A Simple Approximation for the Normal Distribution Function via Variational Iteration Method

V. K. Shchigolev
Ulyanovsk State University, 42 L. Tolstoy Str., Ulyanovsk 432000, Russia
vkshch@yahoo.com

Abstract:
In this paper, we obtain some new approximations for the cumulative distribution function of the standard normal distribution via the He's Variational Iteration Method. For this end, we consider the cumulative distribution function as the unknown function to be determine by solving a certain differential equation of the second order that the cumulative distribution function satisfied subjected with the certain initial conditions. The correction functional in this approach is constructed here in such a manner that we have one real numerical parameter to be tuned for the best result. Our approximations to the cumulative distribution function are comparable to other approximations found in the literature and has the advantage of being a simple expression, that may have potential applications in several areas of applied sciences. Numerical comparison shows that our approximations are very accurate.

Keywords: Normal Distribution, Cumulative Distribution Function, Approximations, Variational Iteration Method

1 Introduction

As well known, the normal distribution is one of the most important distributions in statistics, and it is the result of the central limit theorem [10]. A wide usage of this distribution is due to the fact that it is an infinitely divisible continuous distribution with finite variance. Therefore, some others distributions, for example, the binomial and Poisson ones, approach it in the limit. Moreover, many non-deterministic physical processes are modeled by this distribution. Being at the core of the central limit theorem, the normal distribution is often used to represent random variables whose distributions are not known. It has a wide use in several branches of probability theory, mathematical statistics, statistical physics, engineering, financial mathematics, and so on.

The normal distribution, also called the Gauss or Gauss - Laplace distribution, is the probability distribution, which in the one-dimensional case is given by a probability density function that coincides with the Gauss function:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\dfrac{(x-\mu)^2}{2\sigma^2}}, \tag{1}$$

where the parameter μ is the mean or mathematical expectation of the distribution, the median and the mode of distribution, and the parameter σ is the standard deviation (σ^2 is the dispersion or the variance) of the distribution.

The standard normal distribution is called the normal distribution with the expectation $\mu = 0$ and the standard deviation $\sigma = 1$. If a random variable X is normally distributed with zero mean and the standard deviation, then its probability density function (1) is,

$$f(x) = \frac{1}{\sqrt{2\pi}}\, e^{-\dfrac{x^2}{2}}. \tag{2}$$

The cumulative distribution function (CDF) of the standard normal distribution is the integral

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\dfrac{t^2}{2}}\, dt. \tag{3}$$

Unfortunately, $\Phi(x)$ cannot be expressed in a closed form in terms of elementary functions for all values of x which would be useful for the practical needs, many numerical approximations for $\Phi(x)$ are known. Lots of approximation formulas to the cumulative normal distribution $\Phi(x)$ have been proposed earlier (see for example, [6], and references therein). Most of them are based on the form of series expansion which, in theory, can approximate CDF with arbitrarily high precision by increasing the number of terms [12, 13, 21]. Another approximations are known as the so-called ad hoc approximations. They often take simple forms with few numerical coefficients [5]-[22]. The CDF (3) can be represented in the following form

$$\Phi(x) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int\limits_{0}^{x} e^{-\dfrac{t^2}{2}}\, dt, \tag{4}$$

and one can see that the problem of approximating CDF is equivalent to the problem of approximating the integral in equation (4). There are several methods (see for example, [20]) which provide an approximation of the integral by different numerical methods: Taylor series, asymptotic series, continual fractions, and some others. For more, several other papers should be mentioned where different approximate formulae were obtained for $\Phi(x)$, such as [1]-[22].

It could be mentioned that in [20] the use of the homotopy perturbation method (HPM) is proposed to calculate an approximate analytical solution of the normal distribution integral. Besides, after solving the Gaussian integral by HPM, the result can serve as base to solve other integrals like error function and the cumulative distribution function. The basic idea of [20] is that the integral similar to (4) can be reformulated as a differential equation subjected to a certain initial condition. Being applied to equation (4), this means that it is necessary to solve

$$\Phi'(x) - \frac{1}{\sqrt{2\pi}}\, e^{-\dfrac{x^2}{2}} = 0 \tag{5}$$

with the initial condition $\Phi(0) = \frac{1}{2}$ that follows directly from (4).

At the same time, one more approximate method developed by J.-H. He, the so-called Variational Iteration Method (VIM) [7, 8, 9], can give an approximation of high accuracy in many nonlinear problems, and can be easily implemented. Indeed, this method has been successfully applied for the problems of geodesic motion in the gravitational field of astrophysical objects by the author in Ref. [16].

