International Journal of Statistika and Mathematika, ISSN: 2277- 2790 E-ISSN: 2249-8605, Volume 8, Issue 1, 2013 pp 19-21

Approximations to the t- distribution

Ramu Yerukala[*], Naveen Kumar Boiroju[#], M. Krishna Reddy[#]

[*]Anurag Group of Institutions, Hyderabad, Andhra Pradesh, INDIA.

[#]Department of Statistics, Osmania University, Hyderabad, Andhra Pradesh, INDIA.

Corresponding Addresses:

ramu20@gmail.com, nanibyrozu@gmail.com, reddymk54@gmail.com

Research Article

Abstract: This paper deals with the approximation of cumulative distribution function (CDF) of t-distribution developed using neural networks and the same is compared with the existing functions in the literature. Microsoft Excel function TDIST( ) taken as the benchmark function to compare with the proposed and the existing methods. The best approximation for t-distribution is selected on the basis of minimum error between the selected approximation function and TDIST( ) function. The accuracy of the proposed functions guaranteed that up to three decimal points.

Key Words: Student's t distribution, Neural networks, Error.

1. Introduction

Let the n independent observations $X_1, X_2, \ldots, X_n$ are taken from normal distribution with expected value μ and standard deviation σ, then, $\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}$ (with $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$) has a standard normal distribution. This statistic can be used in the construction of test and confidence intervals relating to the value of μ, provided σ be known. If we estimate σ^2 by the sample variance, $s^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$ then the resulting statistic $t=\frac{\sqrt{n}(\bar{X}-\mu)}{s}$ no longer has a normal distribution. It has a t-distribution on $\nu=n-1$ degrees of freedom. In 1908, W. S. Gosset ('Student') obtained the distribution of $t=\frac{\sqrt{n}(\bar{X}-\mu)}{s}$ and gave a short table of its cumulative distribution.

The probability density function of student's t-distribution is given by

$$f(t)=\frac{1}{\sqrt{\nu}\,B\left(\frac{1}{2},\frac{\nu}{2}\right)}\frac{1}{\left(1+\frac{t^2}{\nu}\right)^{\frac{\nu+1}{2}}};-\infty<t<\infty \quad (1)$$

Cumulative probability distribution function (CDF) is

$$F(t)=\int_{-\infty}^{t}f(x)dx \quad (2)$$

There is no closed form representation of the CDF of t distribution. In view of this handicap, often one has to refer to standard tables for its evaluation. However such tables are cumbersome and do not always suffice (Roy and Choudhury, 2012). Hence, an approximate formula can be used instead of the tables. A large body of literature on approximations is summarized by Johnson et al. (1995) and Brophy (1987).These approximations often play important roles in statistical inference and computations. The aim of this paper is to evaluate the approximations of CDF of t-distribution to determine their accuracy. Consistent with the symmetric property of this distribution, we have restricted our comparison to x > 0. The rest of the paper is organized as follows. Section 2 reviews previous works related to approximation of CDF of t-distribution. Two new functions for the approximation to CDF of t-distribution using feed forward neural networks are proposed in Section 3. Conclusion presented in Section 4.

2. Review of approximation of the CDF of t-distribution

For the Student's t cumulative distribution function $F(x;\nu)$ with ν degrees of freedom, Li and De Moor (1999) proposed a simple approximation of $F(x;\nu)$ as an alternative to various approximations listed in Johnson et. al. (1995).

It is well-known that the Student's t distribution has an ordinary normal approximation for $n\geq 30$, that is, $F(x;\nu)\approx\Phi(x)$.Li and De Moor (1999) proposed an adjusted normal approximation to the student t family distributions that

$$F_1(x,\nu)\approx\Phi(x\lambda)\text{ for }\nu\geq 3 \quad \text{with} \quad \lambda=\lambda(x,\nu)=\frac{(4\nu+x^2-1)}{(4\nu+2x^2)}$$

satisfies $(0<\lambda<1)$. (3)

for n=1 and 2, they have suggested the following exact formulae

$$F(x,1)=\frac{1}{2}+\frac{1}{\pi}\tan^{-1}x \text{ and } F(x,2)=\frac{1}{2}+\frac{x}{2}\frac{1}{(2+x^2)^{\frac{1}{2}}}$$

.

