Pak.j.stat.oper.res. Vol.21 No. 2 2025 pp 119-128 DOI: https://doi.org/10.18187/pjsor.v21i2.4949

Pakistan Journal of Statistics and Operation Research

New highly accurate improvements for single-term approximations of the standard normal distribution function

Mohammed Obeidat[1*], Rema Al-Jamal[2], Ahmad A Hanandeh[3]

* Corresponding Author

1- Department of Statistics, Yarmouk University, Jordan, mohammad.obidat@yu.edu.jo. Orcid: https://orcid.org/0000-0002-1170-1019
2. Department of Statistics, Yarmouk University, Jordan, 2021107025@ses.yu.edu.jo
3. Department of Mathematics, Faculty of Science, Islamic University of Madinah, Madinah 42351, Saudi Arabia, ahmad.hanandeh@yu.edu.jo

Abstract

This paper proposed new, highly accurate, single-term, and explicitly invertible approximations for the standard normal distribution function and its related functions, such as the error function and the quantile function. The proposed approximations are built based on some existing approximations, however, the proposed ones are much more accurate. The accuracy of the proposed approximations is measured via maximum absolute error and mean absolute error. Some of the proposed approximations are at least five times more accurate than the original ones and two of them have maximum absolute error lower than 1.8×10^{-4}, which is quite sufficient for most of real-world applications. Two real applications are studied to show the applicability of the proposed improvements. These applications showed the superiority of one of the proposed approximations over some of the available single-term approximations even though the latter have smaller maximum absolute error.

Key Words: Absolute Error, Approximation, Distribution Function, Error Function, Standard Normal.

Mathematical Subject Classification: 62G05, 62E10, 62E17

1. Introduction

The standard normal distribution function and other related functions such as the quantile function, the error function, and the complementary error function are considered among the most important functions in statistics and engineering. For example, the standard normal distribution function is widely useful in statistical theory and modelling (Devore, 2011). The Error function frequently appears in transport phenomena in chemical engineering (Bird, 2002) and in digital phase modulation and signal processing (Proakis, 2001). In these applications and many others, it is of significant importance to have a closed form approximation of the aforementioned functions. For a detailed review of applications and approximations of the normal cumulative distribution function and related functions, we strongly recommend Soranzo et al. (2023), Eidous and Abu-Shareefa (2019), and Eidous and Al-Rawwash (2025). In addition to approximating the distribution function of the standard normal distribution, it is of great importance also to have lower and upper bounds of the distribution function, see for example Eidous (2023) and Ananbeh and Eidous (2024).

The standard normal distribution function is defined as the probability that a normally distributed random variable X does not exceed a pre-specified value x and is denoted as:

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

There are three functions that are related to $\Phi(x)$ and are very popular in statistical learning, engineering, mathematics, and machine learning. These functions are

1- The Q-function:

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

2- The error function:

$$\operatorname{erf}(x) = \int_0^x \frac{2}{\sqrt{\pi}} e^{-\frac{t^2}{2}}\, dt, \\ = 2\Phi\left(x\sqrt{2}\right) - 1.$$

3- The complementary error function:

$$\operatorname{erfc}(x) = \frac{1}{2} + \int_x^\infty \frac{2}{\sqrt{\pi}} e^{-\frac{t^2}{2}} \\ = 1 - \operatorname{erf}(x).$$

None of the integrals that appear above can be evaluated in closed form, hence several attempts have been made to approximate these functions and/or finding upper or lower limits for them. Existing approximations can be classified into two types: (1) complex, highly accurate, but not explicitly invertible, and (2) simple, invertible, but less accurate. This research aims to improve some of the already published simple approximations of the standard normal distribution function, $\Phi(x)$, to have relatively high accuracy and yet very simple to deal with. Due to the symmetry property of the normal density, it is sufficient to approximate its distribution function for $x > 0$.