Recently, for the analytical calculation of the cosmological luminosity distance, we offered to proceed from the solution of differential equation with the certain initial conditions instead of calculating the corresponding integral. For this purpose, we obtained the differential equation which the luminosity distance should satisfy to, and define the appropriate initial conditions for this equation. We have showed that by using the VIM, the explicit dependency of luminosity distance on red-shift in arbitrary accuracy can be easily obtained by implementing a simple procedure for the governing equation [17].

The purpose of the present paper is the application of He's VIM in obtaining the approximations for the cumulative normal distribution function. For this end, we consider $\Phi(x)$ as the unknown function to be determine from solving a certain differential equation. Indeed, taking into account equations (4) and (5), one can get the following Cauchy problem for $\Phi(x)$:

$$\Phi''(x) + x\,\Phi'(x) = 0\,; \quad \Phi_{\big|x=0} = \frac{1}{2}, \;\; \Phi'_{\big|x=0} = \frac{1}{\sqrt{2\pi}}, \tag{6}$$

where the prime stands for the derivative with respect to x.

2 Approximating CDF via VIM

The main equation (6) is a linear differential equation of the second order. It can be solved exactly in quadratures, but the result again leads to the formula (4). Therefore, we will solve this equation analytically, but with a certain approximation. Among all kinds of approximate methods we now use the VIM.

The VIM has now become standard and the reader is referred to [7]-[9] for the basic ideas of VIM. We only note that in applications of VIM to differential equations one should undertake the following three steps: (i) establishing the correction functional; (ii) identifying the Lagrange multipliers; (iii) determining the initial iteration. For the convergence criteria and error estimates of the VIM, one can refer the reader to [19]-[18].

The correction functional for equation (6) can be written as [7]

$$\Phi_{n+1}(x)=\Phi_n(x)+\int_0^x \lambda_x(s)\left[\Phi_n''(s)+\beta\Phi_n'(s)+(s-\beta)\tilde{\Phi}_n'(s)\right]ds, \tag{7}$$

where $\lambda_x(s)$ is a Lagrange multiplier, that can be identified optimally via variational iteration method. Here, $\tilde{\Phi}_n$ is considered to be a restricted variation which means that $\delta\tilde{\Phi}_n=0$. Making the correction functional (7) stationary, yields

$$\begin{aligned}\delta\Phi_{n+1}(x) &= \delta\Phi_n(x)+\delta\int_0^x \lambda_x(s)\left[\Phi_n''(s)+\beta\,\Phi_n'(s)+(s-\beta)\;\tilde{\Phi}_n'(s)\right]ds \\ &= \delta\Phi_n(x)+\delta\int_0^x \lambda_x(s)\left[\Phi_n''(s)+\beta\Phi_n'(s)\right]ds. \end{aligned} \tag{8}$$

Its stationary conditions, $\delta\Phi_{n+1}(x)=0$, can be obtained using integration by parts in equation (8) and noticing that $\delta\Phi_n(0)=0$. The Lagrange multipliers can be easily and precisely obtained for linear problems.

The successive approximations $\Phi_{n+1}(x)$ of the solution will be readily obtained upon using the obtained Lagrange multiplier and by using any appropriate function for $\Phi_0(x)$. The zeroth approximation $\Phi_0(x)$ may be selected by any function that just meets, at least, the initial and boundary conditions. Therefore by starting from $\Phi_0(x)$, the exact solution may be obtained as

$$\Phi(x)=\lim_{n\to\infty}\Phi_n(x). \tag{9}$$

Using integration by parts in equation (8) , and noticing that $\delta\Phi_n(0)=\delta\Phi_n'(0)=0$, we can rewrite this equation as follows

$$\delta\Phi_{n+1}(x)=\left[1+\lambda_x(s)-\lambda_x'(s)\right]\delta\Phi_n(s)\big|_{s=x}+\lambda_x(s)\delta\Phi_n'(s)\big|_{s=x}+\int_0^x\left[\frac{d^2\lambda_x(s)}{ds^2}-\beta\frac{\lambda_x(s)}{ds}\right]\delta\Phi_n(s)ds. \tag{10}$$

The stationary conditions $\delta\Phi_{n+1}(x)=0$ applied to this equation yield the following equations for the Lagrange multiplier

$$\begin{cases}\lambda_x'(s)\big|_{s=x}-\beta\,\lambda_x(s)\big|_{s=x}-1=0,\\ \lambda_x(s)\big|_{s=x}=0,\\ \lambda_x''(s)-\beta\,\lambda_x'(s)=0.\end{cases} \tag{11}$$

