Pak.j.stat.oper.res. Vol.17 No. 2 2021 pp 381-385 DOI: http://dx.doi.org/10.18187/pjsor.v17i2.3556

Pakistan Journal of Statistics and Operation Research

A New One-term Approximation to the Standard Normal Distribution

Ahmad A. Hanandeh[1*] and Omar M. Eidous[2]

* Corresponding Author

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

Abstract

This paper deals with a new simple one-term approximation to the cumulative distribution function (cdf) of the standard normal distribution which does not have a closed-form representation. The accuracy of the proposed approximation has been evaluated by using the maximum absolute error (MAE), which was about 0.0016 and this accuracy is sufficient for most practical applications. To compare the proposed approximation with some of the existing one-term approximations, the MAE measure was considered. In addition, the difference between the exact and approximation of the normal cdf was computed for specific values of the normal random variable.

Key Words: Cumulative Distribution Function (cdf), Normal Distribution, Maximum Absolute Error (MAE), Approximation

Mathematical Subject Classification: 62E17

1. Introduction

Perhaps the most important and most widely used continuous probability distribution in statistics is the normal distribution. The main reasons are due to its importance in statistical inference and the existence of the central limit theorem. In addition, it plays a significant role in many applications in scientific fields, like engineering, genetics, psychology, biology, medicine, financial risk management, hydrology, mechanics, physics, reliability, and others.

The use of normal distribution often involves computing the area under its probability curve. Suppose that a random variable Z follows a probability density function (pdf) of the standard normal distribution, then its pdf and cdf are, respectively, defined as:

$$g(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \qquad -\infty < z < \infty,$$

and

$$G(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{y^2}{2}} dy \tag{1}$$

As it is well-known, a closed-form solution for the integral in equation (1) does not exist. Consequently, it has to be approximated.

While there are several complex but accurate algorithms, some simple ones have also been proposed in the literature (See for instance, Eidous and Al-Salman, 2015; Aludaat and Alodat, 2008; Johnson et al. 1994; Bailey, 1981 and

Polya, 1949). Even though the simple ones may not be very accurate, they are nevertheless useful as accuracy has to be gauged vis-à-vis simplicity.

Now, let

$$G^*(z) = \frac{1}{\sqrt{2\pi}} \int_0^z e^{-\frac{y^2}{2}} dy,$$

it follows that

$$G(z) = 0.5 + \frac{1}{\sqrt{2\pi}} \int_0^z e^{-\frac{y^2}{2}} dy = 0.5 + G^*(z), \qquad z \geq 0$$

Because of the symmetry of the standard normal distribution, we get $G(z) = 1 - G(-z)$, which implies that it is sufficient to approximate $G^*(z)$.

This paper presents a simple approximation with one-term for $G(z)$ when $z \geq 0$. The idea of approximating $G(z)$ presented here is similar to that used in Polya (1949); Aludaat and Alodat (2008) and Eidous and Al-Salman (2016). On the other hand, the MAE of the presented approximation is less than the MAE of the other one-term approximation approaches in the previously mentioned works.

The remaining part of this paper is presented as: In Section 2, a brief literature review on one-term approximation functions is given. A discussion on our proposed one-term approximation is provided in Section 3. Section 4 illustrates the performance comparisons between our proposed approximation with many other existing one-term approximations. Finally, section 5 provides a short conclusion on our work.

2. Literature review

There are many existing approximation approaches of normal density introduced in the literature. However, in our work, we are interested in the approximations that are of one-term form. This section presents a historical review of different one-term approximations for the normal cdf.

The following table gives some typical one-term approximation formulas of $G(z)$ when $z \geq 0$ and the corresponding MAE of each formula.

Table 1: Some approximations of $G(z)$ and their corresponding MAE.

Author(s) name	Approximation formula	MAE
Polya (1949)	$G_1(z) = 0.5\left(1 + \sqrt{1 - e^{-\frac{2}{\pi}z^2}}\right)$	0.0031
Aludaat and Alodat (2008)	$G_2(z) = 0.5\left(1 + \sqrt{1 - e^{-\sqrt{\frac{\pi}{8}}z^2}}\right)$	0.0020
Eidous and Al-Salman (2015)	$G_3(z) = 0.5\left(1 + \sqrt{1 - e^{-\frac{5}{8}z^2}}\right)$	0.0018
Bowling et al. (2009)	$G_4(z) = \frac{1}{1 + e^{-1.702z}}$	0.0095
Lin (1990)	$G_5(z) = \frac{1}{1 + e^{-4.2\pi z/(9-z)}}$	0.0067
Tocher (1963)	$G_6(z) = \frac{e^{2\sqrt{2/\pi}z}}{1+e^{2\sqrt{2/\pi}z}}$	0.0177
Ordaz (1991)	$G_7(z) = 1 - 0.6931\, e^{-\left(\frac{-9z-8}{14}\right)^2}$	0.0044
Zogheib and Hlynka (2009)	$G_8(z) = 1 - 0.5\, e^{-1.2\, z^{1.3}}$	0.0112

The above approximation approaches are simple due to using one-term to evaluate $G(z)$. Several other approximation approaches for $G(z)$ that used more than one term have been reported in the literature (see for instance Eidous and Abu-Shareefa, 2020).

