Science and Technology, 6 (2001) 81-87

Approximating the Tail Probability of the t-Distribution: A Bayesian Approach

Mohammad Fraiwan Al-Saleh

Department of Mathematics and Statistics, College of Science, Sultan Qaboos University, P.O. Box 36, Al Khod 123, Muscat, Sultanate of Oman.

تقريب الاحتمال الذيلي لتوزيع ت: أسلوب بييز

محمد فريوان الصالح

خلاصة : لقد تم استخدام أسلوب بييز في تقريب الاحتمال الذيلي لتوزيع ت ، حيث تم ايجاد مجموعة من الحدود الدنيا والعليا لهذا الاحتمال . وبالاعتماد على بساطة هذه الحدود ودقتها ، يمكن القول بان هذه الحدود مناسبة للاستعمال . وقد تمت مقارنة بعض هذه الحدود بما هو متوفر من تقريبات . و قد تم بحث إمكانية استخدام هذا الأسلوب لتوزيعات مهمة أخرى.

ABSTRACT: A Bayesian technique is used to approximate the tail probability of the t-distribution. A set of upper and lower bounds are obtained for this probability. Based on their simplicity and accuracy, these bounds are very adequate to use. Some members of these bounds are compared to some existing approximations. The possibility of using this new procedure for some other distributions is explored.

KEYWORDS: Mill's Ratio, Normal Distribution, t-Distribution and Tail Probability.

1. Introduction

If $X_1, X_2, ...X_n$ is a random sample from the same normal distribution with mean θ and variance σ^2, both being finite but unknown, and if $\bar{X} = \sum(X_i/n)$ and $S^2 = \sum(X_i - \bar{X})^2/(n-1))$, then the statistic $\sqrt{n}(\bar{X} - \theta)/S$ has a t-distribution. This statistic is very useful in the construction of tests and confidence intervals of θ. For a brief recent description of this distribution and its properties, see Stuart and Ord (1994) and Johnson, *et al*. (1995).

The importance of approximating the tail probability of this distribution is due to the fact that this probability is frequently used in constructing confidence intervals or in finding the p-values of some statistical tests. There has been intensive work on approximating the t-distribution which produced approximations of very high accuracy, though some times very complicated. Fisher (1935) gave a direct expansion of the probability density function and hence of $F(t;\nu) = P(t_\nu \le t)$ as a series in ν^{-1}, where ν is the degree of freedom. Elfving (1955) suggested the following approximation:

$$P(t_\nu \le t) = \Phi(\sigma t) + \tfrac{5}{96}(\sigma t^5 \nu^{-2}(1 + .5t^2\nu^{-1})^{-.5(\nu+4)} - \phi(\frac{\sigma t}{\sqrt{2}}))$$

where $\sigma = (\frac{\nu - .5}{\nu + .5t^2})^{.5}$, Φ and ϕ are respectively, the cumulative distribution and the density function of the standard normal distribution. Cucconi (1962) obtained the following approximation:

$$t_{\nu,.975} \approx 1.96\nu(\nu^2 - 3.185\nu + 1.696)^{-.5} \; for \; \nu > 1$$

$$t_{\nu,.995} \approx 2.5758\nu(\nu^2 - 3.185\nu + 4.212)^{-.5} \; for \; \nu > 2,$$

where $t_{v,\alpha}$ is a number with tail area (probability) (1-α). Pinkham and Wilk (1954) suggested the use of the expansion:

$$\int_t^{\infty}(1+y^2v^{-1})^{-.5(v+1)}dy=\sum_{i=1}^{m-1}w_i+R_m(t);\ m<.5(v+1),$$

where $w_1=v(v-1)^{-1}t^{-1}(1+t^2v^{-1})^{-.5(v-1)}$; $w_{i+1}=w_i(1+t^2v^{-1})\dfrac{2i-1}{2i-v+1}$; $i=1,2,...,m-1$, and R_m is the remainder term.

Abu-Dayyeh and Ahmed (1993) considered a similar problem. They provided an upper bound for Mill's ratio $R(x)=(1-\Phi(x))/\phi(x)$. They showed that $R(x)\le(\max(x,x^2)+\frac{2}{\pi})^{-.5}$. Al-Saleh (1994) initiated a Bayesian approach to approximate Mill's ratio and hence to approximate the tail probability of the standard normal distribution. He obtained a sequence of upper bounds and a sequence of lower bounds of $R(x)$, each converges to $R(x)$.

