Journal of Statistical and Econometric Methods, vol. 2, no.1, 2013, 1-12
ISSN: 2051-5057 (print version), 2051-5065(online)
Scienpress Ltd, 2013

Tail Approximation of the Skew-Normal by the Skew-Normal Laplace: Application to Owen's T Function and the Bivariate Normal Distribution

Werner Hürlimann[1]

Abstract

By equal mean, two skew-symmetric families with the same kernel are quite similar, and the tails are often very close together. We use this observation to approximate the tail distribution of the skew-normal by the skew-normal-Laplace, and accordingly obtain a normal function approximation to Owen's T function, which determines the survival function of the skew-normal distribution. Our method is also used to derive skew-normal-Laplace approximations to the bivariate standard normal distribution valid for large absolute values of the arguments and correlation coefficients, a situation difficult to handle with traditional numerical methods.

Mathematics Subject Classification: 60E15, 62E17, 62E20
Keywords: Skew-symmetric, skew-normal, skew-normal-Laplace, Owen's T function, normal approximation, asymptotic approximation, bivariate normal

1 Introduction

There are many known methods available to extend symmetric families of probability densities to asymmetric models. An important general method is due to [3]. Consider a continuous density $f(x)$, called the kernel, which is symmetric about zero, and let $G(x)$ be an absolutely continuous distribution, called the skewing distribution, that is symmetric about zero. Then, for any skewing parameter $\lambda \in (-\infty, \infty)$, the function

---

[1]Wolters Kluwer Financial Services Switzerland AG,
e-mail: whurlimann@bluewin.ch

Article Info: *Received :* December 3, 2012. *Revised :* December 30, 2012.
Published online : March 1, 2013

$g_\lambda(x) = 2f(x)G(\lambda x)$ defines a probability density. Skew families of distributions have been studied in [13] and [14] for the normal and Cauchy kernels, and in [1] and [2] for the gamma and Laplace kernels.

It can be asked how different two skew-symmetric families with the same kernel can be. This main question has been studied by [21], who argues that the shape of the family of skew-symmetric distributions with fixed kernel is not very sensitive to the choice of the skewing distribution. In fact, by equal mean, two such distributions will be quite similar because they have the first and all even moments in common. While the maximum absolute deviation is often attained at zero, which implies that the distributions will differ significantly only in a neighborhood of zero, the tails are often very close together. Applied to statistical approximation this observation motivates the search of an analytically simple choice of skewing distribution in order to approximate a given skew-symmetric family without closed-form distribution expression. A result of this type is used to derive a normal and an elementary function approximation to Owen's T function (introduced in [16]), which determines the survival function of the skew-normal distribution. The use of the Laplace skewing distribution turns out to be appropriate for this. The paper is organized as follows.

Section 2 presents a brief account of the results from [21] on the similarity of the shape of skew-symmetric families with the same kernel, which motivates our study as explained above. Section 3 derives our normal approximation to Owen's T function by analyzing the signed difference between the skew-normal and the skew-normal-Laplace distribution. Given a small maximum absolute error, and some technical conditions, there exists a common threshold parameter such that the signed difference between Owen's T function and its normal analytical approximation remains doubly uniformly bounded for all parameter values exceeding this threshold. Numerical examples illustrate our findings. Section 4 applies our method to the bivariate normal distribution and derives double uniform bounds especially useful for large absolute values of the arguments and correlation coefficients, a situation difficult to handle with traditional numerical methods.

2 Skew-symmetric Families and their Shape

The density function of the *skew-symmetric family* with continuous (zero) symmetric *density kernel* $f(x)$, absolutely continuous and (zero) symmetric *skewing distribution* $G(x)$, and *skewing parameter* $\lambda \in (-\infty, \infty)$, is defined by $g_\lambda(x) = 2f(x)G(\lambda x)$. This family has been introduced in [3], [4] and includes the *skew normal* by [15]. Skew families of distributions have been studied in [13] and [14] for the normal and Cauchy kernels, and in [1] and [2] for the gamma and Laplace kernels.

It can be asked how different two skew-symmetric families with the same kernel can be. This main question has been studied in [21]. In general, consider besides the $g_\lambda(x)$ with distribution $G_\lambda(x)$ the density $h_\gamma(x) = 2f(x)H(\gamma x)$ for another skewing distribution $H(x)$ with skewing parameter $\gamma \in (-\infty, \infty)$, and let $H_\gamma(x)$ be its distribution. Assume the kernel has a first moment. Then, the means of these skew-symmetric distributions are:

$$\mu_G(\lambda) = \int_{-\infty}^{\infty} x g_\lambda(x)dx = \int_{-\infty}^{\infty} 2xf(x)G(\lambda x)dx,$$
$$\mu_H(\gamma) = \int_{-\infty}^{\infty} x h_\gamma(x)dx = \int_{-\infty}^{\infty} 2xf(x)H(\gamma x)dx. \tag{2.1}$$

