Applied Mathematical Sciences, Vol. 2, 2008, no. 9, 425 - 429

A Note on Approximating the Normal Distribution Function

K. M. Aludaat[1] and M. T. Alodat[2]
Department of Statistics
Yarmouk University, Jordan
Aludaatkm@hotmail.com[1] and malodat@yui.edu.jo[2]

Abstract

In this paper, we propose a one-term-to-calculate approximation to the normal cumulative distribution function. Our approximation has a maximum absolute error of .00197323. We compare our approximation to the exact one.

Keywords: Cumulative distribution function, normal distribution

1. Introduction

Let X be a standard normal random variable, i.e., a random variable with the following probability density function

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \qquad -\infty < x < \infty.$$

The cumulative distribution function of the standard normal is given by

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt. \qquad (1)$$

The last integral has no closed form. Most basic statistical books give the values of this integral for different values of x in a table called the standard normal table. From the this table we can also find the value of x when $\Phi(x)$ is known.

Several authors gave approximations for by polynomials (Chokri, 2003; Johnson, 1994; Bailey, 1981; Polya, 1945).

These approximations give quite high accuracy, but computer programs are needed to obtain their values and they have a maximum absolute error of more than .003. But only the Polya's approximation

$\Phi(x) \approx 0.5(1+\sqrt{1-e^{-\frac{2}{\pi}x^2}})$

has one-term-to-calculate while the others need more than one term. They are reviewed in Johnson et al. (1994) as follows:

1. $\Phi_1(x) \approx 1-0.5(a_1+a_2x+a_3x^2+a_4x^3+a_5x^4+a_6x^5)^{-16}$,

where

$a_1 = 0.9999998582, \quad a_2 = 0.487385796, \quad a_3 = 0.02109811045,$

$a_4 = 0.003372948927, \quad a_5 = -0.00005172897742, \quad a_6 = 0.0000856957942.$

2. $\Phi_2(x) \approx 1-(2\pi)^{-1/2}\exp(-0.5x^2-0.94x^{-2}), \qquad x \geq 5.5$

3. $\Phi_3(x) \approx \dfrac{\exp(2y)}{1+\exp(2y)}, \qquad y = 0.7988x(1+0.04417x^2).$

4. $\Phi_4(x) \approx 1-0.5\exp\left(-\dfrac{(83x+351)x+562}{703/x+165}\right).$

All these approximations are good but they need computer program to be obtained. In addition their inverse can not be obtained easily. In this short note, we propose new approximations for $\Phi(x)$ and it's inverse. We can also find it's maximum absolute error.

2. The approximation

Since $\Phi(x)$ is symmetric about zero, it is sufficient to approximate

$$\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt$$

only for all values of $x>0$. According to Johnson et al. (1994) the Ploya's approximation represents an upper bound for $\Phi(x)$, i.e.,

$$\Phi(x) \leq 0.5\left(1+\sqrt{1-e^{-\frac{2}{\pi}x^2}}\right).$$

If we can find a sharper upper bound on $\Phi(x)$, then we can improve the Polya's approximation. Since $\sqrt{\frac{\pi}{8}} \leq \frac{2}{\pi}$ implies that $1-e^{-\sqrt{\frac{\pi}{8}}x^2} \leq 1-e^{-\frac{2}{\pi}x^2}$, then

$$\sqrt{1-e^{-\sqrt{\frac{\pi}{8}}x^2}} \leq \sqrt{1-e^{-\frac{2}{\pi}x^2}}.$$

So the term

$$0.5\left(1+\sqrt{1-e^{-\sqrt{\frac{\pi}{8}}x^2}}\right)$$

is closer to $\Phi(x)$ than

$$0.5\left(1+\sqrt{1-e^{-\frac{2}{\pi}x^2}}\right).$$

So we propose the following approximation for $\Phi(x)$:

$$\Phi_5(x) \approx 0.5+0.5\sqrt{1-e^{-\sqrt{\frac{\pi}{8}}x^2}}.$$

The inverse cumulative distribution function is approximated by

$$x=\sqrt{-\sqrt{\frac{8}{\pi}}\log(1-(p-0.5)^2)},$$

where $p=\Phi(x)$.

