x^2}{n-1}\right.$

$n = 21$

The t-distribution is flatter and fatter than the normal distribution.See how it nicely approaches the standard normal distribution when sample size, n, increases.

The *t*-distribution with 20 degrees of freedom

One more thing needs to be noticed. The shape of the normal distribution is only decided by μ and σ, while the shape of the *t*-distribution is not only decided by the sample mean and sample deviation, but also the degrees of freedom.

In 2002, Ding ([1]) published an asymptotic expansion for the *t*-distribution density. In his report, he concluded that:

“Let $\mathrm{t(x, n)}$ be density function of *t*-distribution function of degree *n*, we obtain, in this paper that:

$$t(x,n) = \emptyset(x) + \frac{1}{n}\psi_1(x) + \frac{1}{n^2}\psi_2(x) + o(n^{-2}),$$

Where $$\emptyset(x) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{x^2}{2}\right)$$

$$\psi_1(x) = \frac{1}{4}(x^4 - 2x^2 - 1)\emptyset(x);$$

$$\psi_2(x) = \frac{1}{96}(3x^8 - 28x^6 + 30x^4 + 12x^2 + 3)\emptyset(x).$$

This expansion is very important in small number sample study. ”

His research provided the relation between the *t*-distribution and standard normal distribution, but leaving the o(n^{-2}) term undiscussed. His asymptotic expansion is only to the n^{-2}term. It is of interest to see the significance in the improvement if the term of $n^{-3}$is included, which is the focus of this study. When the values of *n* and *x* vary, the remainder term will affect the accuracy of the similarity of the two distributions. More specifically, *x* has influence on the similarity, and *n* does as well. The two values, together, influence the result. We cannot deny the influence of the remainder term and get rid of it assuming it is always quite tiny as uninfluential. So, in this report, we will take a detailed look at the remainder term $o(n^{-2})$.

Let $\mathrm{t(x; n)}$ be the density function of *t*-distribution function with *n* degrees of freedom, of which the probability density function (pdf) is

$$t(x;n) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}}\left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}, \text{for } -\infty < x < \infty.$$

We obtain, in this paper that

$$t(x;n) = \varphi(x) + \frac{1}{n}\left(\frac{x^4}{4} - \frac{x^2}{2} - \frac{1}{4}\right)\varphi(x) + \frac{1}{n^2}\left(\frac{x^8}{32} - \frac{7x^6}{24} + \frac{5x^4}{16} + \frac{x^2}{8} + \frac{1}{32}\right)\varphi(x) +$$

$$\frac{1}{n^3}\left(\frac{x^{12}}{384} - \frac{11x^{10}}{192} + \frac{113x^8}{384} - \frac{23x^6}{96} - \frac{11x^4}{128} - \frac{x^2}{64} - \frac{49}{384}\right)\varphi(x) + \frac{1}{6n^2(n+1)}\varphi(x) + o\left(\frac{1}{n^3}\right),$$

where $\varphi(x)$ is the density function of the standard normal distribution. In this report, we also provide numerical estimations for the t density function by including the n^{-3} term and identifying the error caused by the asymptotic expansion by Ding.

2. Logarithmic Expansion of the Function and Estimation of the Remainder Terms

For convenience, we suppose $t(x;n) = \varphi(x)f(x)$ first, thus $f(x) = t(x;n)/\varphi(x)$. By doing division, we can get $f(x)$. But subtraction is easier than division, so we take the logarithm of both sides of the equation and get $ln\, f(x) = ln\ t(x;n) - ln\, \varphi(x)$. Once we have $\ln \mathrm{f(x)}$, it is easy to go back to $\mathrm{f(x)}$ by taking the exponential of $ln\, f(x)$. In this chapter, we deal with $ln\, f(x)$. As long as we know $ln\, f(x)$, we can have the information of $f(x)$ equivalently.

Let us start from $t(x;n) = \varphi(x)f(x)$.

We know that

$$ln\, f(x) = ln \frac{t(x;n)}{\varphi(x)} = ln\, t(x;n) - ln\, \varphi(x)$$

$$= ln \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} - ln \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$= ln\, \Gamma\left(\tfrac{n+1}{2}\right) - ln\, \Gamma\left(\tfrac{n}{2}\right) - \tfrac{1}{2} ln\left(\tfrac{n}{2}\right) + \tfrac{x^2}{2} - \tfrac{n+1}{2} ln(1 + \tfrac{x^2}{n}) \qquad (1)$$

We have the Stirling's Approximation ([2]),

$$ln\, \Gamma(p) = ln\sqrt{2\pi} + \left(p - \frac{1}{2}\right) ln\, p - p + \frac{b_1}{2p} - \frac{b_2}{12p^3} + \frac{\theta b_3}{30p^5}, (0 < \theta < 1)$$

where $b_1 = \frac{1}{6}, b_2 = \frac{1}{30}, b_3 = \frac{1}{42}$ are referred to as Bernoulli numbers ([2]).