It is shown in [20] that both $\mu_G(\lambda)$ and $\mu_H(\gamma)$ are strictly increasing, odd functions such that $\mu_G(0) = \mu_H(0)$ and $\lim_{\lambda\to\infty} \mu_G(\lambda) = \lim_{\gamma\to\infty} \mu_G(\gamma) = \int_0^{\infty} 2xf(x)dx$ is the mean of the folded distribution. Moreover, it is well-known that the even moments are the same as those of the folded kernel whatever the parameters are. Therefore, every member of the families $\{g_\lambda\}$ and $\{h_\gamma\}$ has the same set of even moments independently of the choice of the skewing distribution. By equal mean the distributions $G_\lambda(x)$ and $H_\gamma(x)$ will be quite similar because they have the first and all even moments in common. Now, the above properties about the mean guarantee the existence of a unique value $\gamma = \phi(\lambda)$ such that $\mu_G(\lambda) = \mu_H(\phi(\lambda))$ for each λ. One notes that $\phi(\lambda) < 0$ for $\lambda < 0$, $\phi(0) = 0$, and $\phi(\lambda) > 0$ for $\lambda > 0$. Under these conditions it is natural to consider the following *measure of discrepancy* between two skew-symmetric families with the same kernel [21]:

$$d(G_\lambda, H_\gamma) = \max_{x,\lambda} \left| G_\lambda(x) - H_{\phi(\lambda)}(x) \right| \tag{2.2}$$

Since the difference in distributions $G_\lambda(x) - H_\gamma(x)$ is an even function whatever the parameters are ([21], Theorem, p.663), the search of the maximum discrepancy can be restricted either to the interval $(-\infty, 0]$ or $[0, \infty)$. By construction (symmetry of the kernel about zero) it is clear that $x = 0$ is a critical value such that $G'_\lambda(0) - H'_{\phi(\lambda)}(0) = g_\lambda(0) - h_{\phi(\lambda)}(0) = 0$.
From the property $\lim_{x\to\pm\infty} \{G_\lambda(x) - H_\gamma(x)\} = 0$ we conclude that $x = 0$ yields the actual maximum $\max_x \left| G_\lambda(x) - H_{\phi(\lambda)}(x) \right| = G_\lambda(0) - H_{\phi(\lambda)}(0)$ in case the difference $G_\lambda(x) - H_{\phi(\lambda)}(x)$ has no other critical points. In the light of these findings we say that a pair $\{g_\lambda, h_{\phi(\lambda)}\}$ of skew-symmetric families with the same kernel, identical mean and even moments, is a *regular* pair provided the following condition is fulfilled:

$$g_\lambda(x) \neq h_{\phi(\lambda)}(x), \quad \forall\, x \neq 0, \quad \forall\, \lambda \in (-\infty, \infty) \tag{2.3}$$

It follows that for regular pairs, the distance between the skew-symmetric families defined by:

$$d(G_\lambda, H_\gamma) = \max_{x,\lambda>0} \left| G_\lambda(x) - H_{\phi(\lambda)}(x) \right| = \max_{\lambda>0} \left| G_\lambda(0) - H_{\phi(\lambda)}(0) \right| \tag{2.4}$$

is a meaningful measure of discrepancy [21].

In general, while the maximum absolute deviation in (2.4) is attained at $x=0$ and the distributions will differ significantly only in a neighborhood of $x=0$, the left and right tails of $G_\lambda(x)$ and $H_{\phi(\lambda)}(x)$ are often very close together. In particular, this means that the survival functions $\overline{G}_\lambda(x)=1-G_\lambda(x)$ and $\overline{H}_{\phi(\lambda)}(x)=1-H_{\phi(\lambda)}(x)$ for pairs with small discrepancy are expected to differ only very slightly by increasing value of x. More precisely, it is possible to investigate whether, for a sufficiently high threshold x_0 and a parameter value λ_δ, the tail difference remains doubly uniformly bounded by a given small maximum absolute error δ such that:

$$\left|\overline{G}_\lambda(x)-\overline{H}_{\phi(\lambda)}(x)\right|\le\delta,\quad \forall\, x\ge x_0,\ \forall\,\lambda\ge\lambda_\delta \tag{2.5}$$

Applied to statistical approximation, this observation motivates the search for an analytically simple choice of skewing distribution $H(x)$ in order to approximate a skew-symmetric family without closed-form distribution expression. A result of this type is used in the next Section to derive a normal and an elementary function approximation to Owen's T function (introduced in [16]), which determines the survival function of the skew-normal distribution. The use of the Laplace skewing distribution turns out to be appropriate for this.

