

## Some improvements for existing simple Approximations of the Normal Distribution Function

Ahmad A. Hanandeh<sup>1\*</sup>, Omar M. Eidous<sup>2</sup>

\* Corresponding Author



1. Department of Statistics, Science Faculty, Yarmouk University, Jordan, ahmad.hanandeh@yu.edu.jo
2. Department of Statistics, Science Faculty, Yarmouk University, Jordan, omarm@yu.edu.jo

### Abstract

Due to the widespread applicability and use of the normal distribution, a need has arisen to approximate its cumulative distribution function (cdf). In this article, five new simple approximations to the standard normal cdf are developed. In order to assess the accuracy of the proposed approximations, both maximum absolute error and mean absolute error were used. The maximum absolute errors of the proposed approximations lie between 0.00095 and 0.00946, which is highly accurate if compared to the existing simple approximations and quite sufficient for many real-life applications. Even though simple approximations may not as accurate as complicated ones, they are, though, fairly good when judged vis-à-vis their simplicity.

**Key Words:** Approximations; Cumulative Distribution Function; Maximum Absolute Error; Normal Distribution.

**Mathematical Subject Classification:** Primary 62E17; Secondary 62J02, 62J12.

### 1. Introduction

One of the most familiar distributions that are ubiquitous in statistics is the normal (Gaussian) distribution since it fits many natural phenomena very well. This distribution describes a family of continuous probability distributions and plays a fundamental role in several applied problems of biology, business, communication theory, engineering, financial risk management, genetics, hydrology, mechanics, medicine, physics, psychology, reliability, signal processing, and many others. The importance and widely spread usage of this distribution come from the fact that it is an infinitely divisible distribution with finite variance, has a bell-shaped density curve, and is the result of the central limit theorem.

The use of the normal distribution often involves computing its cumulative distribution function (cdf) which, unfortunately, does not have a closed-form theoretical expression. To overcome this handicap, many authors introduced approximations to this cdf. However, most of these approximations are complicated which has inspired the research presented in this paper where the aim is to propose several approximations of the normal cdf with simple form and sufficient accuracy.

Now, suppose that a random variable  $X$  follows a probability density function (pdf) of the standard normal distribution, then its pdf and cdf are, respectively, given as follows:

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R},$$

and

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

The integral in (1) is one of the most widely used integrals in the field of applied mathematics and statistics; however, it does not have a closed-form formula which makes statistical analysis limited.

The main goal of this paper is to propose some new approximations to the standard normal cdf which are simple and convenient to enable a straightforward analysis of the normal distribution. The approximations proposed in this paper are useful due to their simple form especially when statistical tables and/or softwares are not readily available. In addition, the level of accuracy of these approximations is superior if compared to the existing simple approximations reported in the literature. The simplicity of the new approximations stands in sharp contrast to existing, mostly complex approximations, which tend to have a larger number of terms which makes them not tractable.

Eventually, one should highlight that approximating the normal cdf may also be helpful in many applied analyses involving functions that are directly related to it such as gamma function and incomplete gamma function. The remainder of this article is organized as follows. Section 2 provides a summary about the criteria that we used to compare the approximations. Section 3 presents a literature review of some previous work related to existing approximations of the normal cdf. Section 4 proposes the five approximations and demonstrates their accuracy in comparison with the other existing approximations from the literature that have similar analytical form complexity. Finally, Section 5 summarizes the key features of the proposed approximations and explores some conclusions and implications.

## 2. Maximum and mean absolute error measures

Let  $\hat{\Phi}(z)$  denotes an approximation of  $\Phi(z)$  at a specific value of  $z$ . To explore the accuracy of  $\hat{\Phi}(z)$ , we can use the absolute error (AE) measure. The AE can be defined as follows:

$$AE(z) = |\hat{\Phi}(z) - \Phi(z)|$$

For comparison purpose, we use the maximum absolute error (Max.AE) of  $\hat{\Phi}(z)$  which is equal to

$$\max_z |\hat{\Phi}(z) - \Phi(z)|$$

The mean absolute error (MAE) can also be used as a measure of accuracy, it can be computed by

$$MAE(z_0, z_1, \dots, z_n) = \frac{\sum_i AE(z_i)}{n}, z_0 < z_1 < \dots < z_n$$

To compute the  $MAE(z)$  for any value of  $z$ , if we take the values of  $z$  to be 0 to 3 with step 0.0001, which indicates that the above sum is evaluated at 30001 points, 0, 0.0001, 0.0002, ..., 2.9999, 3.