According to the above equation, we have

$$ln\, \Gamma\left(\tfrac{n+1}{2}\right) = ln\sqrt{2\pi} + \left(\tfrac{n}{2}\right) ln(\tfrac{n+1}{2}) - \tfrac{n+1}{2} + \tfrac{1}{6(n+1)} - \tfrac{1}{45(n+1)^3} + \tfrac{8\theta_1}{315(n+1)^5},$$

$$ln\, \Gamma\left(\tfrac{n}{2}\right) = ln\sqrt{2\pi} + \left(\tfrac{n-1}{2}\right) ln(\tfrac{n}{2}) - \tfrac{n}{2} + \tfrac{1}{6n} - \tfrac{1}{45n^3} + \tfrac{8\theta_2}{315n^5},$$

where $0 < \theta_1 < 1,\ 0 < \theta_2 < 1$. Then we substitute the logarithms of equation (1), and we can organize the expression in the following way:

$$ln\,\Gamma\left(\frac{n+1}{2}\right) - ln\,\Gamma\left(\frac{n}{2}\right) - \frac{1}{2}ln\left(\frac{n}{2}\right)$$

$$= \left[ln\sqrt{2\pi} + \left(\frac{n}{2}\right)ln\left(\frac{n+1}{2}\right) - \frac{n+1}{2} + \frac{1}{6(n+1)} - \frac{1}{45(n+1)^3} + \frac{8\theta_1}{315(n+1)^5}\right]$$

$$- \left[ln\sqrt{2\pi} + \left(\frac{n-1}{2}\right)ln\left(\frac{n}{2}\right) - \frac{n}{2} + \frac{1}{6n} - \frac{1}{45n^3} + \frac{8\theta_2}{315n^5}\right] - \frac{1}{2}ln\left(\frac{n}{2}\right)$$

$$= \frac{n}{2}ln\left(1+\frac{1}{n}\right) - \frac{1}{2} - \frac{1}{6n(n+1)} + \frac{3n^2+3n+1}{45n^3(n+1)^3} - \frac{8}{315}\left[\frac{\theta_2}{n^5} - \frac{\theta_1}{(n+1)^5}\right] \qquad (2)$$

Since we have the Taylor expansion for the logarithm,

$$ln(1+x) = \sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}x^n = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots, for\ |x| \le 1,$$

then, we can expand the logarithm of the right-hand side of (2), and we get:

$$\frac{n}{2}ln\left(1+\frac{1}{n}\right) - \frac{1}{2} - \frac{1}{6n(n+1)} + \frac{3n^2+3n+1}{45n^3(n+1)^3} - \frac{8}{315}\left[\frac{\theta_2}{n^5} - \frac{\theta_1}{(n+1)^5}\right]$$

$$= \frac{n}{2}\left(\frac{1}{n} - \frac{1}{2n^2} + \frac{1}{3n^3} - \frac{1}{4n^4} + \sum_{k=5}^{\infty}\frac{(-1)^{k-1}}{kn^k}\right) - \frac{1}{2} - \frac{1}{6n(n+1)} + \frac{3n^2+3n+1}{45n^3(n+1)^3}$$

$$- \frac{8}{315}\left[\frac{\theta_2}{n^5} - \frac{\theta_1}{(n+1)^5}\right]$$

$$= \frac{1}{2} - \frac{1}{4n} + \frac{1}{6n^2} - \frac{1}{8n^3} + \sum_{k=5}^{\infty}\frac{(-1)^{k-1}}{2kn^{k-1}} - \frac{1}{2} - \frac{1}{6n(n+1)} + \frac{3n^2+3n+1}{45n^3(n+1)^3} - \frac{8}{315}\left[\frac{\theta_2}{n^5}\right.$$

$$\left. - \frac{\theta_1}{(n+1)^5}\right]$$

$$= -\frac{1}{4n} - \frac{1}{8n^3} + \frac{1}{6n^2(n+1)} + \sum_{k=5}^{\infty}\frac{(-1)^{k-1}}{2kn^{k-1}} + \frac{3n^2+3n+1}{45n^3(n+1)^3} - \frac{8}{315}\left[\frac{\theta_2}{n^5} - \frac{\theta_1}{(n+1)^5}\right] \qquad (3)$$