3 Approximations to Owen's T Function

Consider the skew-normal density $g_\lambda(x)=2f(x)\Phi(\lambda x),\ \lambda>0,$ and the skew-normal-Laplace density $h_\gamma(x)=2f(x)H(\gamma x)$, with $H(x)=\frac{1}{2}e^x,\ x<0$, $H(x)=1-\frac{1}{2}e^{-x},\ x\ge 0$ [21]. Example 1B, shows that $d(G_\lambda,H_\gamma)=0.01287$ is attained at $\lambda=1.47$, $\gamma=\phi(\lambda)=1.86071$, so that these two families are quite similar. The survival function of the skew-normal reads:

$$\overline{G}_\lambda(x)=\overline{\Phi}(x)+2\cdot T(x,\lambda), \tag{3.1}$$

where $\overline{\Phi}(x)=1-\Phi(x)$ is the standard normal survival function, and $T(x,\lambda)$ is Owen's T function defined by:

$$T(x,\lambda)=\frac{1}{2\pi}\cdot\int_0^\lambda\frac{\exp\left\{-\frac{1}{2}x^2(1+t^2)\right\}}{1+t^2}dt=\varphi(0)\cdot\varphi(x)\cdot\int_0^\lambda\frac{\varphi(xt)}{1+t^2}dt,\quad \varphi(x)=\Phi'(x) \tag{3.2}$$

The skew-normal-Laplace survival function is given by:

$$\overline{H}_\gamma(x)=\begin{cases}1-\varphi(0)\cdot M(\gamma)+e^{\frac{1}{2}\gamma^2}\cdot\left\{\Phi(\gamma-x)-\Phi(\gamma)\right\}, & x\le 0,\\ 2\cdot\overline{\Phi}(x)-e^{\frac{1}{2}\gamma^2}\cdot\overline{\Phi}(x+\gamma), & x\ge 0,\end{cases} \tag{3.3}$$

with $M(x)=\overline{\Phi}(x)/\varphi(x)$ the Mill's ratio.

To derive a double uniform approximation of the type (2.5) we proceed as follows. Let

$\Delta(x,\lambda)=\overline{G}_{\lambda}(x)-\overline{H}_{\phi(\lambda)}(x)$ be the signed difference between these survival functions. To analyze its monotonic properties with respect to the parameter λ we consider its partial derivative:

$$\frac{\partial}{\partial\lambda}\Delta(x,\lambda)=\frac{\partial}{\partial\lambda}\overline{G}_{\lambda}(x)-\gamma'(\lambda)\cdot\frac{\partial}{\partial\gamma}\overline{H}_{\gamma}(x)\Big|_{\gamma=\phi(\lambda)} \tag{3.4}$$

The derivative $\gamma'(\lambda)=\phi'(\lambda)$ in (3.4) is obtained as follows. By definition the function $\gamma=\gamma(\lambda)=\phi(\lambda)$ solves the mean equation $\mu_H(\gamma)=\mu_G(\lambda)$ (apply partial integration or [13], Section 5, for the skew-normal-Laplace mean):

$$2\gamma\cdot\exp\left\{\tfrac{1}{2}\gamma^2\right\}\cdot\overline{\Phi}(\gamma)=\sqrt{\frac{2}{\pi}}\cdot\frac{\lambda}{\sqrt{1+\lambda^2}} \tag{3.5}$$

Use Mill's ratio to rewrite this as:

$$\gamma\cdot M(\gamma)=\lambda/\sqrt{1+\lambda^2} \tag{3.6}$$

Through differentiation and noting that $M'(\gamma)=\gamma\cdot M(\gamma)-1$ we obtain the relationship $\gamma'\cdot\left\{(1+\gamma^2)M(\gamma)-1\right\}=\left(1+\lambda^2\right)^{-3/2}$, which by (3.6) implies that $\gamma=\gamma(\lambda)$ solves the differential equation:

$$\gamma'(\lambda)=\frac{\left(1-\gamma^2M(\gamma)^2\right)^{3/2}}{(1+\gamma^2)M(\gamma)-1} \tag{3.7}$$

As a sufficient condition, which implies a double uniform bound (2.5), we show that $\frac{\partial}{\partial\lambda}\Delta(x,\lambda)<0,\ \frac{\partial}{\partial x}\Delta(x,\lambda)<0$ for all sufficiently high values of $x,\ \lambda>0$. Indeed, since $\Delta(x,\lambda)$ is strictly monotonic decreasing in both arguments, if suffices to determine thresholds $x_0,\ \lambda_0$, and for given absolute error $\delta>0$ the value $\lambda=\lambda_\delta\geq\lambda_0$, if it exists, that solves the signed difference equation $\Delta(x_0,\lambda)=\delta$, then automatically $0\leq\Delta(x_0,\lambda)\leq\delta$ for all $x\geq x_0,\ \lambda\geq\lambda_\delta\geq\lambda_0$. To derive this key result, we use (3.4) and note first that:

$$\frac{\partial}{\partial\lambda}\overline{G}_{\lambda}(x)=\int_x^{\infty}2t\varphi(t)\varphi(\lambda t)dt,\quad \frac{\partial}{\partial\gamma}\overline{H}_{\gamma}(x)=\int_x^{\infty}t\varphi(t)e^{-\gamma t}dt \tag{3.8}$$

