JIRSS (2008)
Vol. 7, Nos. 1-2, pp 57-72

Approximations to the Normal Distribution Function and An Extended Table for the Mean Range of the Normal Variables

M. Kiani[1], J. Panaretos[2], S. Psarakis[2], M. Saleem[3]

[1]Department of Statistics, Athens University of Economics and Business & Payame-Noor University. (mk_mehdi.kiani@yahoo.com)
[2] Department of Statistics, Athens University of Economics and Business.
[3]Department of Statistics, QUAID-i-AZAM University, Islamabad, Pakestan.

Abstract. This article presents a formula and a series for approximating the normal distribution function. Over the whole range of the normal variable z, the proposed formula has the greatest absolute error less than $6.5e-09$, and series has a very high accuracy. We examine the accuracy of our proposed formula and series for various values of z's. In the sense of accuracy, our formula and series are superior to other formulae and series available in the literature. Based on the proposed formula an extended table for the mean range of the normal variables is established.

Key words and phrases: Accuracy, error Function, normal Distribution.

1 Introduction

The normal distribution function (NDF) plays a central role in statistical theory, where,

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{\frac{-t^2}{2}} dt, \qquad -\infty < z < +\infty. \tag{1}$$

The approximations for the NDF have been presented by Zelen and Severo (1946), Abramowitz and Stegun (1964), Hart (1966), Schucany and Gray (1968), Strecock (1968), Cody (1969), Badhe (1976), Kerridge and Cook (1976), Derenzo (1977), Hamaker (1978), Parsonson (1978), Heard (1979), Moran (1980), Lew (1981), Martynov (1981), Monahan (1981), Edgeman (1988), Pugh (1989), Vedder (1993), Johnson, et al. (1994), Bagby (1995), Waissi and Rossin (1996), Bryc (2002), Marsaglia (2004), Shore (2004), Shore (2005), and several other authors.

In this paper, a new formula and a new series, for calculating the function $\Phi(z)$, are introduced. The advantages of the proposed approximations over the existing ones in literature are discussed. The new approximations to the $\Phi(z)$ are based on the error function, $erf(z)$. The integration region of the error function is $(0, z)$ for $z > 0$, or $(0, -z]$ for $z \leq 0$ that is simpler than one of the $\Phi(z)$, such that,

$$erf(z) = \int_{0}^{z} \frac{2}{\sqrt{\pi}} e^{-t^2} dt, \qquad -\infty < z < +\infty, \tag{2}$$

$$\Phi(z) = \frac{1}{2}(1 - erf(-z/\sqrt{2})), \qquad -\infty < z < +\infty. \tag{3}$$

The mean range of the random variables $Z_1, Z_2, \ldots, Z_n$ with the normal distribution, $E(R)$, for $n = 2, (1)30$ is tabulated by Montgomery (2005). The present paper sets out an extended table to $E(R)$, for $n = 2, (1)100, (20)1020$ where $E(R)$ is computed according to the proposed formula to the NDF.

2 Formulae to Approximate the NDF

As we mentioned, there are many approximations to the normal distribution function. In this section, some existing formulae will be represented. Additionally, a new formula and a comparative table, corresponding to the existing and the new approximations, are given.

2.1 Existing Formulae

An approximation to $\Phi(z)-0.5$ with absolute error less than 3×10^{-5} when $z>0$ is given by Bagby (1995),

$$\begin{aligned}\Phi(z)-0.5 \quad &\simeq \quad 0.5\{1-(1/30)[7\exp(-z^2/2)+16\exp(-z^2(2-\sqrt{2})) \\ &\quad +(7+\pi z^2/4)\exp(-z^2)]\}^{0.5}. \qquad (4)\end{aligned}$$

In this case, the approximation is obtained by using the polar integral based on $[\Phi(z)-0.5]^2$.

A sigmoid approximation is indicated by

$$\Phi(z)=1/[1+\exp(\sum_{k=0}^{\infty} a_k z^{2k+1})].$$

Based on this approximation, a simple formula with maximum absolute error 4.31×10^{-5} for $z\in[-8,+8]$ was introduced by Waissi and Rossin (1996),

$$\Phi(z)\simeq 1/(1+\exp[-\sqrt{\pi}(\beta_1 z^5+\beta_2 z^3+\beta_3 z)]), \text{ for } z\in[-8,+8]. \qquad (5)$$

where $\beta_1=-0.0004406$, $\beta_2=0.0418198$, $\beta_3=0.9000000$. Bryc (2002) presented a formula with maximum absolute error 1.9×10^{-5}, according to rational approximations to Mill's ratio,

$$1-\Phi(z)\simeq \frac{z^2+5.575192695z+12.77436324}{\sqrt{2\pi}z^3+14.38718147z^2+31.53531977z+2\times 12.77436324}e^{-z^2/2} \qquad (6)$$

This formula gives at least two significant digits precision for all $z>0$.

