

A Simple Approximation to the Area Under Standard Normal Curve

Amit Choudhury

Department of Statistics, Gauhati University, Guwahati 781014, India
 *Corresponding Author: achoudhury@rediffmail.com

Copyright © 2014 Horizon Research Publishing All rights reserved.

Abstract Of all statistical distributions, the standard normal is perhaps the most popular and widely used. Its use often involves computing the area under its probability curve. Unlike many other statistical distributions, there is no closed form theoretical expression for this area in case of the normal distribution. Consequently it has to be approximated. While there are a number of highly complex but accurate algorithms, some simple ones have also been proposed in literature. Even though the simple ones may not be very accurate, they are nevertheless useful as accuracy has to be gauged vis-à-vis simplicity. In this short paper, we present another simple approximation formula to the cumulative distribution function of standard normal distribution. This new formula is fairly good when judged vis-à-vis its simplicity.

Keywords Cumulative Distribution Function, Normal Distribution, Approximations

1. Introduction

In the collection of statistical distributions, the most well known and most frequently used across all domains is the standard normal distribution. In doing so, the area under the standard normal probability curve is calculated. This is often expressed as a function of the cumulative distribution function (cdf) of the distribution. Since a closed form expression of this cdf does not exist, it becomes necessary to approximate it. These days, many softwares like Ms-Excel and statistical softwares like SAS, SPSS provide for computation of this cdf using highly sophisticated algorithms.

Even with the advent of computers, there has been an interest in providing simple approximating this cdf. A review of the literature points to three approaches to the construction of an approximation. The most popular among them is the construction of approximation formulas. The second approach uses distributions that closely resemble the normal distribution under specific conditions. The third approach involves construction of bounds. In all these approaches, the

approximations or bounds have been for the standard normal variety.

Various approximations and bounds are placed below.

- i. Tocher (1963) : $\Phi(x) \approx e^{2kx}/(1 + e^{2kx})$ where $k = \sqrt{2/\pi}$
- ii. Zelen and Severo (1964):
 $\Phi(x) \approx 1 - (0.4361836 t - 0.1201676 t^2 + 0.9372980 t^3)$
 $(\sqrt{2/\pi})^{-1} e^{-x^2/2}$,
 where $t = (1 + 0.33267x) - 1$
- iii. A popular bound found in many probability texts (for example, Feller 1968). For $x > 0$
 $(x^{-1} - x^{-3}) (\sqrt{2/\pi})^{-1} e^{-x^2/2} < 1 - \Phi(x) < x^{-1}$
 $(\sqrt{2/\pi})^{-1} e^{-x^2/2}$
- iv. Page (1977): $\Phi(x) \approx 0.5\{1 + \tanh(y)\}$ where $y = \sqrt{2/\pi} x(1 + 0.044715x^2)$
- v. Hammakar (1978): $1 - \Phi(x) \approx 0.5 \left\{ 1 - \left(1 - e^{-y^2} \right)^{0.5} \right\}$, $y = 0.806x(1 - 0.018x)$
- vi. Abernathy (1988): $\Phi(x) \approx 0.5 + \frac{1}{\sqrt{2\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{2^n n! (2n+1)}$, $x > 0$
- vii. Lin(1989): $1 - \Phi(x) \approx 0.5 \left(e^{-0.717x - 0.416x^2} \right)$, $x > 0$
- viii. Lin(1990): $1 - \Phi(x) \approx 1/(1 + e^y)$ where $y = 4.2 \pi \{x/(9 - x)\}$, $x > 0$
- ix. (ix) Bagby(1995):
 $\phi(x) \approx 0.5 + 0.5 \left[1 - \left(\frac{1}{30} \right) \left\{ 7e^{-x^2/2} + 16e^{-x^2(2-\sqrt{2})} + \left(7 + \frac{\pi}{4} x^2 \right) e^{-x^2} \right\} \right]^{0.5}$, $x > 0$
- x. (x) Szarek bounds (1999): For $x > -1$
 $\frac{0.5 e^{-x^2/2}}{x + (x^2 + 4)^{0.5}} \leq \frac{\int_x^{\infty} e^{-t^2/2} dt}{x} \leq \frac{0.25 e^{-x^2/2}}{3x + (x^2 + 8)^{0.5}}$
- xi. Byrc (2002A):

$$\phi(x) \approx 1 - \frac{(4 - \pi)x + \sqrt{2\pi}(\pi - 2)}{(4 - \pi)x^2 \sqrt{2\pi} + 2\pi x + 2\sqrt{2\pi}(\pi - 2)} e^{-x^2/2}$$

xii. Byrc (2002B): $\Phi(x) \approx 1 -$

$$\frac{x^2 + 5.575192695x + 12.77436324}{x^3 \sqrt{2\pi} + 14.38718147x^2 + 31.53531977x + 25.548726} e^{-x^2/2}$$

xiii. Standard Logistic cdf (Johnson and Kotz, 1994) :

$$\Phi(x) \approx F(x) = \{1 + \exp(-\pi x/\sqrt{3})\}^{-1}$$

xiv. Shore(2005): $\phi(x) \approx \frac{1 + g(-x) - g(x)}{2}$

where

$$g(x) = \exp\left\{-\log(2)\exp\left\{\left[\alpha/(\lambda/S_1)\right]\left((1 + S_1x)^{(\lambda/S_1)} - 1\right) + S_2x\right\}\right\}$$

$$\lambda = -0.61228883, S_1 = -0.11105481,$$

$$S_2 = 0.44334159, \alpha = -6.37309208$$

xv. Aludaat and Alodat (2008): $\phi(x) \approx 0.5 + 0.5 \sqrt{1 - e^{-\sqrt{\frac{\pi}{8}}x^2}}$

xvi. Winitzki (2008):

$$\phi(x) \approx \frac{1 + \left[1 - \exp\left\{-\left(x^2/2\right)\left(\frac{4}{\pi} + 0.147x^2/2\right)\right\} / \left(1 + 0.147x^2/2\right)\right]^2}{2}$$