Expanding the last two terms of (1) using the same method, we obtain

$$\frac{x^2}{2} - \frac{n+1}{2}ln\left(1+\frac{x^2}{n}\right)$$

$$= \frac{x^2}{2} - \frac{n}{2}\left(\frac{x^2}{n} - \frac{x^4}{2n^2} + \frac{x^6}{3n^3} - \frac{x^8}{4n^4}\cdots\right) - \frac{1}{2}\left(\frac{x^2}{n} - \frac{x^4}{2n^2} + \frac{x^6}{3n^3} - \cdots\right)$$

$$= \frac{x^2}{2} - \left(\frac{x^2}{2} - \frac{x^4}{4n} + \frac{x^6}{6n^2} - \frac{x^8}{8n^3}\cdots\right) - \left(\frac{x^2}{2n} - \frac{x^4}{4n^2} + \frac{x^6}{6n^3} - \cdots\right)$$

$$= \frac{1}{n}\left(\frac{x^4}{4}-\frac{x^2}{2}\right)+\frac{1}{n^2}\left(\frac{x^4}{4}-\frac{x^6}{6}\right)+\frac{1}{n^3}\left(\frac{x^8}{8}-\frac{x^6}{6}\right)+\sum_{k=4}^{\infty}\frac{(-1)^{k-1}}{n^k}\left(\frac{x^{2k+2}}{2(k+1)}-\frac{x^{2k}}{2k}\right) \tag{4}$$

We combine (1) ~ (4), we get:

$ln\, f(x)$

$$= -\frac{1}{4n}-\frac{1}{8n^3}+\frac{1}{6n^2(n+1)}+\sum_{k=5}^{\infty}\frac{(-1)^{k-1}}{2kn^{k-1}}+\frac{3n^2+3n+1}{45n^3(n+1)^3}$$

$$-\frac{8}{315}\left[\frac{\theta_2}{n^5}-\frac{\theta_1}{(n+1)^5}\right]+\frac{1}{n}\left(\frac{x^4}{4}-\frac{x^2}{2}\right)+\frac{1}{n^2}\left(-\frac{x^6}{6}+\frac{x^4}{4}\right)$$

$$+\frac{1}{n^3}\left(\frac{x^8}{8}-\frac{x^6}{6}\right)+\sum_{k=4}^{\infty}\frac{(-1)^{k-1}}{n^k}\left(\frac{x^{2k+2}}{2(k+1)}-\frac{x^{2k}}{2k}\right)$$

$$=\frac{1}{n}\left(\frac{x^4}{4}-\frac{x^2}{2}-\frac{1}{4}\right)+\frac{1}{n^2}\left(-\frac{x^6}{6}+\frac{x^4}{4}\right)+\frac{1}{n^3}\left(\frac{x^8}{8}-\frac{x^6}{6}-\frac{1}{8}\right)+\frac{1}{6n^2(n+1)}$$

$$+\sum_{k=5}^{\infty}\frac{(-1)^{k-1}}{2kn^{k-1}}+\frac{3n^2+3n+1}{45n^3(n+1)^3}-\frac{8}{315}\left[\frac{\theta_2}{n^5}-\frac{\theta_1}{(n+1)^5}\right]$$

$$+\sum_{k=4}^{\infty}\frac{(-1)^{k-1}}{n^k}\left(\frac{x^{2k+2}}{2(k+1)}-\frac{x^{2k}}{2k}\right)$$

$$=\frac{1}{n}\left(\frac{x^4}{4}-\frac{x^2}{2}-\frac{1}{4}\right)+\frac{1}{n^2}\left(-\frac{x^6}{6}+\frac{x^4}{4}\right)+\frac{1}{n^3}\left(\frac{x^8}{8}-\frac{x^6}{6}-\frac{1}{8}\right)+\frac{1}{6n^2(n+1)}+o\left(\frac{1}{n^3}\right)$$

$$=\frac{A}{n}+\frac{B}{n^2}+\frac{C}{n^3}+\frac{1}{6n^2(n+1)}+o\left(\frac{1}{n^3}\right) \tag{5}$$

where $A=A(x)=\frac{x^4}{4}-\frac{x^2}{2}-\frac{1}{4}$,

$B=B(x)=-\frac{x^6}{6}+\frac{x^4}{4}$,

$C=C(x)=\frac{x^8}{8}-\frac{x^6}{6}-\frac{1}{8}$.