Using the response modeling methodology, an approximation for the NDF, having greatest absolute error 2×10^{-6} was suggested by Shore (2004). Later, Shore (2005) improved his proposed formula in Shore (2004) to the following formula with a maximum absolute error 6×10^{-7} , such that,

$$\Phi(z)\simeq[1+g(-z)-g(-z)]/2, \quad \text{for} \quad -9<z<9. \qquad (7)$$

In this case,

$$g(z)=\exp\{-\log(2)\exp\{[\alpha/(\lambda/S_1)][(1+S_1z)^{(\lambda/S_1)}-1]+S_2z\}\}$$

where $\lambda=-0.61228883$; $S_1=-0.11105481$; $S_2=0.44334159$; $\alpha=-6.37309208$.

2.2 New Formula

As noted in the preceding section, the integration region to the function erf is simpler than one of the NDF. Hence, the construction of the proposed formula is based on the error function. Replacing z by $-z/\sqrt{2}$ in (2),

$$erf(-z/\sqrt{2}) = \frac{2}{\sqrt{\pi}} \int_0^{-z/\sqrt{2}} e^{-t^2} dt, \quad \text{for} \quad -\infty < z < +\infty. \tag{8}$$

Let,

$$I = \int_0^{-z/\sqrt{2}} e^{-t^2} dt, \qquad \text{for} \quad -\infty < z < +\infty. \tag{9}$$

$$I^2 = \int_0^{-z/\sqrt{2}} \int_0^{-z/\sqrt{2}} e^{(-t_1^2+t_2^2)} dt_1 dt_2, \quad \text{for} \quad -\infty < z < +\infty. \tag{10}$$

The integrand of polar integral is less variable than the original one; thus, using the definition of trigonometric functions $t_1 = r\cos(\beta)$ and $t_2 = r\sin(\beta)$, for $z \le 0$ equation (10) is transformed to

$$\begin{aligned} I^2 &= \int_0^{\pi/4} \int_0^{-z/(\cos(\beta)\sqrt{2})} re^{-r^2} dr d\beta \\ &\quad + \int_{\pi/4}^{\pi/2} \int_0^{-z/(\sin(\beta)\sqrt{2})} re^{-r^2} dr d\beta \\ &= \int_0^{\pi/4} (1 - e^{-(z/(\cos(\beta)\sqrt{2}))^2}) d\beta. \end{aligned}$$

Denote,

$$\omega(\beta) = (1/(\cos(\beta)\sqrt{2}))^2). \tag{11}$$

Transforming $\omega(\beta)$ in polar coordinates to $\omega(z)$ in rectangular coordinates, we get,

$$\omega(z) = \ln\left(1 - \frac{4}{\pi}(\int_0^{-z/\sqrt{2}} e^{-t^2} dt)^2\right) / - z^2. \tag{12}$$

Therefore $I^2 = \int_0^{\pi/4}(1 - e^{-(z)^2\omega(z)})d\beta$, and,

$$I = \sqrt{\frac{\pi}{4}(1 - e^{-(z)^2\omega(z)})}, \quad \text{for} \quad z \le 0. \tag{13}$$

In the sequel, combining (8), (9), and (13), for $z \le 0$, we have

$$erf(-z/\sqrt{2}) = 2I/\sqrt{\pi}, \quad \text{for} \quad z \le 0. \tag{14}$$

Now, equation (10) is evaluated under the assumption that $z > 0$. When $0 > t_1 \geq t_2 \geq -z/\sqrt{2}$ then $\pi \leq \beta \leq 5\pi/4$ and $0 < r \leq -z/(cos(\beta)\sqrt{2})$, whereas, when $0 > t_1 \geq t_2 \geq -z/\sqrt{2}$ then $5\pi/4 < \beta \leq 3\pi/2$ and $0 < r \leq -z/(sin(\beta)\sqrt{2})$. As a result, the calculations on I^2, on behalf of $z > 0$, yield and $I = -[\pi(-e^{-(z)^2\omega(z)} + 1)/4]^{1/2}$

$$erf(-z/\sqrt{2}) = -2I/\sqrt{\pi}, \quad \text{for} \quad z > 0. \tag{15}$$

Define,

$$sign(-z) = \begin{cases} -1 & \text{if} \quad z > 0 \\ 1 & \text{if} \quad z \leq 0 \end{cases}$$