The accuracy of the proposed improvements will be measured via the maximum absolute error (Max.AE) and mean absolute error (Mean.AE) which are defined as follows. Let $\Phi(x)$ and $\widehat{\Phi}(x)$ be the true and approximate value of the distribution function respectively. Then

$$\text{Max.AE} = max\left|\Phi(x) - \widehat{\Phi}(x)\right|.$$

The Mean.AE is defined as follows:

$$\text{Mean.AE} = \frac{\sum|\Phi(x_i) - \widehat{\Phi}(x_i)|}{M}$$

where x_i is taken to be between 0 to 7 with step 0.0001, and M is the number of x_i points. Note that the built-in R function "*pnorm*" is used to find the true value $\Phi(x)$ for any given x (R Core Team, 2023).

The rest of this paper is organized as follows. Section 2 discusses some literature review on simple approximations. The proposed improvements are discussed in Sections 3. In Section 4 some applications are presented. Section 5 concludes the paper with discussion.

2. Literature Review

Pólya (1945) suggested a formula of the form $\widehat{\Phi}_1(z) = a + b\sqrt{(1 - e^{-c\,z^2})}$, where $a = b = 0.5$, and $c = \frac{2}{\pi}$ with Max.AE= 3.2×10^{-3}. This approximation was originally proposed by Pólya as an upper bound not as an approximation of the standard normal distribution function. However, many authors have revised and improved it as an approximation, see for example Aludaat and Alodat (2008), Eidous and Al-Salman (2016), and Hanandeh and Eidous (2021). Among these approximations, Abderrahmane and Kamel (2017) approximation has the smallest Max.AE, where a = 0.50103, $b = 0.49794$, and $c = 0.62632$, with Max.AE = 1.03×10^{-3} and Mean.AE = 6.64×10^{-4}.

Burr (1967) gave the following approximation: $\widehat{\Phi}_2(z) = 1 - (1 + (a + bz)^c)^{-d}$ where $a = 0.644693$, $b = 0.161984$, $c = 4.874$, $d = 6.158$. This approximation produced Max.AE = 3.97×10^{-3} and Mean.AE = 5.89×10^{-4}

Ordaz (1991) proposed approximation of the form $\widehat{\Phi}_3(z) = 1 - ae^{-\left(-\frac{9z}{14}-\frac{8}{14}\right)^2}$, with a = 0.6931. The Max.AE and Mean.AE of this approximation are 4.1×10^{-3} and $5.5*10^{-4}$. respectively. This formula is then revised by many authors including Hanandeh and Eidous (2022) who improved Ordaz approximation by choosing a = 0.688182 which gives Max.AE = 3.53×10^{-3} and Mean.AE = 4.1×10^{-4}.

Chernoff (1952) proposed the approximation of the form

$$\widehat{\Phi}_4(z) = 1 - ae^{-b\,z^c},$$

where a = 0.5, b = 0.5 and c = 2. The Max.AE and Mean.AE of this approximation are 1.51×10^{-1} and 3.2×10^{-2}, respectively. Many authors revised this formula and the one with the smallest Max.AE was Hanandeh and Eidous (2022) with a = 0.5, $b = 1.2$, $c = 1.275247$ and produces Max.AE = 9.17×10^{-3}, Mean.AE = 2.22×10^{-3}

Soranzo and Epure (2014) proposed the approximation

$$\widehat{\Phi}_5(z) = 2^{-a^{1-b^{cz}}},$$

where a = 22, $b = 41$, $c = \frac{1}{10}$. This approximation gave a Max.AE = 1.274×10^{-4} and Mean.AE = 3.834×10^{-05}

Kundu et al. (2006) proposed approximation of the form

$$\widehat{\Phi}_6(z) = \left(1 - e^{-e^{a\,z+b}}\right)^c,$$

where a = 0.3820198, $b = 1.0792510$, $c = 12.8$, with Max.AE = 3.2×10^{-4}.