## 3. Literature review on some approximations of the normal cdf

There is notable number of approximations to the cdf of normal distribution available in the literature. Most recent approximations include Malki (2017), Maticet al. (2018), Alkhazali et al. (2019), Shchigolev (2020) and Hanandeh and Eidous (2021). To provide a fair comparison, this section presents a brief overview of the previous work related some existing simple approximations to the normal cdf. We give here the five approximations that we have updated so that the reader may assess the new ones (developed in this article) relative to original approximations. The result of this review is summarized in Table 1. For further work and deep review on some existing approximations, we refer the reader to Eidous and Abu-Shareefa (2020) and the references therein.

To measure the accuracy and performance of our proposed approximations, MAE and Max.AE are computed and used for further comparison with some existing approximations. Note that we consider the case of  $z \geq 0$ , we can easily get it for the case  $z < 0$  using the symmetry property of the normal distribution:  $\Phi(z) = 1 - \Phi(-z)$ . The results for MAE and Max.AE appears in this paper has been done for the values of  $z$  between 0 to 7 with step 0.00001.

**Table 1:** Summary of some existing approximations to the normal cdf.

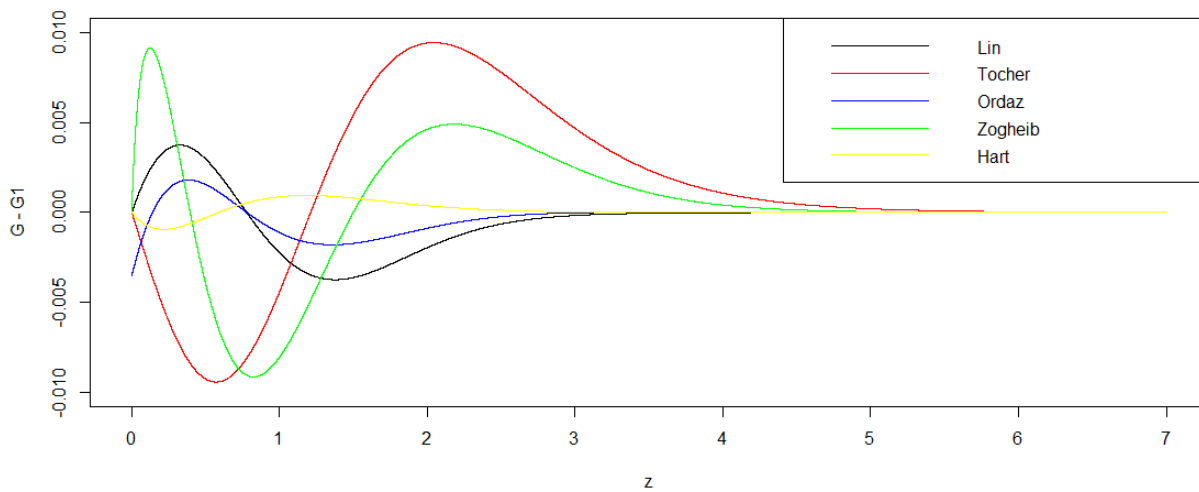
Author(s) name	Approximation formula	Max.AE	MAE
Lin (1990)	$\hat{\Phi}_1(z) = \frac{1}{1 + e^{-4.2\pi z/(9-z)}}$	0.00668	0.00078
Tocher (1963)	$\hat{\Phi}_2(z) = \frac{e^{2\sqrt{2/\pi}z}}{1 + e^{2\sqrt{2/\pi}z}}$	0.01767	0.00506
Ordaz (1991)	$\hat{\Phi}_3(z) = 1 - 0.6931 e^{-\left(\frac{-9z-8}{14}\right)^2}$	0.00440	0.00055
Zogheib and Hlynka (2009)	$\hat{\Phi}_4(z) = 1 - 0.5 e^{-1.2 z^{1.3}}$	0.01120	0.00203
Hart (1957)	$\hat{\Phi}_5(z) = 1 - \frac{\phi(z)}{z + \sqrt{2/\pi}e^{-0.4z}}$	0.00363	0.00040

**4. The New proposed approximations**

In this section, we present five new (updated) approximation to the cdf of standard normal distribution with the following forms:

**Table 2:** Summary of new approximations to the normal cdf.

New updated approximation name	Approximation formula	Max.AE	MAE
Updated Lin (1990)	$\hat{\Phi}_1^*(z) = \frac{1}{1 + e^{-13.48266z/(9-z)}}$	0.00375	0.00087
Updated Tocher (1963)	$\hat{\Phi}_2^*(z) = \frac{e^{1.7017z}}{1 + e^{1.7017z}}$	0.00946	0.00326
Updated Ordaz (1991)	$\hat{\Phi}_3^*(z) = 1 - 0.688182e^{-\left(\frac{-9z-8}{14}\right)^2}$	0.00353	0.00041
Updated Zogheib and Hlynka (2009)	$\hat{\Phi}_4^*(z) = 1 - 0.5 e^{-1.2 z^{1.275247}}$	0.00917	0.00208
Updated Hart (1957)	$\hat{\Phi}_5^*(z) = 1 - \frac{\phi(z)}{z + \sqrt{2/\pi}e^{-0.43432z}}$	0.00095	0.00021



**Figure 1:** The difference between  $\hat{\Phi}^*(z)$  and  $\Phi(z)$  for the proposed approximations.

Our code has been done on R-software and is available upon request.