3. The proposed method

The main purpose for this paper is to present a new formula to approximate the cdf of standard normal distribution $G(z)$, which is defined as following:

$$G(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{y^2}{2}} dy \quad , \quad z \geq 0.$$

Due to Johnson et al. (1994), the Ploya's approximation represents an upper bound for $G(z)$, i.e.,

$$G(z) \leq 0.5\left(1 + \sqrt{1 - e^{-\frac{2}{\pi}z^2}}\right)$$

Aludaat and Alodat (2008) updated the upper bound of $G(z)$ by showing that

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

On the other hand, if the above inequality is true; it does not necessarily indicate that

$$G(z) \leq 0.5\left(1 + \sqrt{1 - e^{-\sqrt{\frac{\pi}{8}}z^2}}\right).$$

So, the conjectured by Aludaat and Alodat (2008) that their upper bound approximation is sharper than Polya's one is not accurate. However, it is still true that their approximation has a lower MAE compared to Polya's one. Now to approximate $G(z)$, we know that

$$\text{MAE}\left\{0.5\left(1 + \sqrt{1 - e^{-\sqrt{\frac{\pi}{8}}z^2}}\right)\right\} < \text{MAE}\left\{0.5\left(1 + \sqrt{1 - e^{-\frac{2}{\pi}z^2}}\right)\right\}$$

Next, let

$$G_9(z) = 0.5\left(1 + \sqrt{1 - e^{-\frac{81}{130}z^2}}\right)$$

be the proposed approximation for $G(z)$.
Now, since

$$\frac{81}{130} < \frac{5}{8} < \sqrt{\frac{\pi}{8}} < \frac{2}{\pi}$$

Therefore,

$$\sqrt{1 - e^{-\frac{81}{130}z^2}} < \sqrt{1 - e^{-\frac{5}{8}z^2}} < \sqrt{1 - e^{-\sqrt{\frac{\pi}{8}}z^2}} < \sqrt{1 - e^{-\frac{2}{\pi}z^2}}.$$

Consequently,

$$G_9(z) < G_3(z) < G_2(z) < G_1(z)$$

where $G_1(z)$, $G_2(z)$ and $G_3(z)$ are given in Table 1.
Suppose that $k = G_9(z)$, it follows that approximation of the inverse cdf is given by:

$$z = \sqrt{-\frac{130}{81} log(1 - [2(k - 0.5)]^2)}.$$

Now, to find the MAE of $G_9(z)$ as an approximation of $G(z)$, define $u(z)$ as

$$u(z) = G(z) - G_9(z) = \frac{1}{\sqrt{2\pi}} \int_{0}^{z} e^{-\frac{y^2}{2}} dy - 0.5\sqrt{1 - e^{-\frac{81}{130}z^2}} \quad , \quad z \geq 0.$$

Hence, by differentiating $u(z)$ with respect to z we get $u'(z) = \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} - \frac{81ze^{-\frac{81}{130}z^2}}{260\sqrt{1-e^{-\frac{81}{130}z^2}}}$.

Next, setting $u'(z)$equal to zero and solving with respect to z, we get two solutions

$$z = 0.600925 \text{ and } z = 1.96739$$

Plugging these values in $u(z)$, we find that $u(z)$ is equal to 0.00162038 and -0.00162702 at $z = 0.600925$ and $z = 1.96739$ respectively. Hence, the MAE of $G_9(z)$ is 0.00162702. By comparing this MAE with the one's in the others approaches, we can see that this MAE is the smallest comparing with the MAE of $G_1(z)$, $G_2(z)$, ..., $G_8(z)$ that are stated in Table (1), i.e., $G_9(z)$ gives higher accuracy approximation.

4. Comparison

In this section, we have carried out a comparison between our approximation $G_9(z)$ and the other one-term approximation approaches $G_1(z)$, $G_2(z)$, ..., $G_8(z)$ that are presented in Table (1). This comparison has been done for the values of z between 0.2 to 4.2 with step 0.4. The results of the different approximation approaches are given in Table (2) below. The corresponding error between the exact $G(z)$ and each of the different approximations is also provided to compare the accuracy of the approaches easily.

From Table (2), we can see that our approximation $G_9(z)$ is the best one for $1.4 \leq z \leq 3$. In this case, the corresponding MAE of $G_9(z)$ is less than that of the other approximation approaches.

Table (2): The exact value of $G(z)$ and the corresponding values of its one-term approximation using the different approaches $G_1(z)$, $G_2(z)$, ..., $G_8(z)$and $G_9(z)$. The differences between $G(z)$ and the other approximation approaches are given in the parentheses.