Tocher (1963) proposed the approximation

$$\widehat{\Phi}_8(z) = 1 - \frac{1}{1+e^{a\,z}},$$

where a = $\sqrt{\frac{8}{\pi}}$ and with Max.AE = 1.8×10^{-2}. This formula is then revised by many authors including Bowling et al. (2009) by choosing a = 1.702 which gives Max.AE = 9.5×10^{-3}.

Recently, Lipoth et al. (2022) proposed the approximation

$$\widehat{\Phi}_7(z) = \left(1 + a\left(ln\left(1 + e^{-\frac{z}{h}+c}\right)\right)^b\right)^{-d}$$

where a = 0.00161826615, b = 3.38692114553, c = 3.26862849061, d = 7.80500878654, and h = 0.82116764005, with Max.AE = 2.4×10^{-5}

3. The Proposed Improvements

In this section, we propose five updated and highly accurate single-term approximations of the standard normal distribution function. These approximations can be easily manipulated to obtain approximations to the error function, complementary error function, and quantile function. Moreover, the proposed improvements can be easily inverted for the purpose of random numbers generation and finding critical values. Note that to obtain the new approximations, we first introduced new unknown constants to the available approximations. These constants along with the original constants in the approximations are then obtained by optimizing the Kolmogrov-Smirnov statistics. The initial guess of these constants are taken to be the constants of the original approximations. We then search for better values of the constants around the optimal solution using R loops to improve the Max.AE.

3.1 Improvement of Pólya's approximation

Our new proposed improvement of Pólya's formula takes the general form:

$$\hat{\Phi}_{1new}(z) = a + b\,(1 - e^{-c\,z^{h}})^{d},$$

with the following constants, $a = b = 0.5, c = 0.668224, d = 0.531504$, and $h = 1.89375$. This approximation produced a Max.AE = 1.73×10^{-4} and Mean.AE = 5.25×10^{-5}. The newly proposed approximation is approximately six times more accurate than the old best one

3.2 Improvement of Burr's approximation

Our proposed approximation of Burr's formula is,

$$\hat{\Phi}_{2new}(z) = 1 - (1 + (a + bz)^{c})^{-d}$$

Where the constants are $a = 0.6446348$, $b = 0.1566217$, $c = 4.9187329$, and $d = 6.3597984$. This approximation produce a Max.AE = 6.06×10^{-4} and Mean.AE = 2.12×10^{-4}. The newly proposed approximation is approximately 6.5 times more accurate than Burr's approximation.

3.3 Improvement of Ordaz' approximation

The proposed improvement is,

$$\hat{\Phi}_{3new}(z) = 1 - ae^{-(bz+c)^{d}},$$

where $a = 0.74$; $b = 0.6154381$; $c = 0.6272601$; $d = 2$. This approximation produced a Max.AE = 7.27×10^{-4}, Mean.AE = 2.2×10^{-4}. This approximation is about 4.8 times more accurate than Hanandeh and Eidous (2022) approximation

3.4 Improving Chernoff's approximation

The proposed improvement takes the same form of Chernoff's approximation, that is

$$\hat{\Phi}_{4new}(z) = 1 - ae^{-b\,z^{c}}$$

with $a = 0.493376, b = 1.1729, c = 1.3037$. This choice of the constants produce a Max.AE = 6.65×10^{-3}, Mean.AE = 1.81×10^{-3}.

3.5 Improvement Soranzo and Epure approximation

The proposed improvement takes the form,

$$\hat{\Phi}_{5new}(z) = 2^{-d*a^{1-b^{cz}}},$$

where a = 22.2118, b=39.59191, c=0.100688, d=1.0001336. The new approximation gave a Max.AE = 1.11×10^{-4} and Mean.AE = $3.988 10^{-5}$.