These equations can be readily solved that yields the following Lagrange multiplier

$$\lambda_x(s) = \frac{1}{\beta}\left(e^{\beta(s-x)} - 1\right). \tag{12}$$

Then, according to (7), the successive approximations $\Phi_n(x)$ of the solution for the main equation (6) can be readily obtained upon using the obtained Lagrange multiplier (12) and by using any appropriate function for $\Phi_0(x)$ as

$$\Phi_{n+1}(x)=\Phi_n(x) + \frac{1}{\beta}\int_0^x\left(e^{\beta(s-x)} - 1\right)\left[\Phi_n''(s) + s\Phi_n'(s)\right]ds. \tag{13}$$

Taking into account equations (6) and (7), let us now determine the initial iteration $\Phi_0(x)$ as the exact solution of the following linear equation

$$\Phi_0''(x) + \beta\,\Phi_0'(x) = 0$$

with the following initial conditions

$$\Phi_0(0) = \frac{1}{2},\ \ \Phi_0'(0) = \frac{1}{\sqrt{2\pi}}.$$

One can easily obtain

$$\Phi_0(x) = \frac{1}{2} + \frac{1}{\beta\sqrt{2\pi}}\left(1 - e^{-\beta x}\right). \tag{14}$$

Substituting $\Phi_0(x)$ from equation (14) into the formula (13) at $n = 0$, we can obtain the simplest approximate solution $\Phi \approx \Phi_1$ as follows

$$\Phi(x) = \frac{1}{2} + \frac{1}{\beta\sqrt{2\pi}}\left[\frac{2\beta^2 - 1}{\beta^2} + \left(\frac{x^2}{2} - \frac{\beta^2 - 1}{\beta}x - \frac{2\beta^2 - 1}{\beta^2}\right)e^{-\beta x}\right]. \tag{15}$$

In principle, this procedure can be continued as far as desired, and the approximation will convergence to its exact solution. In order to demonstrate the accuracy of the method applied, the graphs of the approximate solutions for $\Phi(x)$ according (15) for different values of β and the numerical solution to equation (4) via the Maple package are given in Fig. 1.

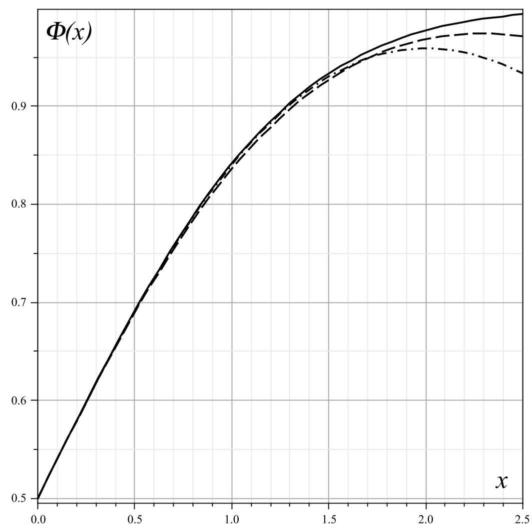

Figure 1: Comparison of the approximate solutions, given by formula (15) for $\beta = 0.5$ (dash-dot line) and $\beta = 1/\sqrt{2}$ (dash line), with the exact numerical solution to Eq. (4) (solid line).

The only parameter that can be adjusted in (15) (for example, by using the NonlinearFit command from Maple Release 15) to obtain a good approximation is λ. As a consequence, this adjustment allows us to ignore a large number of successive terms in a good approximation. Here, we present expression (15) only in two cases as the illustrative examples, when the formulae take the most simple form. Indeed, from (15) we have

$$\Phi(x) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \left[\left(x^2 + 3\,x + 4\right) e^{-\frac{x}{2}} - 4 \right], \tag{16}$$

when $\beta = 0.5$, and

$$\Phi(x) = \frac{1}{2} + \frac{1}{2\sqrt{\pi}} \left(x^2 + \sqrt{2}\,x\right) e^{-\frac{x}{\sqrt{2}}}, \tag{17}$$

when $\beta = 1/\sqrt{2}$. The absolute errors of these approximations compared to the numerical solution of equation (4) are plotted in Fig. 2.

