APPROXIMATION TO THE CUMULATIVE NORMAL DISTRIBUTION USING HYPERBOLIC TANGENT BASED FUNCTIONS

Beong In Yun

ABSTRACT. This paper presents a method for approximation of the standard normal distribution by using hyperbolic tangent based functions. The presented approximate formula for the cumulative distribution depends on one numerical coefficient only, and its accuracy is admissible. Furthermore, in some particular cases, closed forms of inverse formulas are derived. Numerical results of the present method are compared with those of an existing method.

1. Introduction

For a random variable X following a probability density function(p.d.f.) of the standard normal distribution,

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

the cumulative distribution function (or the cumulative normal distribution) is defined as

$$(2) \qquad P(\xi) := \Pr(X \le \xi) = \frac{1}{2} + \int_0^\xi g(x)dx, \qquad \xi \ge 0.$$

Since analytical integration of the p.d.f. $g(x)$ is not possible, lots of approximation formulas to the cumulative normal distribution $P(\xi)$ have been developed. Most of them are based on the form of series expansion which, in theory, can approximate $P(\xi)$ with arbitrarily high precision by increasing the number of terms [6, 10, 9, 13, 14]. Another approximations are the so-called "ad hoc approximations" which often take simple forms with few numerical coefficients

[1, 4, 5, 7, 8]. General classification and further comments on these approximations are given by Waissi and Rossin [14] and Bryc [2].

In this paper, employing hyperbolic tangent based functions such as $\psi_{r,j}(t)$ in (11) and $\phi_r(t)$ in (15), we propose a p.d.f. resembling $g(x)$. This produces a kind of ad hoc approximation formula for $P(\xi)$ which includes only one independent numerical coefficient $r > 0$. The value of r is determined by a condition upon the corresponding variance. The present formula seems to be rather simple and, by comparison of the numerical results, one can observe that it produces a better approximation error than an existing formula in the work of Lin [7] that also contains one independent numerical coefficient.

Some outstanding approximate formulas for the cumulative normal distribution $P(\xi)$ are summarized in the next section. In Section 3 we develop the hyperbolic tangent based functions and then the related approximation formulas for $P(\xi)$ are induced in Section 4. Numerical results of the present approximations are given in Section 5 including comparison with an existing approximation formula.

2. Overview of approximations to the cumulative distribution

In this section we summarize some typical approximation formulas, denoted by $Q(\xi)$, for the cumulative normal distribution $P(\xi)$ in (2) as follows.

2.1. Approximations based on the series expansion

· Page's approximation [10]:

$$(3) \qquad Q^{\mathrm{Pag}}(\xi) = 1 - \left\{ 1 + \exp \left(\sum_{k=0}^{1} a_k \xi^{2k+1} \right) \right\}^{-1}$$

with $a_0 \xi + a_1 \xi^3 = (1.5976)\xi + (0.070565992)\xi^3$. The absolute error is less than 1.4×10^{-4} over $0 \le \xi < \infty$.

· Waissi and Rossin's approximation [14]:

$$(4) \qquad Q^{\mathrm{Wai}}(\xi) = \left\{ 1 + \exp \left(-\sqrt{\pi} \sum_{k=0}^{2} b_k \xi^{2k+1} \right) \right\}^{-1}$$

with

$$b_0 \xi + b_1 \xi^3 + b_2 \xi^5 = (0.9000000)\xi + (0.0418198)\xi^3 - (0.0004406)\xi^5 .$$

The absolute error is less than 4.3×10^{-5} over $0 \le \xi < 8$ (Blows up for $\xi \ge 10$).

2.2. Ad hoc approximations

· Lin's approximation [7]:

$$(5) \qquad Q^{\mathrm{Lin}}(\xi) = 1 - \left\{ 1 + \exp \left(\frac{4.2\pi\xi}{9 - \xi} \right) \right\}^{-1} .$$

The absolute error is less than 6.8×10^{-3} over $0 \leq \xi < 9$ (Blows up for $\xi \geq 9$).