Our approximation has a maximum absolute error of 0.00197323. To prove this, we need to study the difference between the $\Phi(x)$ and $0.5+0.5\sqrt{1-e^{-\sqrt{\frac{\pi}{8}}x^2}}$. So we define for $x>0$, a function $g(x)$ as follows

$$g(x)=\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt-0.5\sqrt{1-e^{-\sqrt{\frac{\pi}{8}}x^2}}.$$

The first derivative of $g(x)$ is

$$g'(x)=\frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}-\frac{0.313329xe^{-\frac{1}{2}\sqrt{\frac{\pi}{2}}x^2}}{\sqrt{1-e^{-\frac{1}{2}\sqrt{\frac{\pi}{2}}x^2}}}.$$

We used the Mathematica Software to find the roots of the first derivative, i.e., the roots of $g'(x)=0$. We got the following three roots 0, 0.533811555441412 and 1.8783147540026042. Testing the derivative for the sign leads to the following facts: $g(x)$ is increasing in the interval [0, 0.533811555441412] $\cup$ [1.8783147540026042, ∞[and $g(x)$ is decreasing in [0.533811555441412, 1.8783147540026042]. So $g(x)$ has the two absolute extreme values -0.00197323 and 0.0010676. So the absolute maximum error is 0.00197323.

3. Comparison

In this section, we compare the exact value of $\Phi(x)$ with its approximated one. For positive values of x, Figure 1 shows the values of $\Phi(x)-.5$ and its approximation against $x>0$. We see from the Figure that $\Phi(x)-.5$ is very close to its approximation, which means that our approximation is very accurate.

Table 1. Comparison between $\Phi(x)$ and its approximations

x	$\Phi(x)$	$\Phi_1(x)$	$\Phi_2(x)$	$\Phi_3(x)$	$\Phi_4(x)$	$\Phi_5(x)$
0.6	**0 .7257**	0.7229	0.7437	0.7257	0.7259	**0.7247**
1.5	**0.9332**	0.9316	0.6427	0.9332	0.9331	**0.9347**
2.5	**0.9938**	0.9936	0.9621	0.9938	0.9939	**0.9950**

From Table 1 we see that our approximation is good. The best approximation is $\Phi_3(x)$, but computing its inverse requires computer programs, while the inverse of our approximation needs simple algebraic calculations. We note that the approximation $\Phi_2(x)$ is valid only for $x \geq 5.5$.

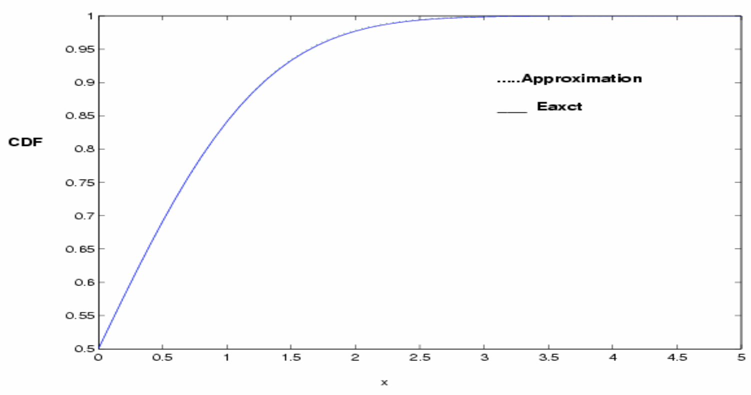

Figure 1 The exact value of $\Phi(x)$ -0.5 (dotted) and its approximation (smooth).

4. Conclusion

In this paper, we proposed an approximation to the cumulative distribution function of the standard normal distribution. Our approximation is one-term-to calculate and is better than the Polya's approximation. Numerical comparison shows that our approximation is very accurate. Moreover, it does not require computer programs to calculate both cumulative distribution function and its inverse.

Acknowledgements. This research has been supported by a grant from Yarmouk University".

References

[1] Bailey, B. J. R. (1981). Alternatives to hasting's approximation to the inverse of the normal cumulative distribution function. Applied statistics, 30, (3) 275-276.

[2] Chokri (2003). A short note on the numerical approximation of the standard normal cumulative distribution and its inverse. *Real 03-T-7, online manuscript*.

[3] Johnson, N. I., Kotz, S. and Balakrishnan, N. (1994) *Continuous univariate distributions*. John wiley & sons.

[4] LeBoeuf, C., Guegand, J., Roque, J. L. and Landry, P. (1987) *Probabilites* Ellipses.

[5] Polya, G. (1945) Remarks on computing the probability integral in one and two dimensions. *Proceeding of the first Berkeley symposium on mathematical statistics and probability*, 63-78.

[6] Renyi, A. (1970). *Probability theory*. North –Holland.

Received: August 11, 2007