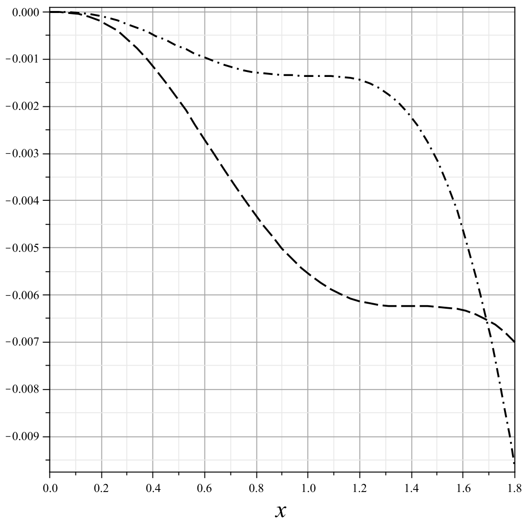

Figure 2: The absolute errors of the approximate solutions, given by equation (16) (dash-dot line) and equation (17) (dash line), compared to the exact numerical solution to equation (4).

It should be noted that approximation (14), and, consequently, particular formulas (15) and (16), were obtained in a single iteration. It is clear that subsequent iterations will significantly improve the accuracy of the approximation, while at the same time complicating the expression for CDF.

In view of the complexity of the general expression for the second iteration, we give the following approximation for the case of the $\beta = 1/\sqrt{2}$, which demonstrates the best result with relative simplicity of the expression in the above result. Substituting equation (16) into (12), one can obtain the next approximation $\Phi \approx \Phi_2$ for the case $\beta = 1/\sqrt{2}$ as follows

$$\Phi(x) = \frac{1}{2} + \frac{1}{\sqrt{\pi}} \left[7 - \left(7 + 3\sqrt{2}x + \frac{5}{4}x^2 + \frac{\sqrt{2}}{4}x^3 + \frac{x^4}{8} \right) e^{-\frac{x}{\sqrt{2}}} \right]. \tag{18}$$

The graphs of the approximate solutions for $\Phi(x)$ given by equations (17) and (18) compared to the numerical solution to equation (4) are plotted in Fig. 3. The absolute error of the best approximation (18) compared to the numerical solution of equation (4) is plotted in Fig. 4. From this graph, one can see that formula (18) gives a rather accurate result along with relative simplicity of its expression.

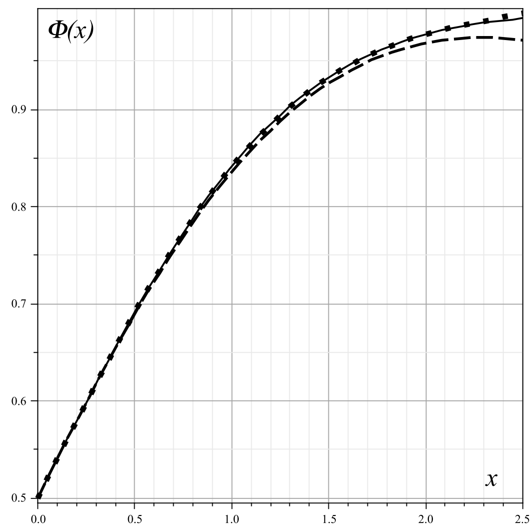

Figure 3: Comparison of the approximate solutions, given by equations (17) (dash line) and (18) (point line), with the exact numerical solution to Eq. (4) (solid line).

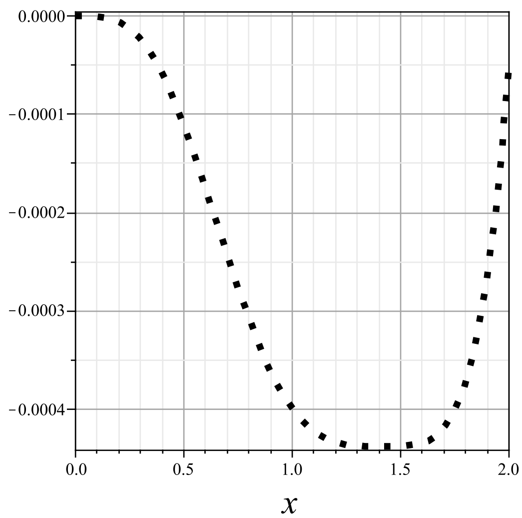

Figure 4: The absolute error of approximate solution, given by equation (18), compared to the exact numerical solution to equation (4).

3 Conclusions

This work presented an approximate analytic solution for the cumulative normal distribution (by using VIM), providing low order absolute errors. For example, the best approximation is given by a single formula (18) for $0 < x < 2$ with an absolute error of less then 4.5×10^{-4}. Our approximations to CDF are comparable to other approximations found in the literature and has the advantage of being a simple expression, that may have potential applications in several areas of applied sciences. It should be noted that the approximate CDFs obtained in this article can be used to solve various engineering and scientific problems with a fairly good approximation.

