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

M. Kiani\textsuperscript{1}, J. Panaretos\textsuperscript{2}, S. Psarakis\textsuperscript{2}, M. Saleem\textsuperscript{3}

\textsuperscript{1}Department of Statistics, Athens University of Economics and Business & Payame-Noor University. (mk.mehdi.kiani@yahoo.com)
\textsuperscript{2} Department of Statistics, Athens University of Economics and Business.
\textsuperscript{3} Department of Statistics, QUAIID-i-AZAM University, Islamabad, Pakistan.

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,

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


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, \( \text{erf}(z) \). 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,

\[
\text{erf}(z) = \int_{0}^{z} \frac{2}{\sqrt{\pi}} e^{-t^2} dt, \quad -\infty < z < +\infty,
\] (2)

\[
\Phi(z) = \frac{1}{2} (1 - \text{erf}(-z/\sqrt{2})), \quad -\infty < z < +\infty.
\] (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.
Approximations to the Normal Distribution Function and ...

2.1 Existing Formulae

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

$$
\Phi(z) - 0.5 \approx 0.5 \{ 1 - (1/30)[7 \exp(-z^2/2) + 16 \exp(-z^2(2 - \sqrt{2})) + (7 + \pi z^2/4) \exp(-z^2)] \}^{0.5}.
$$

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

$$
\Phi(z) = \frac{1}{1 + \exp\left[ -\sqrt{\pi} \left( \beta_1 z^5 + \beta_2 z^3 + \beta_3 z \right) \right]}, \quad \text{for } z \in [-8, +8].
$$

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

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

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 $2 \times 10^{-6}$ was suggested by Shore (2004). Later, Shore (2005) improved his proposed formula in Shore (2004) to the following formula with a maximum absolute error $6 \times 10^{-7}$, such that,

$$
\Phi(z) \approx \frac{1 + g(-z) - g(-z)}{2}, \quad \text{for } -9 < z < 9.
$$

In this case,

$$
g(z) = \exp\{-\log(2) \exp\{\alpha/(\lambda/S_1)\}[(1 + S_1 z)^{(\lambda/S_1)} - 1] + S_2 z\}.
$$

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

www.SID.ir
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),

$$erf(-z/\sqrt{2}) = \frac{2}{\sqrt{\pi}} \int_{0}^{-z/\sqrt{2}} e^{-t^2} dt, \quad \text{for } -\infty < z < +\infty. \quad (8)$$

Let,

$$I = \int_{0}^{-z/\sqrt{2}} e^{-t^2} dt, \quad \text{for } -\infty < z < +\infty. \quad (9)$$

$$I^2 = \int_{0}^{-z/\sqrt{2}} \int_{0}^{-z/\sqrt{2}} e^{-(t_1^2 + t_2^2)} dt_1 dt_2, \quad \text{for } -\infty < z < +\infty. \quad (10)$$

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 $z \leq 0$ equation (10) is transformed to

$$I^2 = \int_{\pi/4}^{\pi/2} \int_{0}^{-z/(\cos(\beta)\sqrt{2})} re^{-r^2} dr d\beta + \int_{\pi/2}^{\pi/4} \int_{0}^{-z/(\sin(\beta)\sqrt{2})} re^{-r^2} dr d\beta$$

$$= \int_{\pi/4}^{\pi/2} (1 - e^{-(z/(\cos(\beta)\sqrt{2}))^2}) d\beta.$$  

Denote,

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

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

$$\omega(z) = \ln \left(1 - \frac{4}{\pi} \int_{0}^{-z/\sqrt{2}} e^{-t^2} dt \right) / z^2. \quad (12)$$

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

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

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

$$erf(-z/\sqrt{2}) = 2I/\sqrt{\pi}, \quad \text{for } z \leq 0. \quad (14)$$
Approximations to the Normal Distribution Function and ...

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 \( I = -[\pi(e^{-z^2\omega(z)} + 1)/4]^{1/2} \)

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

Define,

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

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

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

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

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

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

\[ \omega_A(z) = \begin{cases} -6.62e - 6|z|^5 + 4.4166e - 4z^4 - 1.31e - 5|z|^3 \\ \quad - 9.56/17e - 3z^2 - 4.8e - 7|z| + 6.36619771 & 0 \leq |z| < 1.05 \\ -1.401663e - 4|z|^5 + 1.150811e - 3z^4 - 1.582555e \quad - 3|z|^3 - 7.76126e - 3z^2 - 1.0608e - 3|z| + 6.368751 & 1.05 \leq |z| < 2.29 \\ 5.8716e - 5|z|^5 - 1.221684e - 3z^4 + 9.841663e - 3|z|^3 \quad - 3.5510e - 2z^2 + 3.29203e - 2|z| + 6.2010268 & 2.29 \leq |z| < 8 \\ 0.5 & 8 \leq |z|. \end{cases} \] (18)