The maximum absolute error (Max. AE) for (3) is 0.006982 observed at x=3.8 for 3 degrees of freedom.

In particular, consider the approximation considered by Gleason (2000),

$$g(v)=\frac{v-1.5}{(v-1)^2} \tag{4}$$

and

$$F_2(x;v)\approx\Phi\left(z(x,v)\right)$$

where $z(x,v)=\sqrt{\frac{\log\left(1+\frac{x^2}{v}\right)}{g(v)}}$ (5)

for $v\geq 3$ and $x\geq 0$; the case $x<0$ is handled by symmetry.

The maximum absolute error for (5) is 0.004951 observed at x=1.0 for 3 degrees of freedom. Moreover, if accuracy at very small v is essential, the performance of (5) is easily improved by altering the definition of $g(v)$ using

$g^*(v)=\frac{v-1.5-\frac{0.1}{v}+\frac{0.5825}{v^2}}{(v-1)^2}$; in place of $g(v)$ in (5) reduces the maximum absolute error at $v=3$ to 0.002501 observed at x=0.9 (Gleason, 2000).

$$F_3(x;v)\approx\Phi\left(z(x,v)\right)$$

where $z(x,v)=\sqrt{\frac{\log\left(1+\frac{x^2}{v}\right)}{g^*(v)}}$ (6)

3. Approximation of student's t distribution developed by neural networks

In this section, a new formula for calculating the function $F_t(x;v)$ is obtained using neural networks. Since $F_t(x;v)$ is symmetric about zero, it is sufficient to approximate only for all values of $x\geq 0$ and $v\geq 3$. Feedforward neural networks are efficient nonparametric models for function approximation (Yerukala, 2012). Feedforward neural networks (FFNN) are known to be universal approximators of nonlinear functions; they are the most popular artificial neural networks structures for generating many input-output nonlinear mappings (Hornik et.al.; 1989). A feedforward neural network developed to approximate the CDF of student's t distribution with the values from $x=0$ with increment of 0.1 and are computed up to its corresponding probability equal to 0.9990 under each degrees of freedom and $v=3$ to 30 with increment of 1 as inputs and their corresponding $F_t(x;v)$ values computed using Micro-Soft Excel function TDIST( ) are used as the targets of the neural networks. The network structure contains an input layer with two neurons representing x values and degrees of freedom respectively , two hidden layers with four neurons in each and an output layer with one neuron representing the corresponding probability computed from TDIST() function. Back-propagation learning method is used to train and test the network. The following model obtained as an approximation of the cumulative distribution function of t-distribution.

$$F_4(x;v)\approx 0.259-1.435H(2,1)+0.604H(2,2)+0.548H(2,3)+0.75H(2,4). \tag{7}$$

where $H(2,1)=Tanh\left(-2.013+1.718H(1,1)-0.043H(1,2)+1.346H(1,3)+0.39H(1,4)\right)$,

$H(2,2)=Tanh\left(0.936-0.613H(1,1)+1.339H(1,2)-1.148H(1,3)-0.796H(1,4)\right)$

$H(2,3)=Tanh\left(-1.545+1.796H(1,1)-0.115H(1,2)+0.918H(1,3)+0.701H(1,4)\right)$

$H(2,4)=Tanh\left(-0.726+1.769H(1,1)-1.304H(1,2)+0.284H(1,3)+0.78H(1,4)\right)$

$H(1,1)=Tanh\left(0.015481v-0.268557x+0.22856\right)$

$H(1,2)=Tanh\left(0.219778v-0.212474x+0.776667\right)$

$H(1,3)=Tanh\left(-0.011259v-0.31433x+1.125778\right)$

$H(1,4)=Tanh\left(-0.061481v+0.011443x+0.841444\right)$

Under the restriction that $F_4(x,\nu)=1.0000$ if computed $F_4(x,\nu)\geq 0.9999$.