By far we get the expression for $ln\, f(x)$.

3. Expansion of the Function and Estimation of its Remainder

Since our focus is on $f(x)$ instead of $ln\, f(x)$, we go back to $f(x)$ by $f(x) = e^{ln\, f(x)}$, i.e. $f(x)$ can be expressed in the exponential form.

According to Taylor's Theorem:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots, \quad \text{for } |x| \le 1.$$

Hence,

$$\begin{aligned}
f(x) &= e^{ln\, f(x)} \\
&= e^{\frac{A}{n}+\frac{B}{n^2}+\frac{C}{n^3}+\frac{1}{6n^2(n+1)}+o\left(\frac{1}{n^3}\right)} \\
&= e^{\frac{A}{n}} \cdot e^{\frac{B}{n^2}+\frac{C}{n^3}+\frac{1}{6n^2(n+1)}+o\left(\frac{1}{n^3}\right)} \\
&= (1 + \frac{A}{n} + \frac{A^2}{2n^2} + \frac{A^3}{3!\, n^3} + o(\frac{1}{n^3}))(1 + \frac{B}{n^2} + \frac{C}{n^3} + \frac{1}{6n^2(n+1)} + o(\frac{1}{n^3})) \\
&= 1 + \frac{A}{n} + \frac{1}{n^2}(\frac{A^2}{2} + B) + \frac{1}{n^3}(\frac{A^3}{6} + AB + C) + \frac{1}{6n^2(n+1)} + o(\frac{1}{n^3})
\end{aligned}$$

Then we substitute the coefficients of $\frac{1}{\mathrm{n}}$、$\frac{1}{\mathrm{n}^2}$、$\frac{1}{\mathrm{n}^3}$ by functions of *x* one by one:

$$A = A(x) = \frac{x^4}{4} - \frac{x^2}{2} - \frac{1}{4}$$

$$\begin{aligned}
\frac{A^2}{2} + B &= \frac{1}{2}(\frac{x^4}{4} - \frac{x^2}{2} - \frac{1}{4})^2 + (-\frac{x^6}{6} + \frac{x^4}{4}) \\
&= \frac{1}{2}\left(\frac{x^8}{16} - \frac{x^6}{4} + \frac{x^4}{8} + \frac{x^2}{4} + \frac{1}{16}\right) + \frac{x^4}{4} - \frac{x^6}{6} \\
&= \frac{x^8}{32} - \frac{x^6}{8} + \frac{x^4}{8} + \frac{x^2}{4} + \frac{1}{16} + \frac{x^4}{4} - \frac{x^6}{6} \\
&= \frac{x^8}{32} - \frac{7x^6}{24} + \frac{5x^4}{16} + \frac{x^2}{8} + \frac{1}{32}
\end{aligned}$$

$$\begin{aligned}
\frac{A^3}{6} &= \frac{1}{2}(\frac{x^4}{4} - \frac{x^2}{2} - \frac{1}{4})\left(\frac{x^8}{16} - \frac{x^6}{4} + \frac{x^4}{8} + \frac{x^2}{4} + \frac{1}{16}\right) \\
&= \frac{1}{6}\left(\frac{x^{12}}{64} - \frac{x^{10}}{16} + \frac{x^8}{32} + \frac{x^6}{16} + \frac{x^4}{64} - \frac{x^{10}}{32} + \frac{x^8}{8} - \frac{x^6}{16} - \frac{x^4}{8} - \frac{x^2}{16} - \frac{x^8}{64} + \frac{x^6}{16} - \frac{x^4}{32}\right. \\
&\qquad \left. - \frac{x^2}{16} - \frac{1}{64}\right)
\end{aligned}$$

$$= \frac{1}{6}\left(\frac{x^{12}}{64} - \frac{x^{10}}{32} - \frac{9x^8}{64} + \frac{x^6}{16} - \frac{x^4}{64} - \frac{3x^2}{32} - \frac{1}{64}\right)$$

$$= \frac{x^{12}}{384} - \frac{x^{10}}{64} + \frac{3x^8}{128} + \frac{x^6}{96} - \frac{3x^4}{128} - \frac{x^2}{64} - \frac{1}{384}$$

$$AB = (\frac{x^4}{4} - \frac{x^2}{2} - \frac{1}{4})(-\frac{x^6}{6} + \frac{x^4}{4})$$

$$= \frac{x^8}{16} - \frac{x^6}{8} - \frac{x^4}{16} - \frac{x^{10}}{24} + \frac{x^8}{12} + \frac{x^6}{24}$$