· Bryc's uniform approximation [2]:

The two-numerical constants formula is

$$(6) \qquad Q_2^{\mathrm{Bry}}(\xi) = 1 - \frac{\xi + 3.333}{\sqrt{2\pi}\xi^2 + 7.32\xi + 2 \times 3.333} \, e^{-\xi^2/2}.$$

The absolute error is less than 7.1×10^{-4} over $0 \leq \xi < \infty$ (The largest error occurs in the range $1.07 \leq \xi \leq 1.13$). In addition, the four-numerical constants formula is

$$(7)$$
$$Q_4^{\mathrm{Bry}}(\xi) = 1 - \frac{\xi^2 + 5.575192695\xi + 12.77436324}{\sqrt{2\pi}\xi^3 + 14.38718147\xi^2 + 31.53531977\xi + 2 \times 12.77436324} e^{-\xi^2/2}.$$

The absolute error is less than 1.9×10^{-5} over $0 \leq \xi < \infty$ (The largest error occurs in the range $1.43 \leq \xi \leq 1.61$).

· Hart's uniform approximation [5]:

$$(8)$$
$$Q^{\mathrm{Har}}(\xi) = 1 - \frac{e^{-\xi^2/2}}{\sqrt{2\pi}\xi} \left(1 - \frac{\sqrt{1 + b\xi^2}/(1 + a\xi^2)}{P_0\xi + \sqrt{P_0^2\xi^2 + \exp(-\xi^2/2)\sqrt{1 + b\xi^2}/(1 + a\xi^2)}} \right),$$

where

$$a = (1 + \sqrt{1 - 2\pi^2 + 6\pi})/2\pi, \qquad b = 2\pi a^2, \qquad P_0 = \sqrt{\pi/2}.$$

The absolute error is less than 5.4×10^{-5} over $0 \leq \xi < \infty$ (The largest error occurs in the range $1.03 \leq \xi \leq 1.04$).

· Bagby's uniform approximation [1]:

$$(9)$$
$$Q^{\mathrm{Bag}}(\xi) = \frac{1}{2} + \frac{1}{2} \left\{ 1 - \frac{1}{30} \left[7e^{-\xi^2/2} + 16e^{-\xi^2(2-\sqrt{2})} + \left(7 + \frac{\pi}{4}\xi^2 \right) e^{-\xi^2} \right] \right\}^{1/2}.$$

The error $Q^{\mathrm{Bag}}(\xi) - P(\xi)$ varies from -3×10^{-5} near $\xi = 0.30$ to 3×10^{-5} near $\xi = 1.70$, and it vanishes both as $\xi \to 0$ and as $\xi \to \infty$.

· Moran's approximation [8]:

$$(10) \qquad Q^{\mathrm{Mor}}(\xi) = \frac{1}{2} + \frac{1}{\pi} \left\{ \frac{\xi}{3\sqrt{2}} + \sum_{k=1}^{12} \frac{1}{k} e^{-k^2/9} \sin(k\xi\sqrt{2}/3) \right\}.$$

The absolute error is less than $3. \times 10^{-10}$ over $0 \leq \xi \leq 7$ (Blows up for $\xi \geq 11$).

In this work, focusing on the numerical simplicity, we are interested in developing a new ad hoc approximation formula for the cumulative normal distribution $P(\xi)$.

3. Hyperbolic tangent based functions

For any integer $j \geq 1$ and for any real $r > 0$, we define a so-called *hyperbolic tangent based function*

$$(11) \quad \psi_{r,j}(t) = \frac{1}{2}\left\{1 + \tanh\left[\frac{r}{j2^{j+1}}\left(\frac{1}{(1-t)^j} - \frac{1}{t^j}\right)\right]\right\}, \qquad 0 \leq t \leq 1$$

which is a strictly increasing map from an interval $[0,1]$ onto itself with a derivative

$$(12) \ \psi'_{r,j}(t) = \frac{r}{2^{j+2}}\left\{\frac{1}{(1-t)^{j+1}} + \frac{1}{t^{j+1}}\right\}\operatorname{sech}^2\left[\frac{r}{j2^{j+1}}\left(\frac{1}{(1-t)^j} - \frac{1}{t^j}\right)\right].$$

Particularly, when $j = 1$ we have

$$(13) \qquad\qquad \psi_{r,1}(t) = \frac{1}{2}\left\{1 + \tanh\left[\frac{r}{4}\left(\frac{1}{1-t} - \frac{1}{t}\right)\right]\right\}$$

which is the transformation introduced by Sag and Szekeres [12], i.e., an infinite order sigmoidal transformation for any $r > \sqrt{3}$ (See Elliott [3]). Sigmoidal transformations are mainly used for accurate numerical evaluation of the singular integrals via coordinate transformations.