Because of (14) and (15), the $erf(-z/\sqrt{2})$ over the whole range of z will be

$$erf(-z/\sqrt{2}) = sign(-z)\sqrt{1 - e^{-z^2\omega(z)}}, \quad \text{for} \quad -\infty < z < +\infty. \tag{16}$$

Combining equations (3) and (16) we will have,

$$\Phi(z) = \frac{1}{2}(1 - sign(-z)\sqrt{1 - e^{-z^2\omega(z)}}), \quad \text{for} \quad -\infty < z < +\infty \tag{17}$$

The function $\omega(z)$ is approximated by $\omega_A(z)$, such that,

$$\omega_A(z) = \tag{18}$$
$$\begin{cases} -6.62e-6|z|^5 + 4.4166e-4z^4 - 1.31e-5|z|^3 \\ \quad - 9.56/17e-3z^2 - 4.8e-7|z| + /.636619771 & 0 \leq |z| < 1.05 \\ -1.401663e-4|z|^5 + 1.150811e-3z^4 - 1.582555e \\ \quad -3|z|^3 - 7.76126e-3z^2 - 1.0608e-3|z| + 0.6368751 & 1.05 \leq |z| < 2.29 \\ 5.8716e-5|z|^5 - 1.221684e-3z^4 + 9.841663e-3|z|^3 \\ \quad - 3.55101e-2z^2 + 3.29203e-2|z| + 0.62010268 & 2.29 \leq |z| < 8 \\ 0.5 & 8 \leq |z|. \end{cases}$$

Substituting the $\omega_A(z)$ in equations (16) and (17), then the approximations $erf_A(-z/\sqrt{2})$ and $\Phi_A(z)$ is derived. Equivalently, for each z, we have $erf(z) \simeq erf_A(z) = 1 - 2\Phi_A(-z/\sqrt{2})$. It is highly appropriate, for $|z| \geq 5.5$, the functions $erf_A(z)$ and $\Phi_A(z)$ to be constructed by applying rational chebyshev approximations. In this case, the rational function of degree l in the numerator and m in the denominator is approximately defined by $R_{lm}(1/z^2) \simeq 0.5641882$. As a result,

$$\Phi_A(z) \quad = \quad \frac{1}{2}(1 - sign(-z)\sqrt{1 - e{-}z^2\omega_A(z)}), \quad \text{for} \quad |z| \leq -5.5,$$

$$\Phi_A(z) = \frac{e^(-z^2/2)\sqrt{2}}{2}\{\frac{0.5641882}{z^3} - \frac{1}{z\sqrt{\pi}}\} \text{ for } z \leq -5.5, \quad (19)$$
$$\Phi_A(z) = 1 - \frac{e^(-z^2/2)\sqrt{2}}{2}\{\frac{1}{z\sqrt{\pi}} - \frac{0.5641882}{z^3}\} \text{ for } z \geq -5.5$$

Numerical experiments have shown that, for all z, the greatest absolute error to $\Phi_A(z)$ and $erf_A(z)$ is less than 6.5×10^{-9} and 1.6×10^{-8}, respectively.

Table 1 is established to compare the performance of the studied and the proposed formulae.

Table 1. Absolute error of the formulae to approximate the NDF.

Formula	Z=-30	Z=-10	Z=-6.5	Z=-5.5	Z=-4.5
(4)	4.9E-198	7.6E-24	1.6E-13	5.5E-10	1.2E-07
(5)	1.0E+00	6.3E-06	3.5E-10	1.6E-08	3.6E-07
(6)	1.8E-200	3.6E-26	1.6E-13	6.7E-11	9.6E-09
(7)	i	i	3.5E-11	8.3E-09	2.7E-07
(19)	1.8E-203	2.2E-27	6.2E-14	5.5E-11	1.5E-09
Formula	Z=-3.5	Z=-2.5	Z=-1.5	Z=-0.5	Z=0
(4)	2.3E-06	1.1E-05	1.9E-05	2.8E-05	0.0E+00
(5)	3.4E-06	3.4E-05	1.6E-05	2.6E-05	0.0E+00
(6)	4.6E-07	6.5E-06	1.9E-05	1.6E-06	0.0E+00
(7)	5.2E-07	3.1E-07	7.6E-08	5.7E-08	0.0E+00
(19)	3.3E-10	1.1E-09	3.0E-09	2.1E-09	0.0E+00

"i" indicates a complex number.