Substituting the \( \omega_A(z) \) in equations (16) and (17), then the approximations \( \text{erf}_A(-z/\sqrt{2}) \) and \( \Phi_A(z) \) is derived. Equivalently, for each \( z \), we have \( \text{erf}(z) \simeq \text{erf}_A(z) = 1 - 2\Phi_A(-z/\sqrt{2}) \). It is highly appropriate, for \( |z| \geq 5.5 \), the functions \( \text{erf}_A(z) \) 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,

\[ \Phi_A(z) = \frac{1}{2}(1 - \text{sign}(-z)\sqrt{1 - e^{-z^2\omega_A(z)})}, \quad \text{for} \quad |z| \leq -5.5, \]
\[ \Phi_A(z) = \frac{e^{(-z^2/2)}\sqrt{2}}{2} \left\{ \frac{0.5641882}{z^3} - \frac{1}{z\sqrt{\pi}} \right\} \] for \( z \leq -5.5 \), \( (19) \)

\[ \Phi_A(z) = 1 - \frac{e^{(-z^2/2)}\sqrt{2}}{2} \left\{ \frac{1}{z\sqrt{\pi}} - \frac{0.5641882}{z^3} \right\} \] for \( z \geq -5.5 \)

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

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

<table>
<thead>
<tr>
<th>Formula</th>
<th>( Z=-30 )</th>
<th>( Z=-10 )</th>
<th>( Z=-6.5 )</th>
<th>( Z=-5.5 )</th>
<th>( Z=-4.5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>(4)</td>
<td>4.9E-198</td>
<td>7.6E-24</td>
<td>1.6E-13</td>
<td>5.5E-10</td>
<td>1.2E-07</td>
</tr>
<tr>
<td>(5)</td>
<td>1.0E+00</td>
<td>6.3E-06</td>
<td>3.5E-10</td>
<td>1.6E-08</td>
<td>3.6E-07</td>
</tr>
<tr>
<td>(6)</td>
<td>1.8E-200</td>
<td>3.6E-26</td>
<td>1.6E-13</td>
<td>6.7E-11</td>
<td>9.6E-09</td>
</tr>
<tr>
<td>(7)</td>
<td>i</td>
<td>i</td>
<td>3.5E-11</td>
<td>8.3E-09</td>
<td>2.7E-07</td>
</tr>
<tr>
<td>(19)</td>
<td>1.8E-203</td>
<td>2.2E-27</td>
<td>6.2E-14</td>
<td>5.5E-11</td>
<td>1.5E-09</td>
</tr>
</tbody>
</table>

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 com-
Approximations to the Normal Distribution Function and ...  

plemimentary error function, \( \text{erf}(x) \), where,  

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

The presented approximations are,  

\[
\begin{align*}
\text{erf}(x) & \approx x R_{lm}(x^2) = x \sum_{j=0}^{n} p_j x^{2j} / \sum_{j=0}^{n} q_j x^{2j}, \quad |x| \leq 0.5, \\
\text{erfc}(x) & \approx e^{-x^2} R_{lm}(x^2) = e^{-x^2} \sum_{j=0}^{n} p_j x^{2j} / \sum_{j=0}^{n} q_j x^{2j}, \quad 0.46875 \leq x \leq 4.0 \quad (20) \\
\text{erfc}(x) & \approx \frac{e^{-x^2}}{x} \left( \frac{1}{\sqrt{\pi}} + \frac{1}{x^2} R_{lm}(1/x^2) \right) \\
& = \frac{e^{-x^2}}{x} \left( \frac{1}{\sqrt{\pi}} + \frac{1}{x^2} \sum_{j=0}^{n} p_j x^{-2j} / \sum_{j=0}^{n} q_j x^{-2j} \right), \quad x \geq 4.0.
\end{align*}
\]

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 \( 6 \times 10^{-19} \) and \( 6 \times 10^{-20} \) 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,  

\[
\Phi_0(z) = \int_0^z \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt 
\]

\[
\simeq \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \sum_{n=0}^{+\infty} \frac{1}{2n+1} \theta_{2n}(z/2), \quad -\infty < z < +\infty. \quad (21)
\]

In this series, \( \theta_n(z) = z^n H_n(z) / n! \), for \( n = 0, 1, 2, \ldots \), and \( H_n(z) \) implies the \( n \)th Hermite polynomial, such that \( H_0(z) = 1 \), \( H_1(z) = z \), and \( H_{n+1}(z) = z H_n(z) - n H_{n-1}(z) \) 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 \( n \),