It is therefore clear that the topic has been of interest for more than half a decade.

2. A Simple Approximation Formula

We first note that it is enough to present an approximation formula for $x > 0$. We concentrate on the tail area probability and suggest the following approximation formula for it.

$$\int_x^\infty e^{-u^2/2} du = \frac{e^{-x^2/2}}{0.226 + 0.64x + 0.33\sqrt{x^2 + 3}}, \quad x > 0$$

Then a simple approximation to the standard normal cdf is

$$\Phi(x) \approx 1 - \frac{1}{\sqrt{2\pi}} \frac{e^{-x^2/2}}{0.226 + 0.64x + 0.33\sqrt{x^2 + 3}}, \quad x > 0$$

3. Conclusion

In this paper we have proposed a very simple approximation to the cdf of standard normal distribution. We endeavored to present a simplistic approximation formula with reasonable accuracy. Our formula is so simple that it can be implemented in any hand held calculator. Accuracy of any approximation has to be seen vis a vis the simplicity of the approximation. The maximum absolute error of our approximation is 0.0001922. In other words, even though our approximation is very simple, it provides a minimum accuracy of three digits. We calculated mean absolute error of our formula for $x=0.00005$ (0.00005) 3.0 and it was found to be 5.54339E-05. An algorithm for constructing a function for cdf of standard normal distribution can now be designed. We place in the appendix a pseudocode for constructing such

a function.

4. Appendix

Subroutine Std_Normal_CDF(x : float as input, z : float as output)

variables: pi,x,y,z : float

indi : integer

pi = 3.14159265358979

begin case

case: $x < 0$

indi = 1,

$x = (-1)*x$

case: $x \geq 0$

indi = 0

end case

$y = 1 - (1/(2*\sqrt{2*\pi})) * \exp(-\sqrt{x}/2) / (0.226 + 0.64*x + 0.33 * \sqrt{x*x+3})$

begin case

case: indi=1

$z = 1 - y$

case: indi = 0

$z = y$

end case

return (z)

End of Subroutine Std_Normal_CDF

REFERENCES

- [1] Abernathy, R.W. (Nov 1988). Finding Normal Probabilities with Hand held Calculators. Mathematics Teacher,81, 651-652.
- [2] Aludaat,K.M. and Alodat,M.T.(2008).A Note on Approximating Normal Distribution Function. Applied Mathematical Statistics,Vol.2, 2008, no.9, 425-429.
- [3] Bagby, R.J. (1995).Calculating Normal Probabilities. American Math Monthly, 102, 46-49.
- [4] Byrc, W. (2002A and 2002B). A uniform approximation to the right normal tail integral. Applied Mathematics and Computation, 127, 365-374.
- [5] Feller. W. (1968). An Introduction to Probability Theory and its Applications, Vol. 1, Wiley Eastern Limited.
- [6] Hammakar, H. C. (1978). Approximating the Cumulative Normal Distribution and its Inverse. Applied Statistics, 27, 76-77.
- [7] Johnson, N.L. and Kotz, S. (1994). Continuous Univariate Distribution Vol. 1. John Wiley and Sons.
- [8] Page, E. (1977). Approximations to the cumulative normal function and its inverse for use on a pocket calculator. Applied Statistics, 26, 75-76.
- [9] Lin, J.T. (1989). Approximating the Normal Tail Probability and its Inverse for use on a Pocket Calculator. Applied

Statistics, 38, 69-70.

- [10] Lin, J.T. (1990). A Simpler Logistic Approximation to the Normal Tail Probability and its Inverse. *Applied Statistics*, 39, 255-257.
- [11] Shenton, L.R. (1954). Inequalities for the normal integral. *Biometrika*, 41, 177-189.
- [12] Shore, H. (2005). Accurate RMM-based Approximations for the CDF of the Normal Distribution. *Communications in Statistics-Theory and Methods*, 34, 507-513.
- [13] Tocher, K. D. (1963). *The Art of Simulation*. English University Press, London.
- [14] Wall, H. (1948). *Analytic Theory of Continued Fractions*, Van Nostrand, Princeton, N.J. pp 358
- [15] Winitzki, S. (Feb 2008) A handy approximation for the error function and its inverse. Downloaded from <http://homepages.physik.uni-muenchen.de/~Winitzki/erf-approx.pdf> on Nov 18, 2008.
- [16] Zelen, M. and Severo, N.C. (1964). Probability Function (pp 925-995) in *Handbook of Mathematical Functions*, Edited by M. Abramowitz and I. A. Stegun, Applied Mathematics Series, 55.