$$= -\frac{x^{10}}{24} + \frac{7x^8}{48} - \frac{x^6}{12} - \frac{x^4}{16}$$

Thus, $\frac{A^3}{6} + AB + C$

$$= \frac{x^{12}}{384} - \frac{x^{10}}{64} + \frac{7x^8}{128} + \frac{x^6}{96} - \frac{3x^4}{128} - \frac{x^2}{64} - \frac{1}{384} - \frac{x^{10}}{24} + \frac{7x^8}{48} - \frac{x^6}{12} - \frac{x^4}{16} + \frac{x^8}{8}$$

$$- \frac{x^6}{6} - \frac{1}{8}$$

$$= \frac{x^{12}}{384} - \frac{11x^{10}}{192} + \frac{113x^8}{384} - \frac{23x^6}{96} - \frac{11x^4}{128} - \frac{x^2}{64} - \frac{49}{384}$$

Therefore, $f(x) = 1 + \frac{1}{n}\left(\frac{x^4}{4} - \frac{x^2}{2} - \frac{1}{4}\right) + \frac{1}{n^2}\left(\frac{x^8}{32} - \frac{7x^6}{24} + \frac{5x^4}{16} + \frac{x^2}{8} + \frac{1}{32}\right)$

$$+ \frac{1}{n^3}\left(\frac{x^{12}}{384} - \frac{11x^{10}}{192} + \frac{113x^8}{384} - \frac{23x^6}{96} - \frac{11x^4}{128} - \frac{x^2}{64} - \frac{49}{384}\right)$$

$$+ \frac{1}{6n^2(n+1)} + o(\frac{1}{n^3})$$

4. Comparison of the Density Function of the T-distribution and its Asymptotic Expansion

Based on the estimation above,

$$\hat{t}(x;n) = \varphi(x)[\frac{R_1(x)}{n} + \frac{R_2(x)}{n^2} + \frac{R_3(x)}{n^3} + R_4(x)],$$

where $R_1(x) = \frac{x^4}{4} - \frac{x^2}{2} - \frac{1}{4}$,

$$R_2(x)=\frac{x^8}{32}-\frac{7x^6}{24}+\frac{5x^4}{16}+\frac{x^2}{8}+\frac{1}{32},$$

$$R_3(x)=\frac{x^{12}}{384}-\frac{11x^{10}}{192}+\frac{113x^8}{384}-\frac{23x^6}{96}-\frac{11x^4}{128}-\frac{x^2}{64}-\frac{49}{384},$$

$$R_4(x)=\frac{1}{6n^2(n+1)}.$$

Let us evaluate the functions and compare their errors. Since $t(x;n)$ contains the Gamma function in its expression, we want to rewrite $t(x;n)$ in an easier way that excel can run.

According to the properties of the gamma function Γ:

$$\Gamma(p+1)=p\Gamma(p),\ \Gamma(1)=1,\ \Gamma(0.5)=\sqrt{\pi}.$$

When *n* is an even number, let $n=2k, k$ is an intger.

$$t(x;n)=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}$$

$$=\frac{\Gamma\left(\frac{2k+1}{2}\right)}{\Gamma(k)\sqrt{2k\pi}}\left(1+\frac{x^2}{2k}\right)^{-\frac{2k+1}{2}}$$

$$=\frac{\Gamma\left(\frac{2k+1}{2}\right)}{\Gamma(k)\sqrt{2k\pi}}\left(1+\frac{x^2}{2k}\right)^{-\frac{2k+1}{2}}$$

$$=\frac{\frac{2k-1}{2}\cdot\frac{2k-3}{2}\cdot\ldots\cdot\frac{1}{2}\cdot\Gamma\left(\frac{1}{2}\right)}{(k-1)!\sqrt{2k\pi}}\left(1+\frac{x^2}{2k}\right)^{-(k+0.5)}$$

$$=\frac{(2k-1)!!\sqrt{2k}}{2^{k+1}k!}(1+\frac{x^2}{2k})^{-(k+0.5)}$$

When *n* is an odd number, let $n=2k+1, k$ is an intger.