Figure 1 illustrates the error and absolute error of the 5 proposed approximations for $0 \le z \le 7$. It can be seen from that $\hat{\Phi}_{1new}$ and $\hat{\Phi}_{5new}$ are the best very close in performance followed by $\hat{\Phi}_{2new}$ and $\hat{\Phi}_{3new}$.

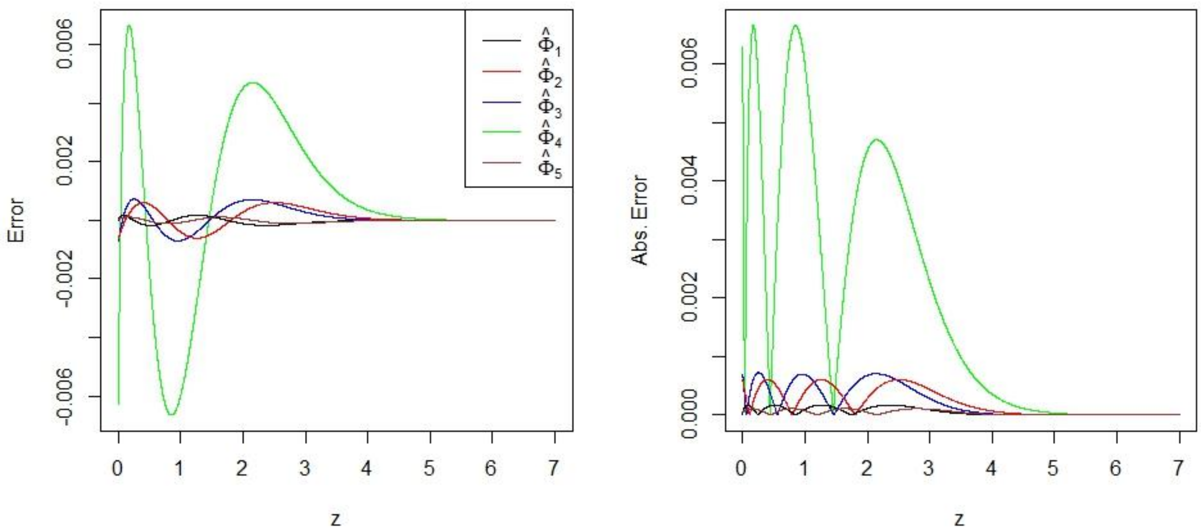

Figure 1: Error (left) and Absolute Error (right) of the 5 proposed approximations

4. Applications

Two real applications will be considered here to compare the performance of the proposed approximations. The first example is in the area of process control and the second is about the application of the error function in digital communications.

Example 1: The first example will discuss a statistical process control of a chemical product. The example was first discussed by Aljebory and Alshebeb (2014) and Alkhazali et al. (2020). The aim is to determine the average run length to detect a shift that may occur in a product pH form a mean of $\bar{X}$ =9.26 to $\bar{X}$ =10. In this case the true shift is 10-9.26=0.74. Figure 2 shows a snapshot of the product control chart (obtained from Aljebory and Alshebeb (2014)).

Using the fact that the data in normal distribution fall within 3 standard deviation from the mean we have

$$UCL - \bar{X} = 3\sigma_{\bar{X}}$$

$$10.278 - 9.26 = 3\sigma_{\bar{X}}$$

$$\sigma_{\bar{X}} = 0.339,$$

Therefore, the shift in terms of $\sigma_{\bar{X}}$ is $\frac{0.74}{0.339} = 2.18$.

Following Alkhazali et al. (2020), the probability of detecting the shift will be $\beta = \Phi(0.82) - \Phi(-5.18) = 0.7938918$. Table 1 gives the approximate value of this probability using the proposed approximations. Although the modified Soranzo and Epure approximation, $\hat{\Phi}_{5new}(\mathrm{z})$, has the smallest Max.AE, $\hat{\Phi}_{1new}$ and $\hat{\Phi}_{2new}$ provided a more accurate approximation in this example.

