The Mernagh Expansion to approximate standard normal distributions

Michael Mernagh, *Proceedings of the 26th Conference on Applied Statistics in Ireland (2006)*

An expansion containing cosine terms can be derived to approximate the equation of the curve of the univariate standard normal distribution

The first method is based on Maclaurin series expansions. The second method uses multiple regression.

A polynomial expression can also be derived equation of the curve of the univariate standard normal distribution.

An expansion involving cosines can be derived for the equations of the outer curves of the slicing planes that bisect the bivariate standard normal distribution transversely and diagonally.

The equation for the surface of the bivariate normal distribution can be derived by multiple regression, based on the product of expansions of two *independent* standard normal distributions.

The equation for the hyper-surface of the trivariate normal distribution (in 4-dimensional space) can be derived by multiple regression, based on the product of expansions of three *independent* standard normal distributions.

An expansion for the hyper-surface of the multivariate normal distribution (in ***n***-dimensional space) is described, involves the product of ***n*–1** *independent* standard normal distributions.

The expression for the Maclaurin cosine expansions to represent the univariate standard normal distribution may be stated generally in the following formula:

$$\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)=\frac{1}{\sqrt{2\pi}}\sum_{i=1}^{n}\beta_i\cos\frac{(2i-1)x}{2}$$

where $\sum_{i=i}^{n}\beta_i=1$

and $|\beta_{i+1}|<|\beta_i|<1$ (1)

For 5 terms

$$y=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)=0.3989\left\{\begin{aligned}&0.704146\cos\frac{x}{2}+0.259003\cos\frac{3x}{2}+0.035081\cos\frac{5x}{2}\\&+0.001735\cos\frac{7x}{2}+0.000034\cos\frac{9x}{2}\end{aligned}\right\} \quad (2)$$

Two comments arise from the foregoing analysis:

(a) When $x = -\pi$ or when $x = +\pi$, then $y = 0$

(b) Frequentists are accustomed to referring to the standard normal distribution as ranging from $\boldsymbol{x} = -\mathbf{3}$ to $\boldsymbol{x} = +\mathbf{3}$. Maybe they should think of it as ranging from $\boldsymbol{x} = -\mathbf{3.14}$ to $\boldsymbol{x} = +\mathbf{3.14}$, which would help to justify the use of a normal prior ranging between $\boldsymbol{x} = -\boldsymbol{\pi}$ and $\boldsymbol{x} = +\boldsymbol{\pi}$ that is often used in Bayesian analyses.

The following is a polynomial expansion for the standard normal distribution, from which a related polynomial expansion for the cumulative standard normal distribution is then derived:

$$\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)=\frac{1}{\sqrt{2\pi}}\sum_{i=0}^{\infty}\frac{(-1)^i x^{2i}}{2^i(i!)}$$

$$=\frac{1}{\sqrt{2\pi}}\left\{1-\frac{x^2}{2}+\frac{x^4}{4(2!)}-\frac{x^6}{8(3!)}+\frac{x^8}{16(4!)}-\frac{x^{10}}{32(5!)}+\frac{x^{12}}{64(6!)}-\frac{x^{14}}{128(7!)}\right\} \quad (24)$$

$$=0.398942\left\{\begin{array}{l}1-0.5x^2+0.125x^4-0.02083x^6+0.0026x^8\\ \qquad -0.00026x^{10}+0.00022x^{12}-0.000002x^{14}\end{array}\right\} \quad (25)$$

$$\begin{aligned}\mathrm{F}(\boldsymbol{t}) = \; & -0.0000000047\,\boldsymbol{t}^{15}\\ & +0.0000002552\,\boldsymbol{t}^{13}\\ & -0.0000064267\,\boldsymbol{t}^{11}\\ & +0.0001015033\,\boldsymbol{t}^{9}\\ & -0.0011488553\,\boldsymbol{t}^{7}\\ & +0.0099153564\,\boldsymbol{t}^{5}\\ & -0.0664499601\,\boldsymbol{t}^{3}\\ & +0.3989339506\,\boldsymbol{t}\\ & +\mathbf{0.5}\end{aligned} \quad (27)$$

and from this expression of $F(t)=\int_{-\infty}^{x} y\,dt$ the probability of any value of $\boldsymbol{x}$ between -3.1 and $+3.1$ can be calculated exactly by substituting any value of $\boldsymbol{x} = \boldsymbol{t}$ in the polynomial expansion for $\boldsymbol{F(t)}$ and hence a table of the probabilities corresponding to all of the $\boldsymbol{x}$ values can be compiled.

Generally, the expansion involving products of three cosine terms for the equation of the hyper-surface of the multivariate normal distribution (in n-dimensional space) is:

$$f\left(\left[\vec{X}_m\right]\right)=\sum_{j=1}^{n-1}\beta_j\prod_{i=1}^{3}\cos\left\{\frac{2i-1}{2}\left(x_{k,m}\right)\right\}+c \quad (38)$$

where $\left[\ \tilde{X}\ \right]=\left[\left(x_{k,m}\right)\right]$ *for* $k=1$ *to* 3 *and* $m=1$ *to* 3^{n-2}

and $\left[\vec{X}_m\right]$ is a row vector of $\left[\ \tilde{X}\ \right]$ and $\boldsymbol{c}$ is a small constant

and $\sum_{j=1}^{n-1}\beta_j+c=f\left(\left[\vec{0}_m\right]^T\right)=\left(\frac{1}{\sqrt{2\pi}}\right)^{n-1}$ at the origin and $\left|\beta_j\right|>\left|\beta_{j+1}\right|$