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

Recently, Marsaglia (2004) provided the approximation below,  

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

(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,
\[
u_0(c) = 1; \quad u_1(c) = -2c; \\
u_k(c) = -2(k-1)u_{k-2}(c) + (-2c)u_{k-1}(c), \quad \text{for} \quad k \geq 2.
\]
Integrating on (23), with respect to $t$ from 0 to $-z/\sqrt{2}$,
\[
I_{erf} = \int_0^{-z/\sqrt{2}} e^{-t^2} dt \simeq I_{Aerf} \\
= \text{sign}(-z) \sum_{i=1}^{B+1} e^{-c_i^2} \sum_{k+1}^{\infty} \frac{u_{(i,k)}(c_i)}{k!} (A_{i+1} - c_i)^k, \tag{24}
\]
where, $B = \text{round}(|z|/\sqrt{2})$, $c_i = A_i = (i-1) \times \text{round}(|z|/B\sqrt{2})$, $A_{B+2} = |z|/\sqrt{2}$, $u_{(i,1)} = 1$, $u_{(i,2)} = -2c_i$, $u_{(i,k)} = -2[(k-2)u_{(i,k-2)} + c_iu_{(i,k-1)}]$, for $i = 1, 2, \ldots, B + 1$ and $k \geq 3$. Exceptionally, $B$ is equated to 1, when $B = \text{round}(|z|/\sqrt{2})$ is equal to zero.
Approximations to the Normal Distribution Function and …  — 65

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

\[
I_{\text{erfe}} = \int_{-z/\sqrt{2}}^{\pm\infty} e^{-t^2} \, dt \simeq I_{\text{Aerf}}
\]

\[
= \text{sign}(-z) \sum_{i=1}^{\infty} \sum_{k=1}^{n-i+\infty} \frac{u(i,k)(c_i)}{k!} (A_i - c_i)^k,
\]

In this case, \( A_i = |z|/\sqrt{2}, \ A_2 = \text{round}(|z|/\sqrt{2}) + 1, \ A_{i+1} = A_i + 1 \) 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

\[
erf(-z/\sqrt{2}) \simeq \frac{2I_{\text{Aerf}}}{\sqrt{\pi}}, \quad (26)
\]

\[
erf(-z/\sqrt{2}) \simeq \frac{2I_{\text{Aerfe}}}{\sqrt{\pi}}, \quad (27)
\]

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

\[
\Phi(z) \simeq \frac{1}{2} \left\{ \frac{\sqrt{\pi}}{2} - 2I_{\text{Aerf}} \right\} \quad \text{for} \quad |z| \leq 4,
\]

\[
\Phi(z) \simeq I_{\text{Aerfe}}/\sqrt{\pi} \quad \text{for} \quad z \leq -4,
\]

\[
\Phi(z) \simeq 1 + I_{\text{Aerfe}}/\sqrt{\pi} \quad \text{for} \quad z \leq -4,
\]

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,

\[
\Phi(-70) = 0.542303960930139932867578667087759716518976172187170282890450e - 1066.
\]
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

$$
\Phi(-600) = 0.6546588205807692852105927713888 \\
1087821194128318531771116943e - 78176.
$$

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).

<table>
<thead>
<tr>
<th>Series</th>
<th>$Z=\pm$18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exact</td>
<td>$0.974094891893715048259189518997\times10^{-72}$</td>
</tr>
<tr>
<td>(20)</td>
<td>$0.974094891893715048708181934747\times10^{-72}$</td>
</tr>
<tr>
<td>(21)</td>
<td>$0.226820907630354110107306715331\times10^{-38}$</td>
</tr>
<tr>
<td>(22)</td>
<td>$0.49999999999757093038012396287\times10^{-00}$</td>
</tr>
<tr>
<td>(28)</td>
<td>$0.974094891893715048259189518997\times10^{-72}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Series</th>
<th>$Z=\pm$9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exact</td>
<td>$0.112858840595384064773550207597\times10^{-18}$</td>
</tr>
<tr>
<td>(20)</td>
<td>$0.112858840595384064738093247631\times10^{-18}$</td>
</tr>
<tr>
<td>(21)</td>
<td>$0.112858840595384064773550207597\times10^{-18}$</td>
</tr>
<tr>
<td>(22)</td>
<td>$0.989596251047682032099597869127\times10^{-08}$</td>
</tr>
<tr>
<td>(28)</td>
<td>$0.112858840595384064773550207597\times10^{-18}$</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Series</th>
<th>$Z=\pm$3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exact</td>
<td>$0.134989803163009452665181476759\times10^{-02}$</td>
</tr>
<tr>
<td>(20)</td>
<td>$0.134989803163009452631102368374\times10^{-02}$</td>
</tr>
<tr>
<td>(21)</td>
<td>$0.134989803163009452665181476759\times10^{-02}$</td>
</tr>
<tr>
<td>(22)</td>
<td>$0.134989803163009452665181476759\times10^{-02}$</td>
</tr>
<tr>
<td>(28)</td>
<td>$0.134989803163009452665181476759\times10^{-02}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Series</th>
<th>$Z=\pm$1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exact</td>
<td>$0.15865525393145705144767454368$</td>
</tr>
<tr>
<td>(20)</td>
<td>$0.158655253931457051377370713583$</td>
</tr>
<tr>
<td>(21)</td>
<td>$0.15865525393145705144767454368$</td>
</tr>
<tr>
<td>(22)</td>
<td>$0.15865525393145705144767454368$</td>
</tr>
<tr>
<td>(28)</td>
<td>$0.15865525393145705144767454368$</td>
</tr>
</tbody>
</table>