This table shows the absolute errors involved in the formulae, such that approximation (19) has minimum absolute error over the wide range of z. The formulae (5) and (7) fail to approximate the NDF for absolute amounts of large z's.

3 Series to Approximate the NDF

In the sequel, series expansions to approximate the NDF will be represented. In addition, a new series with very high accuracy is given.

3.1 Existing Series

Corresponding to the rational chebyshev approximation, Cody (1969) introduced the approximations for the error function and the complementary error function, $erfc(x)$, where,

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

The presented approximations are,

$$\begin{aligned} erf(x) &\simeq xR_{lm}(x^2) = x \sum_{j=0}^{n} p_j x^{2j} / \sum_{j=0}^{n} q_j x^{2j}, \quad |x| \leq 0.5, \\ erfc(x) &\simeq e^{-x^2} R_{lm}(x^2) \\ &= e^{-x^2} \sum_{j=0}^{n} p_j x^j / \sum_{j=0}^{n} q_j x^j, \quad 0.46875 \leq x \leq 4.0 \qquad (20) \\ erfc(x) &\simeq \frac{e^{-x^2}}{x} \{ \frac{1}{\sqrt{\pi}} + \frac{1}{x^2} R_{lm}(1/x^2) \} \\ &= \frac{e^{-x^2}}{x} \{ \frac{1}{\sqrt{\pi}} + \frac{1}{x^2} \sum_{j=0}^{n} p_j x^{-2j} / \sum_{j=0}^{n} q_j x^{-2j} \}, \quad x \geq 4.0. \end{aligned}$$

where, the coefficients p_j and q_j are tabulated for various value n in the paper of Cody (1969). The maximum relative errors, for these approximations, ranging down to between 6×10^{-19} and 6×10^{-20} for all z.

Kerridge and Cook (1976) present a convergent Taylor expansion for computing $\Phi_0(z)$, where $\Phi_0(z) = \Phi(z) - 0.5$ and,

$$\begin{aligned} \Phi_0(z) &= \int_0^z \frac{1}{\sqrt{2\pi}} e^{frac-t^2 2} dt \\ &\simeq \frac{1}{\sqrt{2\pi}} z e^{\frac{-z^2}{8}} \sum_{n=0}^{+\infty} \frac{1}{2n+1} \theta_{2n}(z/2), \quad -\infty < z < +\infty. (21) \end{aligned}$$

In this series, $\theta_n(z) = z^n H_n(z)/n!$, for $n = 0, 1, 2, \ldots$, and $H_n(z)$ implies the nth Hermite polynomial, such that $H_0(z) = 1$, $H_1(z) = z$, and $H_{n+1}(z) = zH_n(z) - nH_{n-1}(z)$ for $n = 1, 2, \ldots$. They suggest some advantages for using $\theta_n(z)$ over $H_n(z)$, such that $\theta_n(z)$ are easier to handle numerically and relatively small for large n,

$$\theta_0(z) = 1; \ \theta_1(z) = z^2; \ \theta_{n+1}(z) = \frac{z^2[\theta_n(z) - \theta_{n-1}(z)]}{n+1}, \text{ for } n = 1, 2, \ldots$$

Recently, Marsaglia (2004) provided the approximation bellow,

$$\Phi(z) \simeq 0.5 + (2\pi)^{-1/2} e^{-z^2/2} \left(z + \frac{z^3}{3} + \frac{z^5}{3.5} + \frac{z^7}{3.5.7} + \frac{z^9}{3.5.7.9} + \cdots \right). \qquad (22)$$

Marsaglia's series based on the Taylor expansion about zero for function $B(z)$,

$$B(z) = \int_0^z e^{-t^2/2} dt e^{-z^2/2} \simeq z + \frac{z^3}{3} + \frac{z^5}{3.5} + \frac{z^7}{3.5.7} + \cdots$$

He provided the following C function, using C compiler libraries, for the computation of $\Phi(z)$,

```
double Phi(double z)
{Long double=z,t=0,b=z,q=z*z,i=1;
while(s!=t)s=(t=s)+(b*=q/(i+=2));
return 0.5+s*exp(-0.5*q-0.91893853320467274178L);}.
```

The accuracy of the proposed series by Kerridge and Cook (1976) and Marsaglia (2004) will be discussed, where the accuracy of these series rely on the terms used of series and the digits predefined for computing the approximations.