As mentioned by Johnson *et al*. (1995), the available tables of the t-distribution are more than sufficient for almost all applications. However, a major concern raised by the above authors, is how to quickly evaluate the tail probability. It is well known that the t-distribution converges to the normal distribution as v goes to infinity. Thus, for $v\ge30$, if $F(x)$ is the distribution function of the t-distribution then $F(x)\approx\Phi(x)$. However this approximation is not so accurate for small v. Recently, Li and Moor (1999) suggested the approximation of $F(x)$ by $\Phi(\lambda x)$, where λ is a shrinkage factor and its value is given by

$$\lambda=\frac{4v+x^2-1}{4v+2x^2}.$$

This approximation has a much simpler form and very accurate when compared with many of the approximations listed in Johnson *et al*. (1995). A weak point of this approximation is that it is written in terms of Φ, which has no closed form and has to be obtained from tables. For other approximations of the t-distribution see Johnson *et al*. (1995) and Gleason (2000).

In this paper we use the Bayesian approach introduced by Al-Saleh (1994) to obtain new approximations of the tail probability of the t-distributions. These approximations, which turned out to be of a simple form, can give very accurate values. The possibility of applying this approach to some other distributions is discussed.

2. Derivation of the Bounds

Assume that X is a random variable, which has a t- density with parameters θ and v. Then the density of X is

$$g(x;\theta,v)=\frac{\Gamma(.5(v+1))}{\Gamma(.5v)\sqrt{\pi v}}\frac{1}{(1+\frac{1}{v}(x-\theta)^2)^{.5(v+1)}},$$

where v is a positive number and x is any real number. Assume further that the median θ is positive. Let $\pi(\theta)$ be an improper uniform prior of θ, defined by $\pi(\theta)=1$ for $\theta>0$ and zero otherwise. Then the posterior density of θ for given x can be written as:

$$\pi(\theta\mid x)=\frac{f(x-\theta)}{\int_0^{\infty}f(x-\theta)d\theta}=\frac{f(x-\theta)}{F(x)},$$

where $f(.)$ stands for the t density with parameter ν and zero median and F stands for the corresponding cumulative distribution. The main object of this paper is to approximate the tail probability, $1-F(x)$.

Now, if ν=1, the distribution function of the t-distribution is the same as that of the Cauchy which has a closed form. For $\nu \geq 2$, the posterior expected value of θ^k is finite for $k = 1,\cdots, \nu-1$ and is given by:

$$E(\theta^k \mid x) = \frac{1}{F(x)}\int_0^{\infty} \theta^k f(x-\theta)d\theta = \frac{1}{F(x)}\int_{-\infty}^{x} (x-t)^k f(t)dt$$

and hence,

$$E(\theta^k \mid -x) = \frac{(-1)^k}{1-F(x)}\int_x^{\infty} (x-t)^k f(t)dt \tag{1}$$

Since $\theta > 0$, we have $E(\theta^k \mid -x) > 0$ for all values of x. Thus,

$$\int_x^{\infty} (x-t)^k f(t)dt > 0 \qquad \text{for even values of } k,$$

and

$$\int_x^{\infty} (x-t)^k f(t)dt < 0 \qquad \text{for odd values of } k.$$

Now, the last integral can be written as:

$$\int_x^{\infty} (x-t)^k f(t)dt = \sum_{i=0}^{k} (-1)^i \binom{k}{i} x^{k-i}\mu_i(x) \tag{2}$$

where

$$\mu_i(x) = \int_x^{\infty} t^i f(t)dt.$$

Integrating by parts, it can be shown that for $i = 2,\cdots,k;\ k \leq \nu-1$ and $\nu \geq 2$ we have

$$\mu_i(x) = \frac{\nu-1}{\nu-i}x^{i-1}\mu_1(x) + \frac{\nu(i-1)}{\nu-i}\mu_{i-2}(x) \tag{3}$$

where,

$$\mu_0(x) = 1-F(x)\ ;\ \mu_1(x) = \frac{\nu}{\nu-1}(1+\tfrac{x^2}{\nu})f(x).$$

Thus, using (3), upper and lower bounds can be obtained for the tail probability of the standard t-distribution, i.e. for the quantity $\mu_0(x) = 1-F(x)$.

For i even, μ_i can be written as

$$\mu_i(x) = A_i(x)\mu_1(x) + B_i\mu_0(x),$$

where,

$$A_i(x) = a_i(x) + \sum_{k=2}^{i/2}(a_{2k-2}(x)\prod_{j=k}^{i/2} b_{2j});\ B_i = \prod_{j=1}^{i/2} b_{2j};$$

$$a_i(x) = \frac{v-1}{v-i}x^{i-1} \text{ and } b_i = \frac{(i-1)v}{v-i}.$$

For i odd, $\mu_i(x)$ can be written as

$$\mu_i(x) = C_i(x)\mu_1(x),$$

where

$$C_i(x) = a_i(x) + \prod_{j=1}^{(i-1)/2} b_{2j+1} + \sum_{k=1}^{(i-3)/2} (a_{2k+1}(x) \prod_{j=k}^{(i-3)/2} b_{2j+3}).$$