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 particular, 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 condi-
tions, 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 evalu-
ation to the mean range of the random variables with normal distri-
bution 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. How-
ever, although, the considered series are much more accurate than
the formulæ, 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.
Approximations to the Normal Distribution Function and ...  _  69

Table 3. The mean range of normal distribution (d₂).

<table>
<thead>
<tr>
<th>n</th>
<th>d₂</th>
<th>n</th>
<th>d₂</th>
<th>n</th>
<th>d₂</th>
<th>n</th>
<th>d₂</th>
<th>n</th>
<th>d₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.12838</td>
<td>31</td>
<td>4.11293</td>
<td>60</td>
<td>4.63856</td>
<td>89</td>
<td>4.93131</td>
<td>460</td>
<td>6.02251</td>
</tr>
<tr>
<td>3</td>
<td>1.69257</td>
<td>32</td>
<td>4.13934</td>
<td>61</td>
<td>4.65112</td>
<td>90</td>
<td>4.93940</td>
<td>480</td>
<td>6.04853</td>
</tr>
<tr>
<td>4</td>
<td>2.05875</td>
<td>33</td>
<td>4.16482</td>
<td>62</td>
<td>4.66346</td>
<td>91</td>
<td>4.94739</td>
<td>500</td>
<td>6.07340</td>
</tr>
<tr>
<td>5</td>
<td>2.32593</td>
<td>34</td>
<td>4.18943</td>
<td>63</td>
<td>4.67557</td>
<td>92</td>
<td>4.95529</td>
<td>520</td>
<td>6.09721</td>
</tr>
<tr>
<td>6</td>
<td>2.53441</td>
<td>35</td>
<td>4.21322</td>
<td>64</td>
<td>4.68747</td>
<td>93</td>
<td>4.96309</td>
<td>540</td>
<td>6.12004</td>
</tr>
<tr>
<td>7</td>
<td>2.70436</td>
<td>36</td>
<td>4.23625</td>
<td>65</td>
<td>4.69916</td>
<td>94</td>
<td>4.97079</td>
<td>560</td>
<td>6.14198</td>
</tr>
<tr>
<td>8</td>
<td>2.84720</td>
<td>37</td>
<td>4.25855</td>
<td>66</td>
<td>4.71065</td>
<td>95</td>
<td>4.97841</td>
<td>580</td>
<td>6.16308</td>
</tr>
<tr>
<td>9</td>
<td>2.97003</td>
<td>38</td>
<td>4.28018</td>
<td>67</td>
<td>4.72194</td>
<td>96</td>
<td>4.98593</td>
<td>600</td>
<td>6.18340</td>
</tr>
<tr>
<td>11</td>
<td>3.17287</td>
<td>40</td>
<td>4.32155</td>
<td>69</td>
<td>4.74397</td>
<td>98</td>
<td>5.00073</td>
<td>640</td>
<td>6.22194</td>
</tr>
<tr>
<td>12</td>
<td>3.25846</td>
<td>41</td>
<td>4.34136</td>
<td>70</td>
<td>4.75472</td>
<td>99</td>
<td>5.00800</td>
<td>660</td>
<td>6.24023</td>
</tr>
<tr>
<td>13</td>
<td>3.33598</td>
<td>42</td>
<td>4.36063</td>
<td>71</td>
<td>4.76529</td>
<td>100</td>
<td>5.01519</td>
<td>680</td>
<td>6.25794</td>
</tr>
<tr>
<td>14</td>
<td>3.40676</td>
<td>43</td>
<td>4.37938</td>
<td>72</td>
<td>4.77570</td>
<td>120</td>
<td>5.14417</td>
<td>700</td>
<td>6.27509</td>
</tr>
<tr>
<td>15</td>
<td>3.47183</td>
<td>44</td>
<td>4.39764</td>
<td>73</td>
<td>4.78595</td>
<td>140</td>