We can observe that the function $\psi_{r,j}(t)$ has the following properties.

(P1) $\psi_{r,j}(t) \in \mathbb{C}^\infty[0,1]$ with

$$\psi_{r,j}(0) = 0\,, \qquad \psi_{r,j}\left(\frac{1}{2}\right) = \frac{1}{2}\,, \qquad \psi_{r,j}(1) = 1.$$

(P2) $\psi_{r,j}(t) + \psi_{r,j}(1-t) = 1\,, \quad 0 \leq t \leq 1.$

(P3) Every order of derivatives of $\psi_{r,j}(t)$ vanish at both end points, that is,

$$\psi_{r,j}^{(k)}(0) = \psi_{r,j}^{(k)}(1) = 0\,, \qquad k = 1,\, 2,\, \ldots.$$

From (12) we note that the first derivative of $\psi_{r,j}(t)$ has a local maximum at $t = \frac{1}{2}$ as

$$(14) \qquad\qquad\qquad \psi'_{r,j}\left(\tfrac{1}{2}\right) = r$$

independently of j.

On the other hand we may consider another form of the hyperbolic tangent based function

$$(15) \qquad \phi_r(t) = \frac{1}{2}\left\{1 + \tanh\left[r\tanh^{-1}(2t-1)\right]\right\}, \qquad 0 \leq t \leq 1$$

with a derivative

$$(16) \qquad\qquad \phi'_r(t) = \frac{r}{4t(1-t)}\operatorname{sech}^2\left[r\tanh^{-1}(2t-1)\right].$$

In fact, since $\tanh^{-1}(t) = \frac{1}{2}\log\left(\frac{1+t}{1-t}\right)$, $\phi_r(t)$ becomes

$$\phi_r(t) = \frac{t^r}{t^r + (1-t)^r}$$

which is the so-called simple sigmoidal transformation presented by Prössdorf and Rathsfeld [11]. The function $\phi_r(t)$ also satisfies the properties (P1) and (P2) with

$$\phi_r^{(k)}(t) = O\left(t^{r-k}\right), \qquad k = 1, 2, \ldots, \lfloor r \rfloor$$

near $t = 0$ and $t = 1$. Moreover, $\phi_r(t)$ has its inverse such as

$$(17) \qquad \left(\phi_r\right)^{-1}(s) = \phi_{1/r}(s), \qquad 0 \leq s \leq 1$$

and a local maximum of the first derivative is

$$(18) \qquad \phi_r'\left(\tfrac{1}{2}\right) = r.$$

4. The probability density function

For a real $a > 0$, employing the hyperbolic tangent based function $\psi_{r,j}$, we define a function $h_{r,j}$ on the real line $\mathbb{R}$ such as

$$(19) \qquad h_{r,j}(x) := \begin{cases} \dfrac{1}{2a}\, \psi_{r,j}'\left(\dfrac{a+x}{2a}\right), & |x| \leq a, \\ 0, & |x| > a. \end{cases}$$

We can see that $h_{r,j}$ is compact supported with $h_{r,j} \in \mathbb{C}^\infty(\mathbb{R})$ and it is a well-defined p.d.f. on $\mathbb{R}$ by observing that

$$\int_{-\infty}^{\infty} h_{r,j}(x)\, dx = \int_0^1 \psi_{r,j}'(t)\, dt = 1$$

from the property (P1).

Moreover, from the properties of $\psi_{r,j}(t)$ and the formula (14), we note that a local maximum of $h_{r,j}$ is

$$h_{r,j}(0) = \frac{r}{2a}$$

independently of j. In order to approximate the p.d.f of the standard normal distribution given in (1) we use a constraint, $h_{r,j}(0) = g(0) = 1/\sqrt{2\pi}$. Then the half length of the support of $h_{r,j}$ satisfies

$$(20) \qquad a = \sqrt{\frac{\pi}{2}}\, r.$$

Therefore, for any integer j chosen, we only have to determine the coefficient r of the function $h_{r,j}$ in (19). Let X be a random variable related to the p.d.f. $h_{r,j}(x)$. Then, by change of variables and integration by parts, the second moment of X (or the variance of X) becomes

$$\begin{aligned} E(X^2) &= \frac{1}{2a} \int_{-a}^{a} x^2 \psi_{r,j}'\left(\frac{a+x}{2a}\right) dx \\ &= a^2 \left\{ 1 - 4 \int_0^1 (2t-1)\, \psi_{r,j}(t)\, dt \right\}. \end{aligned}$$