Table 1: Approximated value and error of the probability for the new five approximations.

Approximation	Approximated value	Error ($\Phi(x) - \hat{\Phi}(x)$)
$\hat{\Phi}_{1new}$	0.7939175	-0.000025

$\widehat{\Phi}_{2new}$	0.7939525	-0.00006
$\widehat{\Phi}_{3new}$	0.794508	-0.00062
$\widehat{\Phi}_{4new}$	0.8005355	-0.0066
$\widehat{\Phi}_{5new}$	0.7940018	-0.00011

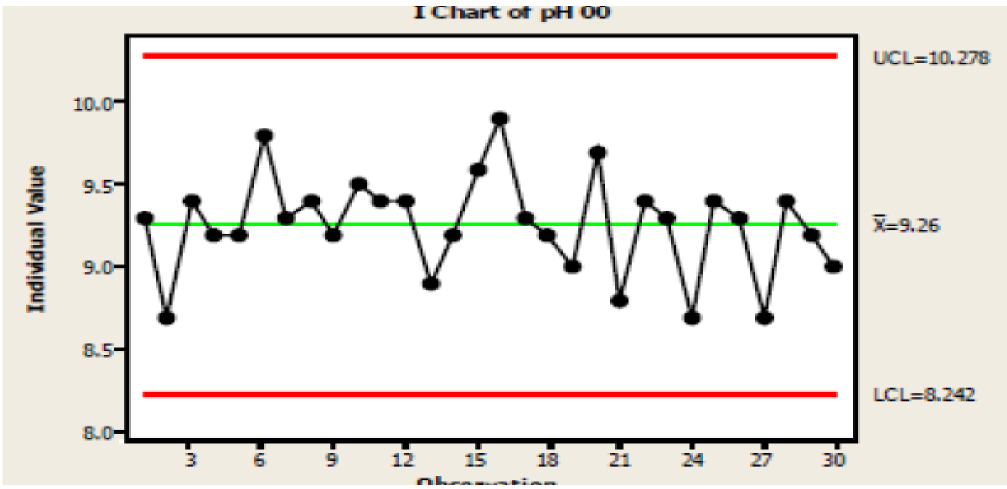

Figure 2: Control chart of the product pH levels (Aljebory and Alshebeb, 2014).

Example 2: : The second example is in the field of digital communications, see Sandoval-Hernandez et al. (2019). In digital phase modulation, Bit Error Rate (BER) is a key performance metric. It measures the ratio of erroneous bits to the total transmitted bits over a communication channel. It is considered a crucial metric for evaluating and optimizing the reliability of data transmission. The probability of BER in binary Phase Shift Keying is defined as

$$P = \frac{1}{2} erfc\left(\sqrt{\frac{E_b}{N_0}}\right)$$

$$= 1 - \Phi\left(\sqrt{\frac{2E_b}{N_0}}\right),$$

Where, E_b and N_0 are the energy in one bit and the additive white Gaussian noise, respectively. The ratio $\frac{E_b}{N_0}$ is called the signal to noise ratio (SNR).

Figure 3 shows the performance of the proposed approximation in estimating BER. It is clear that $\widehat{\Phi}_{1new}$ perform the best of the five proposed approximations followed by $\widehat{\Phi}_{5new}$ and $\widehat{\Phi}_{3new}$. It can be also observed that in this application $\widehat{\Phi}_{1new}$outperform Lipoth's approximation even though the latter has smaller Max.AE. Note that the R code used to produce this figure are converted from the *Matlab* code of Sandoval-Hernandez et al. (2019).