Z	$G(z)$	$G_1(z)$	$G_2(z)$	$G_3(z)$	$G_4(z)$	$G_5(z)$	$G_6(z)$	$G_7(z)$	$G_8(z)$	$G_9(z)$
0.2	0.579260	0.579283 (-0.000024)	0.578668 (0.000592)	0.578565 (0.000694)	0.584288 (-0.005028)	0.570918 (0.008341)	0.579118 (0.000142)	0.575389 (0.003871)	0.569885 (0.009374)	0.578446 (0.000813)
0.6	0.725747	0.726284 (-0.000537)	0.72470 (0.001047)	0.724435 (0.001312)	0.735206 (-0.009459)	0.710456 (0.015291)	0.722613 (0.003134)	0.722713 (0.003034)	0.73317 (-0.007420)	0.724127 (0.001620)
1.0	0.841345	0.843119 (-0.00177)	0.841184 (0.000161)	0.840859 (0.000486)	0.845796 (-0.004451)	0.827897 (0.013448)	0.831426 (0.009919)	0.841353 (0.000008)	0.852385 (-0.011040)	0.840481 (0.000864)
1.4	0.919243	0.922154 (-0.002900)	0.920474 (-0.001231)	0.920191 (-0.000950)	0.915506 (0.003737)	0.91010 (0.009140)	0.903268 (0.015975)	0.920475 (-0.001230)	0.92442 (-0.005180)	0.919861 (-0.000618)
1.8	0.964070	0.967141 (-0.003071)	0.966024 (-0.001954)	0.965834 (-0.001765)	0.955366 (0.008704)	0.958576 (0.005494)	0.946464 (0.017605)	0.965075 (-0.001000)	0.963579 (0.000490)	0.965613 (-0.001543)
2.2	0.986097	0.988390 (-0.002293)	0.987809 (-0.001712)	0.987709 (-0.001613)	0.976897 (0.009200)	0.983136 (0.002960)	0.97099 (0.015106)	0.986562 (-0.00047)	0.983317 (0.002779)	0.987593 (-0.001497)
2.6	0.995339	0.996608 (-0.001269)	0.996371 (-0.001032)	0.99633 (-0.000991)	0.988170 (0.007169)	0.993971 (0.001368)	0.984465 (0.010874)	0.99547 (-0.00013)	0.992687 (0.002652)	0.996282 (-0.000943)
3.0	0.998650	0.999187 (-0.000537)	0.999111 (-0.000461)	0.999098 (-0.000447)	0.993976 (0.004674)	0.998136 (0.000514)	0.991734 (0.006916)	0.998662 (0.000012)	0.996917 (0.001733)	0.999082 (-0.000432)
3.4	0.999663	0.999841 (-0.000178)	0.999821 (-0.000158)	0.999818 (-0.000160)	0.996942 (0.002722)	0.999514 (0.000150)	0.995617) (0.004046	0.999654 (0.000009)	0.998746 (0.000917)	0.999814 (-0.000151)
3.8	0.999928	0.999975 (-0.000046)	0.999971 (-0.000043)	0.999970 (-0.000040)	0.998449 (0.001478)	0.999894 (0.000030)	0.99768 (0.002248)	0.999922 (0.000006)	0.999506 (0.000422)	0.999969 (-0.000041)
4.2	0.999987	0.999997 (-0.000010)	0.999996 (-0.000009)	0.999996 (-0.000009)	0.999214 (0.000772)	0.999983 (0.000003)	0.998773 (0.001213)	0.999984 (0.000002)	0.999811 (0.000176)	0.999996 (-0.000009)

Our code is available upon request

 DOI: http://dx.doi.org/10.18187/pjsor.v17i2.3556

5. Conclusion

In this paper, we have proposed a new one-term approximation for the cumulative distribution function of standard normal distribution. We have compared our approximation with many other existing one-term approximations, the MAE of the proposed approximation is less than that of all the existing approximations stated in this paper.

Acknowledgment

The authors are thankful to the anonymous reviewers for constructive comments and criticism that improved the paper.

References

1. Aludaat, K. M. and Alodat, M. T. (2008). A note on approximating the normal distribution function. Applied Mathematical Science, 2(9), 425-429.
2. Bailey, B. J. (1981). Alternatives to Hasting's approximation to the inverse of the normal cumulative distribution function. Applied Statistics, 30(3), 275-276.
3. Bowling, S. R., Khasawneh, M. T., Kaewkuekool, S. and Cho, B. R. (2009). A logistic approximation to the cumulative normal distribution. Journal of Industrial Engineering and Management, 2(1), 114-127.
4. Eidous, O. M. and Al-Salman, S. A. (2015). One-term approximation for normal distribution function. Mathematics and Statistics, 4(1), 15-18.
5. 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.
6. Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994). Continuous univariate distributions. Volume 1, 2nd Edition. John Wiley, New York.
7. Lin, J. T. (1990). A simpler logistic approximation to the normal tail probability and its inverse. Applied Statistics, 39, 255-257.
8. Ordaz, M. (1991). A simple approximation to the Gaussian distribution. Structural Safety, 9(4), 315-318.
9. Polya, G. (1949). Remarks on computing the probability integral in one and two dimensions. Proceedings of the First Berkeley Symposium on Mathematical and Statistical Probability. 63-78.
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.