Thus, for $k < v$ with $B_0 = C_1 = 1$ and $A_0 = 0$ we have

$$\int_x^\infty (x-t)^k f(t)dt = \mu_0(x)\sum_{i\,even}\binom{k}{i}x^{k-i}B_i - \mu_1(x)(\sum_{i\,odd}\binom{k}{i}x^{k-i}C_i(x) - \sum_{i\,even}\binom{k}{i}x^{k-i}A_i(x)).$$

Hence, for k even we have:

$$\mu_0(x) = 1 - F(x) \geq L_k(x) = \frac{\mu_1(x)(\sum_{i\,odd}\binom{k}{i}x^{k-i}C_i(x) - \sum_{i\,even}\binom{k}{i}x^{k-i}A_i(x)}{\sum_{i\,even}\binom{k}{i}x^{k-i}B_i} \quad (4)$$

while for k odd we have

$$\mu_0(x) = 1 - F(x) \leq U_k(x) = \frac{\mu_1(x)(\sum_{i\,odd}\binom{k}{i}x^{k-i}C_i(x) - \sum_{i\,even}\binom{k}{i}x^{k-i}A_i(x))}{\sum_{i\,even}\binom{k}{i}x^{k-i}B_i} \quad (5)$$

where $L_k(x)$ is the k^{th} lower bound of $1-F(x)$ and $U_k(x)$ is k^{th} upper bound of $1-F(x)$.

3. Numerical Calculations of $L_k(x)$ and $U_k(x)$

To see how accurate $L_k(x)$ and $U_k(x)$ are, the two bounds have been obtained for some values of v and k. For v=10 and k=3, 4, the two consecutive bounds are:

$$L_4(x) = \left(\frac{.8929x^3 + 5.8036x}{x^4 + 7.5x^2 + 6.25}\right)\mu_1(x) \;\; ; \; U_3(x) = \left(\frac{.9107x^2 + 2.8571}{x^3 + 3.75x}\right)\mu_1(x).$$

For v=15, the two bounds are:

$$L_4(x) = \left(\frac{.9324x^3 + 5.4944x}{x^4 + 6.9231x^2 + 4.7203}\right)\mu_1(x) \; ; \; U_3(x) = \left(\frac{.9359x^2 + 2.5}{x^3 + 3.4615x}\right)\mu_1(x).$$

And for v=20, we have:

$$L_4(x) = \left(\frac{.9498x^3 + 5.553x}{x^4 + 6.6667x^2 + 4.1667}\right)\mu_1(x) \; ; \; U_3(x) = \left(\frac{.9510x^2 + 2.3529}{x^3 + 3.333x}\right)\mu_1(x)..$$

Here,

$$\mu_1 = f(x)(1+\tfrac{x^2}{v})\tfrac{v}{v-1} \text{ and } f(x) = \frac{\Gamma(.5(v+1))}{\Gamma(.5v)\sqrt{\pi v}} \frac{1}{(1+\frac{1}{v}x^2)^{.5(v+1)}}.$$

For a given v and suitable k, we take the average of the two bounds as an approximation of the tail probability $1-F(x)$, i.e. for k even

$$1-F(x) \approx \alpha_k^*(x) = \frac{L_k(x)+U_{k-1}(x)}{2} \tag{6}$$

and for k odd we have

$$1-F(x) \approx \alpha_k^*(x) = \frac{L_{k-1}(x)+U_k(x)}{2} \tag{7}$$

Table 1: Values of $\alpha_4^*(x), \alpha_1, \alpha_2$ and α_3

v	x	Exac α	$\alpha_4^*(x)$	α_1	α_2	α_3
10	1.812	.050	.0505	.0583	.0639	.0500
	2.228	.025	.0252	.0288	.0309	.0249
	2.764	.010	.0100	.0093	.0123	.0098
	3.169	.005	.0051	.0041	.0063	.0048
15	1.753	.050	.0506	.0595	.0609	.0500
	2.131	.025	.0252	.0310	.0289	.0250
	2.602	.010	.0101	.0132	.0111	.0100
	2.947	.005	.0050	.0054	.0054	.0050
20	1.725	.050	.0505	.0599	.0610	.0500
	2.086	.025	.0252	.0318	.0288	.0250
	2.528	.010	.0101	.0136	.0111	.0100
	2.845	.005	.0050	.0063	.0054	.0050