To estimate r, we propose a condition that the variance of X above should be equal to that of the standard normal distribution as below.

$$(21) \qquad \frac{\pi}{2} r^2 \left\{ 1 - 4 \int_0^1 (2t - 1)\, \psi_{r,j}(t)\, dt \right\} = 1$$

for each integer $j \geq 1$ fixed. For some values of j, corresponding numerical solutions for r of the equation (21) are given in Table 1. Therein, r^* indicates the optimal value for which approximation of the cumulative distribution function $P(\xi)$ proposed in the next section results in the best absolute error, numerically. It is seen that the values of r and a are increasing as j goes higher.

Table 1. Numerical solutions for r of the equation (21) and related values of a based on the formula (20).

j	r	a	r^*	a^*
1	4.1369	5.1848	4.04	5.0759
2	5.7248	7.1749	5.60	7.0186
4	8.9047	11.160	8.76	10.966
6	12.085	15.146	11.9	14.914
8	15.265	19.132	15.1	18.925
10	18.445	23.117	18.2	22.810

For instance, the approximation errors of $h_{r,j}(x)$, $j = 1, 4$, to the p.d.f. $g(x)$ in (1) are shown in Figure 1. The value of r for each j was chosen from Table 1. From the numerical results it was observed that, as j goes higher, the error becomes somewhat better and this tendency is limited, however.

5. The cumulative normal distribution

We denote by $H_{r,j}(\xi)$ a cumulative distribution function corresponding to the p.d.f. $h_{r,j}$ in the form of

$$(22)$$

$$H_{r,j}(\xi) = \int_{-\infty}^{\xi} h_{r,j}(x)\, dx = \psi_{r,j}\left(\frac{a + \xi}{2a} \right)$$

$$= \frac{1}{2} \left\{ 1 + \tanh\left[\frac{r}{2j} \left(\frac{1}{(1 - \xi/a)^j} - \frac{1}{(1 + \xi/a)^j} \right) \right] \right\}, \quad 0 \leq \xi \leq a$$

with a satisfying the relation (20). Since $H_{r,j}(a) = 1$, we may define $H_{r,j}(\xi) = 1$ for all $\xi > a$. Then $H_{r,j}(\xi) \in \mathbf{C}^{\infty}(0, \infty)$ and it can be used for approximation to the cumulative normal distribution $P(\xi)$ given in (2).

By the numerical experiment using a symbolic program, _Mathematica_ v.5, we have the following results for approximation errors: For $j = 1$

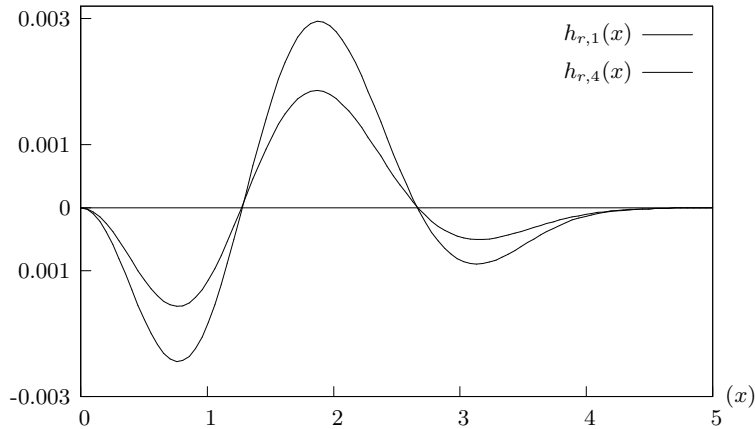$$\| H_{r^*,1} - P \|_{\infty} \leq 1.8 \times 10^{-3}$$