The maximum absolute error for the above function (7) is 0.000742 observed at x=0.10 for 13 and14 degrees of freedom.We propose one more function, which reduces the maximum absolute error of the proposed neural networks function (7).

$$F_5(x;v)\approx\begin{cases}F_1(x;v); & if \quad x\in[0.0\ \ 0.5)\\ F_4(x;v); & if \quad x\in[0.5\ \ 4.5)\\ F_2(x;v); & if \quad x\in[4.5\ \ 19.5]\end{cases} \tag{8}$$

The maximum absolute error for the above function (8) is 0.000582 observed at x=0.4 for 3 degrees of freedom. Approximation accuracy of the given functions measured using mean absolute error (MAE), mean absolute percentage error (MAPE) and root mean squared error (RMSE) and the results presented in the following table.

Table 3.1: Error measures for different approximations

Function	MAX. AE	MAE	RMSE	MAPE
$F_1(x;v)$	0.0070	0.0004	0.0323	0.0424
$F_2(x;v)$	0.0050	0.0001	0.0226	0.0157
$F_3(x;v)$	0.0025	0.0001	0.0161	0.0113
$F_4(x;v)$	0.0007	0.0003	0.0103	0.0390
$F_5(x;v)$	0.0006	0.0002	0.0003	0.0243

It is clearly observed that the proposed approximations having minimum of maximum absolute error and root mean squared errors as compared with the existing methods. The proposed two approximations are better than the function $F_1(x;v)$ with respect to the given error measures.

4. Conclusion

It is observed that the maximum absolute error of the functions $F_1(x;v)$, $F_2(x;v)$ and $F_3(x;v)$ is decreasing for higher degrees of freedom. The maximum absolute error for the function $F_4(x;v)$ is observed at 13 and 14 degrees of freedom and for remaining functions, it is observed at 3 degrees of freedom. Maximum absolute error for the function $F_3(x;v)$ is observed at $x=0.9$ for all ν. With respect to maximum absolute error of the functions, $F_3(x;v)$ is better than the $F_1(x;v)$ and $F_2(x;v)$ for all $\nu \le 5$; $F_1(x;v)$ is better than the $F_2(x;v)$ and $F_3(x;v)$ for all $\nu > 10$ or $x \le 0.5$; $F_2(x;v)$ is better than the $F_1(x;v)$, $F_3(x;v)$ and $F_4(x;v)$ if $x \ge 4.5$; and $F_5(x;v)$ is better than the $F_1(x;v)$, $F_2(x;v)$, $F_3(x;v)$ and $F_4(x;v)$ for all x and ν. The error behavior of the proposed functions is random. The two proposed functions performing well at lower degrees of freedom (up to $\nu = 10$). The proposed functions are well suited for the approximation to CDF of t-distribution up to the three decimal points as compared with the functions discussed in this paper.

References:

1. Brophy, A.L., "Efficient estimation of probabilities in the t distribution", Behavior Research Methods, Instruments & Computers, 19(5), pp. 462-466, 1987.
2. Gentleman, W. M. and M. A. Jenkins, "An Approximation for Student's t-Distribution, Biometrika", Vol. 55, No. 3, pp. 571-572, 1968.
3. Gleason, J.R. "A note on a proposed student t approximation", Computational Statistics & Data Analysis, 34, pp. 63-66, 2000.
4. Honik, K., Stinchcombe, M., White, H. Multilayer, "Feedforward Networks are Universal Approximators", Neural networks, Vol.2, pp. 359-366, 1989.
5. Johnson, N.L., Kotz, S. and Balakrishnan, N., "Distributions in Statistics: Continuous Univariate Distributions", Vol. 2, Second edition, New York. Wiley, 1995.
6. Li, B., De Moor, B., "A corrected normal approximation for student's t distribution", Computational Statistics & Data Analysis, 29, pp. 213-216, 1999.
7. Roy, P. Choudhury, A., "Approximate Evaluation of Cumulative Distribution Function of Central Sampling Distributions: A Review", Electronic Journal of Applied Statistical Analysis, Vol. 5, 121 – 131, 2012.
8. Yerukala, R., "Functional approximation using neural networks", Unpublished Ph.D. Thesis, Department of Statistics, Osmania University, Hyderabad, 2012.