It is easy to show that  $\hat{\Phi}_1^*(z)$ ,  $\hat{\Phi}_2^*(z)$ ,  $\hat{\Phi}_3^*(z)$  and  $\hat{\Phi}_4^*(z)$  are invertible as shown in the following table:

**Table 3:** Invertible formula for some of the new approximations to the normal cdf.

New updated approximation	Invertible formula
Updated Lin (1990)	$z = \frac{9 \ln\left(\frac{1}{\hat{\Phi}_1^*(z)} - 1\right)}{-13.4826 + \ln\left(\frac{1}{\hat{\Phi}_1^*(z)} - 1\right)}$
Updated Tocher (1963)	$z = \frac{\ln\left(\frac{\hat{\Phi}_2^*(z)}{1 - \hat{\Phi}_2^*(z)}\right)}{1.7017}$
Updated Ordaz (1991)	$z = \frac{14 \sqrt{\ln\left(\frac{0.688182}{1 - \hat{\Phi}_3^*(z)}\right)} - 8}{9}$
Updated Zogheib and Hlynka (2009)	$z = \left(-\frac{\ln(-2\hat{\Phi}_4^*(z) + 2)}{1.2}\right)^{\frac{1}{1.275247}}$

### 5. Concluding Remarks

In this article, some simple approximations to the standard normal cdf are presented. The accuracy of these approximations is evaluated based on the error computed from the respective cdf approximation available in the R software.

It is remarkable to mention that, although other more accurate approximations are existing in the literature, those approximations are very complicated. The features of our approximations are the simplicity of their analytical form which makes them easily programmable as well as their relatively high accuracy. They also don't require different forms over different intervals, as a result, these three features make them widely applicable and comparable to or even better than previously reported approximations in the literature of similar analytical form complexity.

The Max.AEs of the proposed approximations lie between 0.00095 and 0.00946, while the MAEs lie between 0.00021 and 0.00326 which makes these approximations fairly good compared to the existing simple approximations and quite sufficient for many real-life applications. Although we have introduced five simple approximations, the final formula  $\hat{\Phi}_5^*$  is the most accurate with Max.AE equal to 0.00095 and MAE equal to 0.00021 and thus we recommend using it.

### References

1. Alkhazali A., Al-Rabayah M. and Hamasha M. M. (2019). A new high accuracy mathematical approximation to the cumulative normal density function. *Jordan Journal of Mechanical and Industrial Engineering*, 13(4), 247-251.
2. Eidous, O. M. and Abu-Shareefa, R., (2020). New approximations for standard normal distribution function. *Communications in Statistics-Theory and Methods*, 49(6), 1357-1374.
3. Hart, R. G. (1957). A formula for the approximation of the normal distribution function. *Mathematical Tables and Other Aids to Computation*, 11(60), 265-265.
4. Hanandeh, A. A., and Eidous, O. M. (2021). A new one-term approximation to the standard normal distribution. *Pakistan Journal of Statistics and Operation Research*, 17(2), 381-385.
5. Lin, J. T. (1990). A simpler logistic approximation to the normal tail probability and its inverse. *Applied Statistics*, 39(2), 255-257.
6. Malki, A. (2017). Improvement Bryc's approximation to the cumulative distribution function. *International Journal of Swarm Intelligence and Evolutionary Computation*, 17(6), 1-7.
7. Matic I. Radoičić R. and Stefanica D. (2018). A sharp Pólya - based approximation to the normal cumulative distribution function. *Applied Mathematics and Computation*, 322, 111-122.
8. Ordaz, M. (1991). A simple approximation to the Gaussian distribution. *Structural Safety*, 9(4), 315-318.
9. Shchigolev V. K. (2020). A Simple Approximation for the normal distribution function via Variational iteration method. *MathLAP Journal*, 6, 45-52

10. Tocher, K. D. (1963). The art of simulation. The English Universities Press: London, UK.
11. Zogheib, B. and Hlynka, M. (2009). Approximations of the standard normal distribution. University of Windsor, Dept. of Mathematics and Statistics, on-line text.
12. Afify, A. Z., Yousof, H. M., Cordeiro, G. M., Ortega, E. M. M. & Nofal, Z. M. (2016b). The Weibull Fréchet distribution and its applications. Journal of Applied Statistics, 43(14), 2608-2626.