<td>5.25118</td>
<td>720</td>
<td>6.29172</td>
</tr>
<tr>
<td>16</td>
<td>3.53198</td>
<td>45</td>
<td>4.41544</td>
<td>74</td>
<td>4.79604</td>
<td>160</td>
<td>5.34243</td>
<td>740</td>
<td>6.30786</td>
</tr>
<tr>
<td>17</td>
<td>3.58788</td>
<td>46</td>
<td>4.43279</td>
<td>75</td>
<td>4.80598</td>
<td>180</td>
<td>5.42186</td>
<td>760</td>
<td>6.32353</td>
</tr>
<tr>
<td>18</td>
<td>3.64006</td>
<td>47</td>
<td>4.44972</td>
<td>76</td>
<td>4.81578</td>
<td>200</td>
<td>5.49208</td>
<td>780</td>
<td>6.33876</td>
</tr>
<tr>
<td>19</td>
<td>3.68896</td>
<td>48</td>
<td>4.46624</td>
<td>77</td>
<td>4.82543</td>
<td>220</td>
<td>5.55497</td>
<td>800</td>
<td>6.35358</td>
</tr>
<tr>
<td>20</td>
<td>3.73495</td>
<td>49</td>
<td>4.48238</td>
<td>78</td>
<td>4.83493</td>
<td>240</td>
<td>5.61185</td>
<td>820</td>
<td>6.36800</td>
</tr>
<tr>
<td>21</td>
<td>3.77834</td>
<td>50</td>
<td>4.49815</td>
<td>79</td>
<td>4.84431</td>
<td>260</td>
<td>5.66375</td>
<td>840</td>
<td>6.38205</td>
</tr>
<tr>
<td>22</td>
<td>3.81938</td>
<td>51</td>
<td>4.51356</td>
<td>80</td>
<td>4.85355</td>
<td>280</td>
<td>5.71144</td>
<td>860</td>
<td>6.39573</td>
</tr>
<tr>
<td>23</td>
<td>3.85832</td>
<td>52</td>
<td>4.52864</td>
<td>81</td>
<td>4.86266</td>
<td>300</td>
<td>5.75553</td>
<td>880</td>
<td>6.40908</td>
</tr>
<tr>
<td>24</td>
<td>3.89535</td>
<td>53</td>
<td>4.54339</td>
<td>82</td>
<td>4.87165</td>
<td>320</td>
<td>5.79652</td>
<td>900</td>
<td>6.42211</td>
</tr>
<tr>
<td>25</td>
<td>3.93063</td>
<td>54</td>
<td>4.55783</td>
<td>83</td>
<td>4.88051</td>
<td>340</td>
<td>5.83480</td>
<td>920</td>
<td>6.43483</td>
</tr>
<tr>
<td>26</td>
<td>3.96432</td>
<td>55</td>
<td>4.57197</td>
<td>84</td>
<td>4.88926</td>
<td>360</td>
<td>5.87069</td>
<td>940</td>
<td>6.44725</td>
</tr>
<tr>
<td>27</td>
<td>3.99654</td>
<td>56</td>
<td>4.58582</td>
<td>85</td>
<td>4.89789</td>
<td>380</td>
<td>5.90446</td>
<td>960</td>
<td>6.45939</td>
</tr>
<tr>
<td>28</td>
<td>4.02741</td>
<td>57</td>
<td>4.59939</td>
<td>86</td>
<td>4.90641</td>
<td>400</td>
<td>5.93636</td>
<td>980</td>
<td>6.47126</td>
</tr>
<tr>
<td>29</td>
<td>4.05704</td>
<td>58</td>
<td>4.61270</td>
<td>87</td>
<td>4.91481</td>
<td>420</td>
<td>5.96655</td>
<td>1000</td>
<td>6.48287</td>
</tr>
<tr>
<td>30</td>
<td>4.08552</td>
<td>59</td>
<td>4.62575</td>
<td>88</td>
<td>4.92311</td>
<td>440</td>
<td>5.99522</td>
<td>1020</td>
<td>6.49423</td>
</tr>
</tbody>
</table>

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 some existing methods. An extended table for the mean range of the normal variables has been constructed.

References


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


Approximations to the Normal Distribution Function and ...