$$t(x;n)=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}$$

$$=\frac{\Gamma\left(\frac{2k+2}{2}\right)}{\Gamma\left(\frac{2k+1}{2}\right)\sqrt{(2k+1)\pi}}\left(1+\frac{x^2}{2k+1}\right)^{-\frac{2k+2}{2}}$$

$$=\frac{\Gamma(k+1)}{\Gamma\left(\frac{2k+1}{2}\right)\sqrt{(2k+1)\pi}}\left(1+\frac{x^2}{2k+1}\right)^{-(k+1)}$$

$$= \frac{\Gamma(k+1)}{\frac{2k-1}{2}\cdot\frac{2k-3}{2}\cdot\ldots\cdot\frac{1}{2}\cdot\Gamma\left(\frac{1}{2}\right)\sqrt{(2k+1)\pi}}\left(1+\frac{x^2}{2k+1}\right)^{-(k+1)}$$

$$= \frac{2^k k!\sqrt{2k+1}}{(2k+1)!!\,\pi}(1+\frac{x^2}{2k+1})^{-(k+1)}$$

In conclusion, when *n* is not a large number we can rewrite $t(x;n)$ in such form:

$$t(x;n) = \begin{cases} \frac{(2k-1)!!\sqrt{2k}}{2^{k+1}k!}(1+\frac{x^2}{2k})^{-(k+0.5)}, n = 2k \\ \frac{2^k k!\sqrt{2k+1}}{(2k+1)!!\pi}(1+\frac{x^2}{2k+1})^{-(k+1)}, n = 2k+1 \end{cases} \quad (k = 5,6,\ldots\ldots),$$

Take *n*=10, 20, 36 (*k*=5, 10, 18), results are shown in sheet 1.

Sheet 1

Comparison of Density Function of the T-distribution and its Asymptotic Expansion

x	t(x;10)	$\hat{t}(x;10)$	original estimation	error	original error
0.0	0.3891	0.3891	0.3989	0.0000	0.0253
0.2	0.3807	0.3807	0.3910	0.0000	0.0273
0.4	0.3566	0.3565	0.3683	-0.0001	0.0328
0.6	0.3203	0.3202	0.3332	-0.0003	0.0403
0.8	0.2766	0.2765	0.2897	-0.0005	0.0472
1.0	0.2304	0.2304	0.2420	0.0000	0.0504
1.2	0.1857	0.1862	0.1942	0.0030	0.0459
1.4	0.1454	0.1472	0.1497	0.0121	0.0298
1.6	0.1111	0.1147	0.1109	0.0326	-0.0014
1.8	0.0831	0.0891	0.0790	0.0714	-0.0501
2.0	0.0611	0.0694	0.0540	0.1353	-0.1170
2.2	0.0444	0.0545	0.0355	0.2282	-0.2007
2.4	0.0319	0.0430	0.0224	0.3486	-0.2975
2.6	0.0227	0.0338	0.0136	0.4879	-0.4024
2.8	0.0161	0.0263	0.0079	0.6314	-0.5090

3.0	0.0114	0.0201	0.0044	0.7625	-0.6113
3.2	0.0081	0.0150	0.0024	0.8678	-0.7039
3.4	0.0057	0.0110	0.0012	0.9408	-0.7834
3.6	0.0040	0.0080	0.0006	0.9816	-0.8480
3.8	0.0029	0.0057	0.0003	0.9925	-0.8977
4.0	0.0020	0.0040	0.0001	0.9732	-0.9341
4.2	0.0015	0.0028	0.0001	0.9183	-0.9594

x	t(x;20)	$\hat{t}(x;20)$	original estimation	error	original error
0.0	0.3940	0.3940	0.3989	0.0000	0.0126
0.2	0.3858	0.3858	0.3910	0.0000	0.0136
0.4	0.3624	0.3624	0.3683	0.0000	0.0163
0.6	0.3267	0.3267	0.3332	-0.0001	0.0200
0.8	0.2830	0.2830	0.2897	-0.0001	0.0235
1.0	0.2360	0.2360	0.2420	0.0000	0.0251
1.2	0.1899	0.1900	0.1942	0.0007	0.0228
1.4	0.1476	0.1480	0.1497	0.0028	0.0142
1.6	0.1112	0.1121	0.1109	0.0078	-0.0028
1.8	0.0814	0.0829	0.0790	0.0174	-0.0305
2.0	0.0581	0.0601	0.0540	0.0340	-0.0705
2.2	0.0405	0.0429	0.0355	0.0599	-0.1236
2.4	0.0276	0.0303	0.0224	0.0970	-0.1895
2.6	0.0185	0.0212	0.0136	0.1457	-0.2667
2.8	0.0122	0.0147	0.0079	0.2047	-0.3526
3.0	0.0080	0.0101	0.0044	0.2704	-0.4435
3.2	0.0051	0.0069	0.0024	0.3378	-0.5353
3.4	0.0033	0.0046	0.0012	0.4009	-0.6239
3.6	0.0021	0.0030	0.0006	0.4539	-0.7054
3.8	0.0013	0.0020	0.0003	0.4916	-0.7771
4.0	0.0008	0.0012	0.0001	0.5097	-0.8373