Figure 1. Differences of the functions $h_{r,1}(x)$ and $h_{r,4}(x)$ from the p.d.f. $g(x)$ of the standard normal distribution.

and for any $j \geq 2$

$$\|H_{r^*,j} - P\|_\infty \leq 8.9 \times 10^{-4}$$

with r^* and a^* chosen according to each j as given in Table 1. It can be seen that as j goes higher, the error becomes better. For example, Figure 2 shows errors $H_{r^*,j}(\xi) - P(\xi)$, $0 \leq \xi \leq 5$ for each $j = 1, 4, 10$ combined with the values of r^* and a^* in Table 1. In addition, it is observed that the error has extremes near $\xi = 1.2$ and $\xi = 2.5$ for all j.

In the formulas (19)-(22), replacing the hyperbolic tangent based function $\psi_{r,j}$ by ϕ_r defined in (15), we have another simple approximation such as

(23)
$$H_r(\xi) = \phi_r \left(\frac{a + \xi}{2a} \right)$$
$$= \frac{1}{2} \left\{ 1 + \tanh \left[r \cdot \tanh^{-1} \left(\frac{\xi}{a} \right) \right] \right\}, \quad 0 \leq \xi \leq a$$

and $H_r(\xi) = 1$ for all $\xi > a$, where r is determined by solving the equation

(24)
$$\frac{\pi}{2} r^2 \left\{ 1 - 4 \int_0^1 (2t - 1)\, \phi_r(t)\, dt \right\} = 1$$

and a satisfies the relation (20). The numerical solution of (24) can be obtained as $r = 2.5673$ by using the Newton method. Instead of this vale for r, however, a more appropriate value r^* can be chosen to get a better approximation to $P(\xi)$ like the case of $H_{r,j}(\xi)$.

Furthermore, we can obtain a closed form of the inverse formula of H_r as

(25)
$$H_r^{-1}(\eta) = a \cdot \tanh \left[\frac{1}{r} \tanh^{-1} (2\eta - 1) \right]$$

for $\frac{1}{2} \leq \eta \leq 1$. Similarly, from (22) with $j = 1$, we have an inverse of $H_{r,1}$ in the form of

$$(26) \quad H_{r,1}^{-1}(\eta) = \frac{ar}{2\tanh^{-1}(2\eta - 1)} \left\{ -1 + \sqrt{1 + \left(\frac{2\tanh^{-1}(2\eta - 1)}{r} \right)^2} \right\}.$$

It should be noted that the present approximations $H_{r,j}(\xi)$ with j fixed and $H_r(\xi)$ include only one independent numerical coefficient, r and a constant, a related with r via (20) so that comparison with the simplest existing approximation $Q^{\text{Lin}}(\xi)$ containing one numerical coefficient seems to be reasonable. In Figure 3, the errors of $H_{r^*,10}(\xi)$ with $r^* = 18.2$ and $H_{r^*}(\xi)$ with $r^* = 2.48$ are compared with that of the Lin's formula $Q^{\text{Lin}}(\xi)$ given in (5).
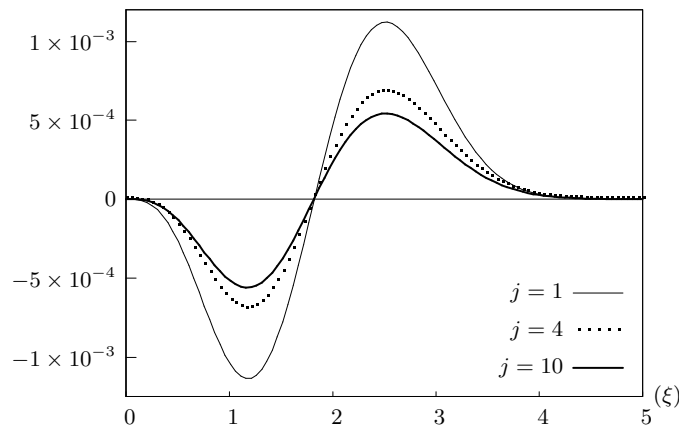

Figure 2. Differences of the present approximations $H_{r^*,j}(\xi)$, $j = 1, 4, 10$, from the cumulative normal distribution $P(\xi)$.