3.2 New Series

The function $\Phi(z)$ can be numerically approximated, utilizing the Taylor expansion to e^{-t^2},

$$e^{-t^2} \simeq e^{-c^2} \sum_{k=0}^{+\infty} \frac{u_k(c)}{k!}(t-c)^k, \quad \text{for} \quad -\infty < t < +\infty \tag{23}$$

where, $c = t + \alpha$ for $0 < \alpha \leq 1$, and,

$$u_0(c) = 1; \; u_1(c) = -2c;$$
$$u_k(c) = -2(k-1)u_{k-2}(c) + (-2c)u_{k-1}(c), \quad \text{for} \quad k \geq 2.$$

Integrating on (23), with respect to t from 0 to $-z/\sqrt{2}$,

$$\begin{aligned} I_{erf} &= \int_0^{-z/\sqrt{2}} e^{-t^2} dt \simeq I_{Aerf} \\ &= sign(-z) \sum_{i=1}^{B+1} e^{-c_i^2} \sum_{k+1}^{n\to+\infty} \frac{u_{(i,k)}(c_i)}{k!}(A_{i+1} - c_i)^k, \end{aligned} \tag{24}$$

where, $B = round(|z|/\sqrt{2})$, $c_i = A_i = (i-1) \times round(|z|/B\sqrt{2}))$, $A_{B+2} = |z|/\sqrt{2}$, $u_{(i,1)} = 1$, $u_{(i,2)} = -2c_i$, $u_{(i,k)} = -2[(k-2)u_{(i,k-2)} + c_i u_{(i,k-1)}]$, for $i = 1, 2, \ldots, B+1$ and $k \geq 3$. Excepti-onally, B is equated to 1, when $B = round(|z|/\sqrt{2})$ is equal to zero.

To avoid the calculation of equation (24) for large value B, integrating on (23) with respect to t from $-z/\sqrt{2}$ to $\pm\infty$, let us define

$$\begin{aligned} I_{erfe} &= \int_{-z/\sqrt{2}}^{\infty} e^{-t^2} dt \simeq I_{Aerf} \\ &= sign(-z) \sum_{i=1}^{m\to+\infty} e^{-c_i^2} \sum_{k+1}^{n\to+\infty} \frac{u_{(i,k)}(c_i)}{k!}(A_i - c_i)^k, \end{aligned} \tag{25}$$

In this case, $A_i = |z|/\sqrt{2}$, $A_2 = round(|z|/\sqrt{2}) + 1$, $A_{i+1} = A_i + 1$ for $i \geq 3$, and $c_i = A_{i+1}$ for $i \geq 1$. Applying (24) and (25), the error function and the complementary error function are approximated by

$$erf(-z/\sqrt{2}) \simeq 2I_{Aref}/\sqrt{\pi}, \tag{26}$$

$$erf(-z/\sqrt{2}) \simeq 2I_{Arefe}/\sqrt{\pi}, \tag{27}$$

As a consequence, corresponding to $\Phi(z) = 0.5 \times erfc(-z/\sqrt{2})$ for $z \leq 0$ and $\Phi(z) = 1 + 0.5 \times erfc(-z/\sqrt{2})$ for $z > 0$, we will have the following approximation with very high precision,

$$\begin{aligned} \Phi(z) &\simeq \frac{1}{2}\frac{\{\sqrt{\pi} - 2I_{Aerf}\}}{\sqrt{\pi}} \quad \text{for} \quad |z| \leq 4, \\ \Phi(z) &\simeq I_{Aerfe}/\sqrt{\pi} \quad \text{for} \quad z \leq -4, \\ \Phi(z) &\simeq 1 + I_{Aerfe}/\sqrt{\pi} \quad \text{for} \quad z \leq -4, \end{aligned} \tag{28}$$

On behalf of $|z| \leq 4$, this approximation is accurate with at least 60 significant digits accuracy, when $n \geq 100$ in (24). The digits for computing (28) is held equal or greater than 65 to achieve at least 60 significant digits accuracy for $0 \leq |z| \leq 70$. Furthermore, the approximation (28) relies on the value m in series (25). Numerical experiments show, when $m \geq 10$ for $4 \leq |z| \leq 45$ and $m \geq 2$ for $45 \leq |z| \leq 70$, the approximation (28) gives the desired accuracy, in at least 60 significant digits.