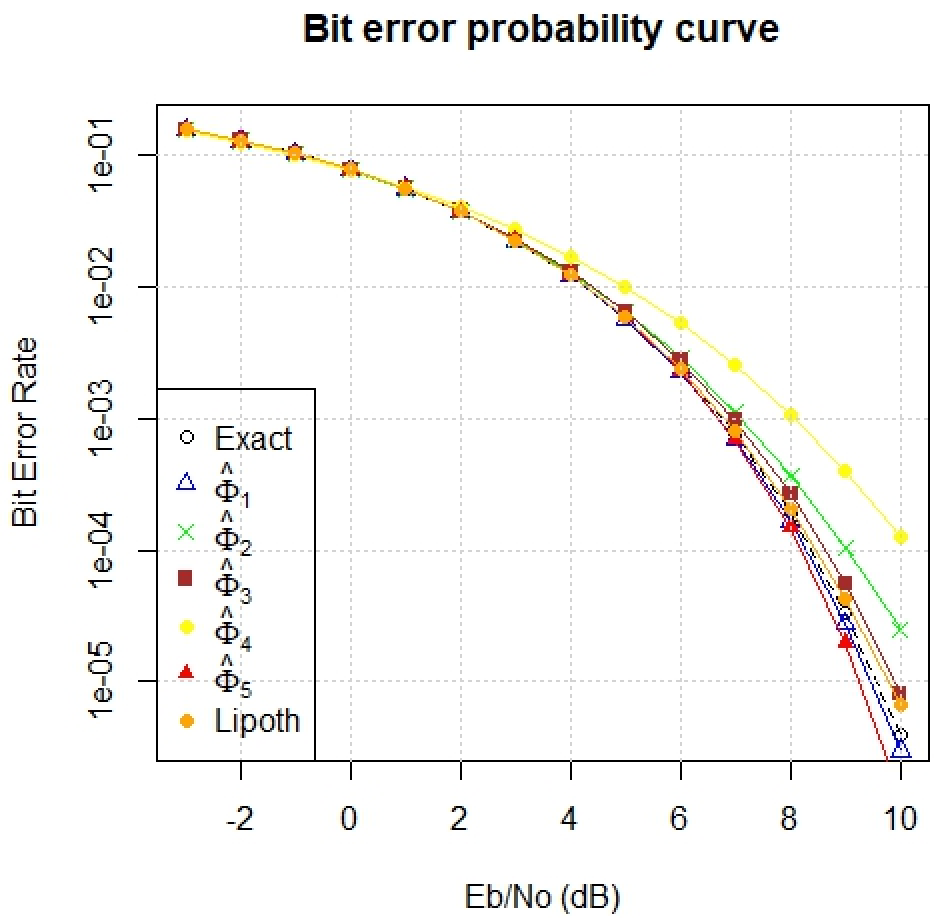

Figure 3: Bit error probability curve for Phase Shift Keying modulation.

5. Conclusion

In this article, five one-term improvements to the standard normal distribution function approximations are proposed and investigated. These improvements are compared with the approximation available in the R software using the maximum and mean absolute error. The R approximation to the standard normal distribution function is considered as the true value.

It is worth mentioning that the improved approximations considered here have simple single-term analytical form which makes them easy to program and easy to invert and yet have sufficiently high accuracy. They also have a simple form, consisting of a single term involving *x*. Therefore, the proposed improvements are considered very competitive approximations to other highly accurate but complex approximations available in the literature. The Max. AEs of some of the proposed improvements is four to six times better than previously proposed improvements which considered relatively high in single-term approximations. Two of the proposed improvements have Max.AE less than 1.8×10^{-4}. Two real examples are discussed to show the applicability of the proposed improvements. These applications showed that $\widehat{\Phi}_{1new}$ outperforms existing approximations even though they have smaller Max.AE.

Funding: No funding was received for conducting this study.

Data availability: There is no data used in this article

Code availability: The codes in this paper represent a new development on R statistical software and can be obtained from the corresponding author upon request.

Declarations Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Ethical statements: We hereby declare that, this manuscript is the result of our independent creation under the reviewers' comments. Except the quoted contents, this manuscript does not contain any research achievements by other individuals or groups. We are the only authors of this manuscript, that have been published or written. The legal responsibility of this statement shall be borne by us.