$\alpha_4^*(x)$ is compared to the approximations provided by Elfving (1955), Pinkham and Wilk (1954), and Li and Moor (1999) denoted by α_1, α_2 and α_3 respectively. Table 1 contains the values of $\alpha_4^*(x), \alpha_1, \alpha_2$ and α_3 for selected values of v and x. The values of x are those values that are used frequently in applications, i.e. values that correspond to exact tail values of .0500, .0250, .0100, and .0050. It can be seen from this table that the value of $\alpha_4^*(x)$ is very accurate and closer to the exact value than the first two approximations. Furthermore, the values of $\alpha_4^*(x)$ are almost as accurate as the values of α_3. Note that more accurate bounds can be obtained using higher values of k. For example if we take k=5, then

$$U_5 = \frac{.9071x^4 + 10.7321x^2 + 22.8571}{x^5 + 12.5x^3 + 31.25x}\mu_1.$$

Table 2 contains the values of $\alpha_5^*(x), \alpha_1, \alpha_2$ and α_3 for selected values of x when v=10. It can be concluded from this table that the values of $\alpha_5^*(x)$ are even more accurate than α_3.

Table 2: Values of $\alpha_5^*(x), \alpha_1, \alpha_2$ and α_3

v	x	Exac α	$\alpha_5^*(x)$	α_1	α_2	α_3
10	1.812	.050	.0501	.0583	.0639	.0500
	2.228	.025	.0250	.0288	.0309	.0249
	2.764	.010	.0100	.0093	.0123	.0098
	3.169	.005	.0050	.0041	.0063	.0048

4. Other Applications of the Technique

The Bayesian approach, which is used in this paper to approximate the t-distribution, was used by the author to approximate the normal distribution. An inspection of the procedure reveals that it can be applied to some other distributions.

If X has a density $f(x)$ that is symmetric around zero and if we let $Y = X - \theta$, where θ is a location parameter then the density of Y is $f(y-\theta)$. If we impose a uniform prior on θ of the type $\pi(\theta) = 1$ for $\theta \geq 0$ and zero otherwise, then the posterior density of θ given x is

$$\pi(\theta \mid x) = \frac{f(x-\theta)}{\int_0^\infty f(x-\theta)d\theta} = \frac{f(x-\theta)}{F(x)}.$$

All moments of this density are nonnegative and hence as in section (2), it can be shown that

$$\int_x^\infty (x-t)^k f(t)dt = \sum_{i=0}^{k} (-1)^i \binom{k}{i} x^{k-i} \mu_i(x)$$

where

$$\mu_i(x) = \int_x^\infty t^i f(t)dt.$$

Now, depending on the functional form of $f(x)$, it may be possible to obtain a recursive formula for $\mu_i(x)$ like the one in equation (3).

We believe that some distributions such as the lognormal, non-central t and other location types-distribution can benefit from this procedure. Another useful application of the procedure is for estimating the cumulative distribution of the bivariate normal and other bivariate distributions.

5. Concluding Remarks

There has been considerable work on the possible approximations of the tail probability of the t-distribution. Simplicity as well as accuracy are important factors in assessing the value of an approximation. In this paper, we use a Bayesian approach to provide a set of upper and lower bounds of this probability; the set consists of $[v-1]$ members. Any member of the set or a combination of members can serve as an approximation. Taking the average of two consecutive lower and upper bounds can be a good choice. It turns out that this approach is a suitable one in providing simple and accurate approximations and can be used for similar problems. Unlike many other approximations, the current procedure doesn't depend on $\Phi(x)$.

6. Acknowledgment

I wish to thank the referees for their constructive comments and suggestions.

References

ABU-DAYYEH, W. and AHMED, M. 1993. Some new bounds on the tail probability of standard normal distribution. *Journal of Information and Optimization Sciences* **14:** 155-159.

AL-SALEH, M. FRAIWAN.1994. Mill's ratio: a Bayesian approach. *Pakistan Journal of Statistics* **10:**629-632.

CUCCONI, O. 1962. On simple relation between the number of degrees of freedom and the critical value of student- t. *Memeorie Academia Patavina* **74:**179-187.

ELFVING, G. 1955. An expansion principle for distribution functions with application to statistics. *Annals Academiae Scientiarum Fennicae*, **series A 204:**1-8.

FISHER, R.A. 1935. The mathematical distributions used in the common tests of significance. *Econometrica* **3:**353-365.

GLEASON, J.R. 2000. A note on a proposed student t approximation. *Computational statistics and data analysis* **34:**63-66.

JOHNSON, N., KOTZ, S. and BALAKRISHNAN, N.1995. *Continuous Univariate Distribution*. John Wiley and sons, New York.

LI, B. and MOOR, B. 1999. A corrected normal approximation for the student t distribution. *Computational Statistics and Data Analysis* **29:**213-216.

PINKHAM, R. and WILK, M. 1954. Tail areas of the t-distribution from a Mills-ratio-like expansion. *Annals of Mathematical Statistics* **34:**335-337.

STUART, A. and ORD, J. 1994. *Kendall Advanced Theory of Statistics*. Edward Arnold, London

---

Received 5 January 2000
Accepted 22 January 2001