Generally, in practice, for $0 \leq |z| \leq 70$, the approximation (28) has in at least 60 significant digits accuracy, whereas in theory arbitrary accuracy can be achieved for all z. For example, according to the approximations (21), (22) and (28), applying Maple or C compiler libraries, we have,

$$\begin{aligned} \Phi(-70) = {} & 0.5423039609301399328675786670 8 \\ & 7759716518976172187170282890450e - 1066. \end{aligned}$$

This means (28) is accurate in at least 60 significant digits. The calculations of $\Phi(-70)$ are based on expansion truncated at 3309, about 13600, and 252 terms and Digits equated to 1127, about 3000, and 64 for series (21), (22), and (28), respectively.

The small terms, m and n, and small digits for computing are good properties for (28), such that the speed of calculation is too fast for either small or large z, $0 \leq |z| \leq 70$. For very large value z i.e., $|z| > 70$, if the significant digits accuracy is only important and the speed is not, then the approximation (28) is proposed, such that we hold $m = 2$, Digits:=70 and only increase n. Otherwise, the approximation (20) is proposed, where the speed of the calculation for this approximation is very fast and its accuracy is between 18 and 20 significant digits. Corresponding to numerical experiments, the number of terms, n, for a specific level of accuracy is almost a linear function of z. For example, for $z = -600$ the approximation (28) is truncated at $n = 1080$ terms, where

$$\begin{aligned}\Phi(-600) &= 0.654658820580769285210592771388 8\\ &\quad 1087821194128318531772111 6943e-78176.\end{aligned}$$

We expect this approximation to be accurate, since larger terms to compute (28), $n > 1080$, gives the same approximation for $\Phi(-600)$, possessing 60 significant digits. Series (20) and formulae (6) and (19) approximate $\Phi(-600)$ having 20, 3, and 10 significant digits accuracy, respectively. It is not easy to calculate $\Phi(-600)$ according to series (21) and (22), because of the convergence of these series suffer from difficulties with very large terms required.

Table 2 is constructed to compare the performance of the studied and the proposed series. The digits to the calculation for this table is defined Digits:=200, for all the approximations, and expansion of series (21), (22) and (28) are truncated at 160 terms. The exact values in comparative table is based on the series (21), (22) and (28) with large terms and digits, such that these series give the same approximation for $\Phi(z)$ having at least 30 significant digits accuracy.

Table 2. Series to approximate the NDF, (160 terms and Digits:=200).

Series	Z=-18
Exact	0.974094891893715048259189518997e-72
(20)	0.974094891893715048708181934747e-72
(21)	0.226820907630354110107306715331e-38
(22)	0.499999999999757093038012396287e-00
(28)	0.974094891893715048259189518997e-72
Series	Z=-9
Exact	0.112858840595384064773550207597e-18
(20)	0.112858840595384064738093247631e-18
(21)	0.112858840595384064773550207597e-18
(22)	0.989596251047682032099597869127e-08
(28)	0.112858840595384064773550207597e-18

Table 2 (continued). Series to approximate the NDF, (160 terms and Digits:=200).

Series	Z=-3
Exact	0.134989803163009452665181476759e-02
(20)	0.134989803163009452631102368374e-02
(21)	0.134989803163009452665181476759e-02
(22)	0.134989803163009452665181476759e-02
(28)	0.134989803163009452665181476759e-02
Series	Z=-1
Exact	0.158655253931457051414767454368
(20)	0.158655253931457051377370713583
(21)	0.158655253931457051414767454368
(22)	0.158655253931457051414767454368
(28)	0.158655253931457051414767454368

Table 2 shows, under the conditions as mentioned, the series (21) and (22) are accurate, for small z's, and series (20) is accurate with 18 to 20 significant digits for wide range of z. Furthermore, this table shows series (28) is accurate in at least 30 significant digits for either small or large z's.