References

1- Abderrahmane, M. and Kamel, B. (2017). A new approximation to standard normal distribution function. Journal of Computational and Applied Mathematics, 6(2), 1-18.

2- Aljeborya, K. M. and Alshebeb, M. (2014). Integration of statistical and engineering process control for quality improvement (A case study: chemical industry - national chlorine industries). Jordan Journal of Mechanical and Industrial Engineering, 8(4), 243 -256.

3- Alkhazali, A., Al-Rabayah, M. and Hamasha, M. M. (2020). A new high accuracy mathematical approximation to the cumulative normal density function. Jordan Journal of Mechanical and Industrial Engineering, 13, 247-251

4- Aludaat, K. M. and Alodat, M.T. (2008). A note on approximating the normal distribution function. Applied Mathematical Sciences, 2(9), 425-429.

5- Ananbeh, E. & Eidous, O. (2024). New Simple Bounds for Standard Normal Distribution Function. *Communications in Statistics – Simulation and Computation*. https://doi.org/10.1080/03610918.2024.2326596.

6- Bird R. B. (2002). Transport phenomena, Applied Mechanics Reviews, 55(1), R1–4.

7- 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.

8- Chernoff, H. (1952). A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. The Annals of Mathematical Statistics, 493-507.

9- Devore J. L. (2011). Probability and statistics for engineering and the sciences. Cengage Learning.

10- Eidous, O. (2023). New Tightness Lower and Upper Bounds for the Standard Normal Distribution Function and Related Functions. Mathematical Methods in the Applied Sciences, 46(14), 15011–15019. https://doi.org/10.1002/mma.9358

11- Eidous, O. & Al-Rawwash, M. (2025). Approximation for Standard Normal Distribution Function and Its Invertible. Journal of Algorithms & Computational Technology, 19, 1–5. https://doi.org/10.1177/17483026251322100

12- Eidous, O. and Al-Salman, S. (2016). One-term approximation for normal distribution function, Mathematics and Statistics 4(1),15-18.

13- Eidous, O. and Abu-Shareefa, R. (2019). New approximations for standard normal distribution function. Communications in Statistics - Theory and Methods, 49(6), 1357-1374.

14- Hanandeh, A. A. and Eidous, O. (2021). A new one-term approximation to the standard normal distribution. Pakistan Journal of Statistics and Operation Research, 1**7**(2), 381-385.

15- Kundu, D., Manglick, A. and Gupta, R. D. (2006). A convenient way of generating normal random variables using generalized exponential distribution. Journal of Modern Applied Statistical Methods, 5, 266-272.

16- Lipoth, J., Tereda, Y., Papalexiou, S. M. and Spiteri, R. J. (2022). A new very simply explicitly invertible approximation for the standard normal cumulative distribution function. AIMS Math, 7(7), 1-9.

17- Pólya, G. (1945). Remarks on computing the probability integral in one and two dimensions. Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability, 1, 63-78.

18- Proakis J. G., (2001). Digital communications, McGraw-Hill Book Co., New York.

19- R Core Team. (2023). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.R-project.org/

20- Sandoval-Hernandez, M. A., Vazquez-Leal, H., Filobello-Nino, U. and Hernandez-Martinez, L. (2019). New handy and accurate approximation for the Gaussian integrals with applications to science and engineering. Open Mathematics, 17(1), 1774-1793.