4.2	0.0005	0.0008	0.0001	0.5042	-0.8856

x	t(x;36)	$\hat{t}(x;36)$	original estimation	error	original error
0.0	0.3962	0.3962	0.3989	0.0000	0.0070
0.2	0.3881	0.3881	0.3910	0.0000	0.0075
0.4	0.3650	0.3650	0.3683	0.0000	0.0090
0.6	0.3296	0.3296	0.3332	0.0000	0.0111
0.8	0.2860	0.2860	0.2897	0.0000	0.0130
1.0	0.2386	0.2386	0.2420	0.0000	0.0139
1.2	0.1918	0.1918	0.1942	0.0002	0.0126
1.4	0.1486	0.1487	0.1497	0.0009	0.0077
1.6	0.1112	0.1114	0.1109	0.0024	-0.0022
1.8	0.0804	0.0809	0.0790	0.0053	-0.0186
2.0	0.0564	0.0570	0.0540	0.0106	-0.0429
2.2	0.0384	0.0391	0.0355	0.0191	-0.0763
2.4	0.0254	0.0262	0.0224	0.0320	-0.1195
2.6	0.0164	0.0172	0.0136	0.0500	-0.1726
2.8	0.0103	0.0111	0.0079	0.0738	-0.2351
3.0	0.0064	0.0070	0.0044	0.1031	-0.3057
3.2	0.0039	0.0044	0.0024	0.1372	-0.3825
3.4	0.0023	0.0027	0.0012	0.1743	-0.4628
3.6	0.0013	0.0016	0.0006	0.2119	-0.5437
3.8	0.0008	0.0010	0.0003	0.2470	-0.6223
4.0	0.0004	0.0006	0.0001	0.2760	-0.6958
4.2	0.0002	0.0003	0.0001	0.2952	-0.7621

From sheet 1 we have the following conclusions:

1. When x<1.2, the error is almost negligible. It is obvious that the accuracy of my approximation is better than that of Ding’s by comparison between the error and the original error.

2. When n=20, the error is always less than the original error. It the same case for n=36. But for n=10, when $1.6 \le x \le 4.2$, the accuracy of my approximation is worse than that of Ding's. It needs further exploration that why the more exquisite expansion leads to a worse approximation.

5. Conclusion

This report provides a more accurate approximation relation between the *t*-distribution and the standard normal distribution. The *t*-distribution is usually regraded as a standard normal distribution when the degrees of freedom go to infinity. In Bangjun Ding's report, he provided the asymptotic expansion for the *t*-distribution, asymptotic expansion for the *t*-distribution density, we get the expansion of the *t*-distribution which is very important in small number sample study. His report only provides the asymptotic expansion for the *t*-distribution to the n^{-2} term. Based on his report, I furthered the asymptotic expansion to include the significant term of order n^{-3} and provided a numerical analysis of the asymptotic expansion, which are my main contributions to the study. It would help people avoid mistakenly replacing the *t*-distribution with standard normal distribution when the level of accuracy is not good. But the work is not done perfectly as we can see from the discussion in chapter 4, for cases when $1.6 \le x \le 4.2$ and n=10, the accuracy of my approximation is worse than that of Ding's. The error cannot be ignored. Further research is needed. Maybe in future we can get a form solution to simplify the remainder term instead of splitting in into several cases.

Basic Concepts:

1. If the random variable $X \sim N(0,1)$ (standard normal distribution), then its pdf is $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, and its CDF(cumulative distribution function) is $\Phi(x) = P(X \le x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt, for\ -\infty < x < \infty$.

2. If the random variables X_1 and X_2 are independent, and $X_1 \sim N(0,1)$, $X_2 \sim \chi^2(n)$,

(χ^2 distribution), then we say $T = \frac{X_1}{\sqrt{X_2/n}}$ is the *t*-distribution with *n* degrees of freedom. Its pdf is $t(x;n) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}, for\ -\infty < x < \infty, n \in N^*$.

3. Γ function $\Gamma(z) = \int_0^{\infty} t^{z-1}e^{-t}dt$ (real number z *>0*)

Properties:

(1) $\Gamma(1) = 1,\ \Gamma(0.5) = \sqrt{\pi}$;

(2) $\Gamma(z+1) = z\Gamma(z)$; when $x = n$ is a natural number, $\Gamma(n+1) = n\Gamma(n) = n!$.