As a result, the existing series have serious disadvantages. Series proposed by Kerridge and Cook (1976) and Marsaglia (2004) are based on the Taylor expansion about zero, in particularly, these approximations are the Maclaurin series. Hence, in practice, these series fail to approximate $\Phi(z)$ for large z's. Series offered by Cody (1969) relies on the fixed amounts of coefficients p_j and q_j. Therefore, the accuracy accomplished of this approximation is constrained on 18 to 21 significant digits. Therefore the proposed series seems to be superior in at least these aspects. Statistical softwares (for example Matlab, S-plus and MS Excel) compute the $\Phi(z)$, for example $\Phi(-1)$, with different significant digits. To overcome this problem, the new series is proposed to be used on the statistical packages, because of its advantages.

4 The Mean Range for Normal Distribution

To approximate the mean range of the normal variables, corresponding to equation (1), define the random variable Z_i's, for $i = 1, 2, \ldots, n$. Define, also, the range of order statistics $Z_{(1)}, Z_{(2)}, \ldots, Z_{(n)}$ by $R = Z_{(n)} - Z_{(1)}$, and its mean by d_2 i.e. $E(R) = d_2$. Under these conditions, the probability density function to R is

$$\varphi_R(r) = \int_{-\infty}^{+\infty} n(n-1)[\Phi(r+z) - \Phi(z)]^{n-2}\varphi(z)\varphi(r+z)dz, \quad r \geq 0.$$

where, the $\varphi(z)$ denotes the normal density function. The evaluation to the mean range of the random variables with normal distribution is given by Johnson, et al. (1994), where,

$$E(R) = \int_{-\infty}^{+\infty} \{1 - (F(z))^n - (1 - F(z))^n\}dz.$$

To construct an extended table for d_2 the $\Phi(z)$ is evaluated by the proposed formula (19) with maximum absolute error 6.5e-09. However, although, the considered series are much more accurate than the formulae, but it is no possible or at least very difficult to evaluate the $E(R)$, by using these series. Table 3 exhibits the mean range of the normal variables Z_i for various values $n = 2(1)100$, 120(20)1020. Application of $E(R)$ is given by Montgomery (2005) to establish the mean control charts for monitoring a quality characteristic.

Table 3. The mean range of normal distribution (d_2).

n	d_2	n	d_2	n	d_2	n	d_2	n	d_2
2	1.12838	31	4.11293	60	4.63856	89	4.93131	460	6.02251
3	1.69257	32	4.13934	61	4.65112	90	4.93940	480	6.04853
4	2.05875	33	4.16482	62	4.66346	91	4.94739	500	6.07340
5	2.32593	34	4.18943	63	4.67557	92	4.95529	520	6.09721
6	2.53441	35	4.21322	64	4.68747	93	4.96309	540	6.12004
7	2.70436	36	4.23625	65	4.69916	94	4.97079	560	6.14198
8	2.84720	37	4.25855	66	4.71065	95	4.97841	580	6.16308
9	2.97003	38	4.28018	67	4.72194	96	4.98593	600	6.18340
10	3.07751	39	4.30117	68	4.73305	97	4.99337	620	6.20301
11	3.17287	40	4.32155	69	4.74397	98	5.00073	640	6.22194
12	3.25846	41	4.34136	70	4.75472	99	5.00800	660	6.24023
13	3.33598	42	4.36063	71	4.76529	100	5.01519	680	6.25794
14	3.40676	43	4.37938	72	4.77570	120	5.14417	700	6.27509
15	3.47183	44	4.39764	73	4.78595	140	5.25118	720	6.29172
16	3.53198	45	4.41544	74	4.79604	160	5.34243	740	6.30786
17	3.58788	46	4.43279	75	4.80598	180	5.42186	760	6.32353
18	3.64006	47	4.44972	76	4.81578	200	5.49208	780	6.33876
19	3.68896	48	4.46624	77	4.82543	220	5.55497	800	6.35358
20	3.73495	49	4.48238	78	4.83493	240	5.61185	820	6.36800
21	3.77834	50	4.49815	79	4.84431	260	5.66375	840	6.38205
22	3.81938	51	4.51356	80	4.85355	280	5.71144	860	6.39573
23	3.85832	52	4.52864	81	4.86266	300	5.75553	880	6.40908
24	3.89535	53	4.54339	82	4.87165	320	5.79652	900	6.42211
25	3.93063	54	4.55783	83	4.88051	340	5.83480	920	6.43483
26	3.96432	55	4.57197	84	4.88926	360	5.87069	940	6.44725
27	3.99654	56	4.58582	85	4.89789	380	5.90446	960	6.45939
28	4.02741	57	4.59939	86	4.90641	400	5.93636	980	6.47126
29	4.05704	58	4.61270	87	4.91481	420	5.96655	1000	6.48287
30	4.08552	59	4.62575	88	4.92311	440	5.99522	1020	6.49423

5 Conclusion

New methods for the approximation of the normal distribution function have been introduced. The accuracy and the speed of the calculations are advantages of the proposed methods over the some existing methods. An extended table for the mean range of the normal variables has been constructed.