21- Soranzo, A. and Epure, E. (2014). Very simply explicitly invertible approximations of normal cumulative and normal quantile function. Applied Mathematical Sciences, 8(87), 4323-4341.
22- Soranzo, A., Vatta, F., Comisso, M., Buttazzoni, G. and Babich, F. (2023). Explicitly invertible approximations of the Gaussian q-function: a survey. IEEE Open Journal of the Communications Society, 4, 3051-3101.
23- Tocher, K. D. (1963). The art of simulation. The English Universities Press: London, UK

Appendix 1: The R code used to produce Figure 3 in Example 2.

```
# Polya
f1=function(x)
{
par=c(0.668224,1.89375,0.531504)
res=0.5*(1+(1-exp(-par[1]*x^par[2]))^par[3])
return(res)
}

# Bur
f2=function(x)
{
par=c(0.6446348, 0.1566217, 4.9187329, 6.3597984)
res=  1-(1+(par[1]+par[2]*x)^par[3])^(-par[4])
return(res)
}

#Ordaz
f3=function(x)
{
a1=a = 0.74; a2=  0.6154381; a3=0.6272601; a4=2;a5=1
res= (1-a1*exp(-(a2*x+a3)^a4))^a5
return(res)
}


#Chernof
f4=function(x)
{
a=0.493376;b=1.1729;c=1.3037
res=  1-a*exp(-b*x^c)
return(res)
}


#soranzo
f5=function(x)
{
a=22.2118;b= 39.59191;c=  0.100688;d=  1.0001336
res=  2^(-d*a^(1-b^(c*x)))
return(res)
}
```

```
# Lipoth
f6=function(x)
{
a=0.00161826615;b=3.38692114553
c=3.26862849061;d=7.80500878654;h=0.82116764005
res=(1+a*(log(1+exp(-x/h+c)))^b)^(-d)
return(res)
}
set.seed(100)
N <- 10^7

ip <- runif(N) > 0.5
s <- 2 * as.numeric(ip) - 1
set.seed(200)
n <- (1 / sqrt(2)) * (rnorm(N) + 1i * rnorm(N))
Eb_N0_dB <- -3:10

for (ii in 1:length(Eb_N0_dB)) {
  y <- s + 10^(-Eb_N0_dB[ii] / 20) * n

  ipHat <- as.numeric(Re(y) > 0)
}

library(ggplot2)
library(pracma)

nErr <- numeric(length(Eb_N0_dB))
for (ii in 1:length(Eb_N0_dB)) {
  nErr[ii] <- sum(ip != ipHat)
}

simBer2 <- 0.5 * (1 - erf(sqrt(10^(Eb_N0_dB / 10))))
theoryBer <- 0.5 * erfc(sqrt(10^(Eb_N0_dB / 10)))

plot(Eb_N0_dB, theoryBer, type="o", col="black", lty=2, log="y",
     xlab="Eb/No (dB)", ylab="Bit Error Rate", main="Bit error probability curve")
#points(Eb_N0_dB, simBer2, type="o", col="blue", pch=4)
points(Eb_N0_dB, 1-f1(sqrt(2*10^(Eb_N0_dB / 10))), type="o", col="blue", pch=2)
points(Eb_N0_dB, 1-f2(sqrt(2*10^(Eb_N0_dB / 10))), type="o", col="green", pch=4)
points(Eb_N0_dB, 1-f3(sqrt(2*10^(Eb_N0_dB / 10))), type="o", col="brown", pch=15)
points(Eb_N0_dB, 1-f4(sqrt(2*10^(Eb_N0_dB / 10))), type="o", col="yellow", pch=16)
points(Eb_N0_dB, 1-f5(sqrt(2*10^(Eb_N0_dB / 10))), type="o", col="red", pch=17)
points(Eb_N0_dB, 1-f6(sqrt(2*10^(Eb_N0_dB / 10))), type="o", col="orange", pch=19)


grid()
legend("bottomleft",                                   legend=c("Exact",
expression(hat(Phi)[1]),expression(hat(Phi)[2]),expression(hat(Phi)[3]),expression(hat(Phi)[4]),expression(hat(Phi)[
5]),"Lipoth"), col=c("black", "blue","green","brown","yellow","red","orange"), pch=c(1,2,4,15,16,17,19))
```