4. Stirling's Approximation

$$ln\,\Gamma(z) = ln\sqrt{2\pi} + \left(z - \frac{1}{2}\right) ln\,z - z + \frac{B_1}{1\cdot 2}\cdot\frac{1}{z} - \frac{B_2}{3\cdot 4}\cdot\frac{1}{z^3} + \cdots + \frac{(-1)^{m-1}B_m}{(2m-1)\cdot 2m}$$

$$\cdot\frac{1}{z^{2m-1}}\ (0 < \theta < 1)$$

Proof (too complicated) refer to [2] Grigorii Mikhailovich Fichtenholz, Differential and Integral Calculus, Higher Education Press, 2006: 656-659 (41) .

Where B_i are called Bernoulli numbers.

5. Bernoulli numbers（refer to [2] Grigorii Mikhailovich Fichtenholz, Differential and Integral Calculus, Higher Education Press, 2006: 656-659 (41)）

Define: $\frac{z}{e^z-1} = \frac{1}{1+\frac{z}{2!}+\frac{z^2}{3!}+\cdots+\frac{z^{n-1}}{n!}+\cdots} = \frac{1}{\sum_{n=1}^{\infty}\frac{z^{n-1}}{n!}} = \sum_{n=0}^{\infty}\frac{\beta_n}{n!}z^n, z \in \mathbb{C}$.

Since $\left(\sum_{n=1}^{\infty}\frac{z^{n-1}}{n!}\right)\left(\sum_{n=0}^{\infty}\frac{\beta_n}{n!}z^n\right) = 1$, we have $\beta_0 = 1$, and when $n \geq 1$ we can get:

$\frac{\beta_n}{n!1!} + \frac{\beta_{n-1}}{(n-1)!2!} + \cdots + \frac{\beta_{n-k}}{(n-k-1)!k!} + \cdots + \frac{\beta_1}{1!n!} + \frac{\beta_0}{0!(n+1)!} = 0$, and we can solve β_n one by one.

For example, let $n = 1$, $\frac{\beta_1}{1!1!} + \frac{\beta_0}{0!(1+1)!} = 0$, we can solve $\beta_1 = -\frac{1}{2}$; from$\frac{\beta_2}{2!1!} + \frac{\beta_1}{1!2!} + \frac{\beta_0}{0!(2+1)!} = 0$, we can solve $\beta_2 = \frac{1}{6}$; so on, …, $\beta_3 = 0, \beta_4 = -\frac{1}{30}, \beta_5 = 0, \beta_6 = \frac{1}{42}, \beta_7 = 0, \beta_8 = -\frac{1}{30}, \ldots.$

It can be proved that, except for $\beta_1 = -\frac{1}{2}$, all the $\beta_{2n+1} = 0$ (n ≥ 1). As for all $\beta_{2n} = 0$ $(n \geq 1)$, we denote $\beta_n = (-1)^{n-1}\beta_{2n}$, they are called Bernoulli numbers. $\beta_1 = \frac{1}{6}, \beta_2 = \frac{1}{30}, \beta_3 = \frac{1}{42}, \beta_4 = \frac{1}{30} \cdots$

6. The *t*-distribution and its difference from standard normal distribution

（1）the *t*-distribution and standard normal distribution look alike in shape, and both are symmetric about y-axis;

（2）the *t*-distribution has a relatively lower peak than standard normal distribution and a thicker tail. P289 graph 5.4.3.

Refer to Mao Shisong, Probability Theory and Mathematical Statistics, second edition.

*though the *t*-distribution and standard normal distribution has tiny difference in shape (especially for $n \geq 30$), they are essentially different.

For naturally generated, large population data, according to the central limit theorem, normal distribution is useful. But for data (usually not large population) generated from manmade experiments, the *t*-distribution is more suitable, especially there is difference of the probability on the tails.

References:

[1] Ding bangjun, Asymptotic Expansion for T-distribution Density, Mathematical Statistics and Applied Probability. 2002,13(4):307-311.

[2] Grigorii Mikhailovich Fichtenholz, Differential and Integral Calculus, Higher Education Press, 2006:656-659 (41).

[3] Mao Shisong, Probability Theory and Mathematical Statistics, second edition.

[4] Journal of Chemometrics, Volume: 29, Issue: 9, Pages: 481-483, First published: 07 June 2015, DOI: (10.1002/cem.2713)