6. Conclusions

We proposed a procedure to approximate the cumulative normal distribution $P(\xi)$ by using the so-called hyperbolic tangent based function. Despite containing only one independent numerical coefficient, the presented formula in (22) gives better results than the existing Lin's approximation formula which also includes one independent numerical coefficient.

The use of another hyperbolic tangent based function $\phi_r(t)$, defined in (15), instead of $\psi_{r,j}(t)$ in the present method is also available. The resultant approximation formula in (23) is rather simpler. However, the approximation error is slightly worse than the case combined with $\psi_{r,j}(t)$.

Additionally, in the case of $\phi_r(t)$ or $\psi_{r,1}(t)$, we can use the inverse formula (25) or (26) for a given value of the cumulative normal distribution $P(\xi)$.
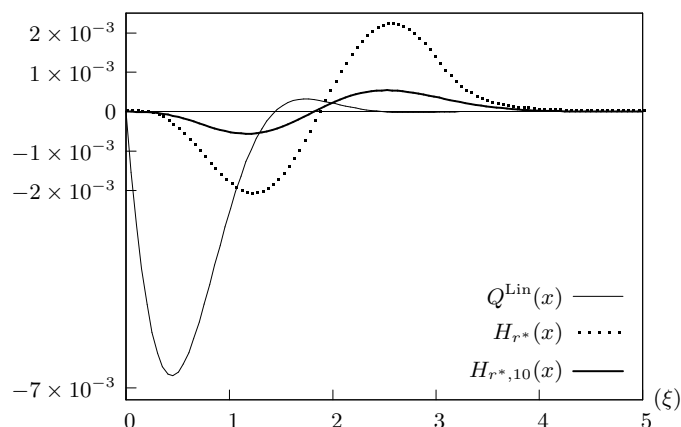
Figure 3. Differences of the approximations $Q^{\mathrm{Lin}}(x)$, $H_{r^*}(\xi)$ and $H_{r^*,10}(\xi)$ from the cumulative normal distribution $P(\xi)$.

References

[1] R. J. Bagby, *Calculating normal probabilities*, Amer. Math. Monthly **102** (1995), no. 1, 46–48.

[2] W. Bryc, *A uniform approximation to the right normal tail integral*, Appl. Math. Comput. **127** (2002), no. 2-3, 365–374.

[3] D. Elliott, *Sigmoidal transformations and the trapezoidal rule*, J. Aust. Math. Soc. Ser. B **40** (1998/99), (E), E77–E137.

[4] H. Hamaker, *Approximating the cumulative normal distribution and its inverse*, Appl. Statist. **27** (1978), 76–77.

[5] R. G. Hart, *A closed approximation related to the error function*, Math. Comp. **20** (1966), 600–602.

[6] R. G. Kerridge and G. W. Cook, *Yet another series for the normal integral*, Biometrika **63** (1976), 401–407.

[7] J. T. Lin, *A simpler logistic approximation to the normal tail probability and its inverse*, J. Roy. Statist. Soc. Ser. C **39** (1990), no. 2, 255–257.

[8] P. A. P. Moran, *Caculation of the normal distribution function*, Biometrika **67** (1980), 675–676.

[9] R. M. Norton, *Pocket-calculator approximation for areas under the standard normal curve*, Amer. Statist. **43** (1989), 24–26.

[10] E. Page, *Approximation to the cumulative normal function and its inverse for use on a pocket calculator*, Appl. Statist. **26** (1977), 75–76.

[11] S. Prössdorf and A. Rathsfeld, *On an integral equation of the first kind arising from a cruciform crack problem*, Integral equations and inverse problems (Varna, 1989), 210–219, Pitman Res. Notes Math. Ser., 235, Longman Sci. Tech., Harlow, 1991.

[12] T. W. Sag and G. Szekeres, *Numerical evaluation of high-dimensional integrals*, Math. Comp. **18** (1964), 245–253.

[13] J. D. Vedder, *An invertible approximation to the normal-distribution function*, Comput. Statist. Data Anal. **16** (1993), 119–123.

[14] G. R. Waissi and D. F. Rossin, *A sigmoid approximation to the standard normal integral*, Appl. Math. Comput. **77** (1996), 91–95.

School of Mathematics, Informatics and Statistics
Kunsan National University
Kunsan 573-701, Korea
E-mail address: biyun@kunsan.ac.kr