REFERENCES

[1] Aludaat K. M., Alodat M. T., A Note on Approximating the normal distribution function, *Applied Mathematical science*, 2(9): 425-429, 2008.

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

[3] Bowling S. R., Khasawneh M. T., Kaewkuekool S., and Cho B. R., A Logistic approximation to the cumulative normal distribution, *Journal of Industrial Engineering and Management*, 2(1): 114-127, 2009.

[4] Chesneau Ch. and Navarro F., On some applicable approximations of Gaussian type integrals, *Journal of Mathematical Modeling*, Vol. 7, No. 2, 2019, pp. 221-229. DOI: 10.22124/jmm.2019.12897.1250.

[5] Eidous O. M., Al-Salman S. A., One-Term Approximation for Normal Distribution Function, *Mathematics and Statistics*, 4(1): 15-18, 2016.

[6] Edous M., Eidous O., A Simple Approximation for Normal Distribution Function, *Mathematics and Statistics*, 6(4): 47-49, 2018.

[7] He J.-H., Variational iteration method - a kind of non-linear analytical technique: some examples, *International Journal of Non-Linear Mechanics*, Vol. 34, No. 4, 699-708, 1999. http://dx.doi.org/10.1016/S0020-7462(98)00048-1

[8] He J.-H., Variational iteration method for autonomous ordinary differential systems, *Applied Mathematics and Computation*, Vol. 114, No. 2-3, 115-123, 2000. http://dx.doi.org/10.1016/S0096-3003(99)00104-6

[9] He J.-H., Variational iteration method-Some recent results and new interpretations, *Journal of Computational and Applied Mathematics*, Vol. 207, No. 1, 3-17, 2007. http://dx.doi.org/10.1016/j.cam.2006.07.009

[10] Johnson N., Kotz S., and Balakrishnan N., Continuous Univariate Distributions, Vol. 1, *John Wiley* & *Sons*, New York, 1994.

[11] Lin J. T., A simple Logistic approximation to the normal tail probability and its inverse, *Applied Statistics*, 39: 255-257, 1990.

[12] Kerridge R. G. and Cook G. W., Yet another series for the normal integral, *Biometrika*, 63: 401407, 1976.

[13] Page E., Approximation to the cumulative normal function and its inverse for use on a pocket calculator, *Appl. Statist.*, 26: 7576, 1977.

[14] Polya G., Remarks on computing the probability integral in one and two dimensions, Proceeding of the first Berkeley symposium on mathematicalstatistics and probability, 63-78, 1945.

[15] Ramos J. I., On the Variational Iteration Method and Other Iterative Techniques for Nonlinear Differential Equations, *Applied Mathematics and Computation*, Vol. 199, No. 1, 39-69, 2008. http://dx.doi.org/10.1016/j.amc.2007.09.024

[16] Shchigolev V. K., Variational iteration method for studying perihelion precession and deflection of light in General Relativity, *International Journal of Physical Research*, Vol. 4, No. 2, 52-57, 2016. http://dx.doi.org/10.14419/ijpr.v4i2.6530

[17] Shchigolev V. K., Analytic Approximation of Luminosity Distance in Cosmology via Variational Iteration Method, *Universal Journal of Computational Mathematics*, 5(3): 68-74, 2017. http://dx.doi.org/10.13189/ujcmj.2017.050303

[18] Scheiber Ernest, On the Convergence of the Variational Iteration Method, arxiv:1509.01779.

[19] Tatari M. and Dehghan M., On the Convergence of Hes Variational Iteration Method, *Journal of Computational and Applied Mathematics*, Vol. 207, No. 1, 121-128, 2007. http://dx.doi.org/10.1016/j.cam.2006.07.017

[20] Vazquez-Leal H., Castaneda-Sheissa R., et al, High Accurate Simple Approximation of Normal Distribution Integral, *Mathematical Problems in Engineering*, Vol. 2012, Article ID 124029, 2012. doi:10.1155/2012/124029.

[21] Waissi G. R. and Rossin D. F., A sigmoid approximation to the standard normal integral, *Appl. Math. Comput.*, 77: 9195, 1996.

[22] Yerukala R., and Boiroju N.K., Approximations to Standard Normal Distribution Function, *International Journal of Scientific & Engineering Research*, 6(4): 515-518, 2015.