References

Abramowitz, M. and Stegun, I. A. (1964), Handbook of Mathematical Functions. New York: Dover Publications, Inc.

Badhe, S. K. (1976), New approximation of the normal distribution function. Communications in Statistics-Simulation and Computation, **5**, 173-176.

Bagby, R. J. (1995), Calculating normal probabilities. The American mathematical monthly. **102**, 46-49.

Bryc, W. (2002), A uniform approximation to the right normal tail integral. Applied Mathematics and Computation, **127**, 365-374.

Cody, W. J. (1969), Rational Chebyshev approximations for the error function. Mathematics of Computation, **23**(107), 631-637.

Derenzo, S. E. (1977), Approximations for hand calculators using small integral coefficients. Mathematics of Computation, **31**, 214-225.

Divgi, D. R. (1979), Calculation of univariate and bivariate normal probability function. The Annals of Statistics, **7**(4), 903-910.

Edgeman, R. L. (1988), Normal distribution probabilities and quantiles without tables. Mathematics of Computer Education, **22**, 95-99.

Hamaker, H. (1978), Approximating the cumulative normal distribution and its inverse. Applied Statistics, **27**, 76-77.

Hart, R. G. (1966), A close approximation related to the error function. Mathematics of Computation, **20**, and 600-602.

Heard, T. J. (1979), Approximation to the normal distribution function. Mathematical Gazette, **63**, 39-40.

Johnson, N. l., Kotz, S., and Balakrishnan, N. (1994), Continuous Univariate Distributions. 2nd edition. New York: Wiley.

Kerridge, D. F. and Cook, G. W. (1976), Yet another series for the normal integral. Biometrika, **63**, 401-403.

Lew, R. A. (1981), An approximation to the cumulative normal distribution with simple coefficients. Applied Statistics, **30**(3), 299-301.

Marsaglia, G. (2004), Evaluating the normal distribution. Journal of Statistical Software, **11**(4), 1-11.

Martynov, G. V. (1981), Evaluation of the normal distribution function. Journal of Soviet Mathematics, **17**, 1857-1876.

Monahan, J. F. (1981), Approximation the log of the normal cumulative. Computer Science and Statistic: Proceeding of the Thirteenth Symposium on the Interface W. F. Eddy (editor), 304-307, New York: Springer-Verlag.

Moran, P. A. P. (1980), Calculation of the normal distribution function. Biometrika, **67**, 675-676.

Montgomery, D. C. (2005), Introduction to Statistical Quality Control. 5th edition. New York: John Wiley.

Parsonson, S. L. (1978), An approximation to the normal distribution function. Mathematical Gazette, **62**, 118-121.

Pugh, G. A. (1989), Computing with Gaussian Distribution: A Survey of Algorithms. ASQC Technical Conference Transactions, **46**, 496-501.

Schucany, W. R. and Gray, H. L. (1968), A new approximation related to the error function. Mathematics of Computation, **22**, 201-202.

Shore, H. (2004), Response modeling methodology (RMM)-current statistical distributions, transformations and approximations as special cases of RMM. Communications in Statistics-Theory and Methods, **33**(7), 1491-1510.

Shore, H. (2005), Accurate RMM-based approximations for the CDF of the normal distribution. Communications in Statistics-Theory and Methods, **34**, 507-513.

Strecock, A. J. (1968), On the calculation of the inverse of the error function. Mathematics of Computation, **22**, 144-58.

Vedder, J. D. (1993), An invertible approximation to the normal-distribution function. Computational Statistics & Data Analysis, **16**, 119-123.

Waissi, G. and Rossin, D. F. (1996), A Sigmoid approximation to the standard normal integral. Applied Mathematics and Computation, **77**, 91-95.

Zelen, M. and Severo, N. C. (1964), Probability Function. Handbook of Mathematical Functions, M. Abramowitz and I. A. Stegun (editors), 925-995, U. S. Department of Commerce, Applied Mathematics Series, 55.