Inserted into (3.4) we obtain the integral expression:

$$\frac{\partial}{\partial\lambda}\Delta(x,\lambda)=2\cdot\int_x^{\infty}t\varphi(t)\left\{\varphi(\lambda t)-\frac{1}{2}\gamma'(\lambda)e^{-\gamma(\lambda)t}\right\}dt \tag{3.9}$$

Now, we certainly have $\frac{\partial}{\partial\lambda}\Delta(x,\lambda)<0$ provided the curly bracket is negative for all

$t > x_0$ for some appropriate threshold x_0, which is equivalent with the inequality:

$$\chi(\lambda,t) = \frac{1}{2}\left(\lambda^2 t^2 - 2\gamma(\lambda)t + \ln\left\{\tfrac{\pi}{2}\gamma'(\lambda)^2\right\}\right) > 0, \quad \forall\, t > x_0 \tag{3.10}$$

Let $D(\lambda) = \gamma(\lambda)^2 - \lambda^2 \cdot \ln\left\{\frac{\pi}{2}\gamma'(\lambda)^2\right\}$ be the discriminant of the quadratic form in (3.10). If $D(\lambda) < 0$ the inequality $\chi(\lambda,t) > 0$ holds $\forall\, t > 0, \forall\, \lambda > 0$. Otherwise, the inequality $\chi(\lambda,t) > 0$ holds $\forall\, t > x_0 = \max_{\lambda>0}\{t_0(\lambda)\}$, where $t_0(\lambda)$ denotes the greatest zero of the quadratic equation $\chi(\lambda,t) = 0$. By (3.5)-(3.6) and since $\gamma = \gamma(\lambda)$ is uniquely defined, we can express $t_0(\lambda)$ as function of γ only, namely:

$$\begin{aligned} t_0(\lambda) &= t_0(\gamma) = \left(\frac{1-\gamma^2 M(\gamma)^2}{\gamma^2 M(\gamma)^2}\right) \cdot \left\{\gamma + \sqrt{D(\gamma)}\right\}, \\ D(\gamma) &= \gamma^2 - \left(\frac{\gamma^2 M(\gamma)^2}{1-\gamma^2 M(\gamma)^2}\right)\left(\ln\left\{\tfrac{\pi}{2}\right\} + 3\ln\left\{1-\gamma^2 M(\gamma)^2\right\} - 2\ln\left\{(1+\gamma^2)M(\gamma) - 1\right\}\right). \end{aligned} \tag{3.11}$$

A graphical numerical analysis shows that:

$x_0 = \max_{D(\gamma)\ge 0}\{t_0(\gamma)\} = t_0(\gamma^*) = 2.43527, \gamma^* = 1.40625$.

It remains to show that $\frac{\partial}{\partial x}\Delta(x,\lambda) < 0$ for all $x \ge x_0$ provided $\lambda > 0$ is sufficiently high.

A partial differentiation yields $\frac{\partial}{\partial \lambda}\Delta(x,\lambda) = h_\gamma(x) - g_\lambda(x) = 2\varphi(x)\left\{\overline{\Phi}(\lambda x) - \frac{1}{2}e^{-\frac{1}{2}\gamma x}\right\}$ for all $x > 0$. A sufficient condition for $\frac{\partial}{\partial \lambda}\Delta(x,\lambda) < 0$, is the inequality $\overline{\Phi}(\lambda x) = M(\lambda x)\varphi(\lambda x) < \frac{1}{2}e^{-\frac{1}{2}\gamma x}$. Since $M(x) \le x^{-1}$ by the inequality of [9], [5], and $x \ge x_0 > 1$, it suffices that $\varphi(\lambda x) < \frac{1}{2}\lambda e^{-\frac{1}{2}\gamma x}, \forall\, x \ge x_0$, which is equivalent to the quadratic inequality:

$$q(\lambda,x) = \frac{1}{2}\left(\lambda^2 x^2 - 2\gamma(\lambda)t + 2\ln\left\{\sqrt{\tfrac{\pi}{2}}\lambda\right\}\right) > 0, \forall\, x \ge x_0 \tag{3.12}$$

Now, we argue as above. Let $D(\lambda) = \gamma(\lambda)^2 - \lambda^2 \cdot \ln\left\{\frac{\pi}{2}\gamma'(\lambda)^2\right\}$ be the discriminant of the quadratic form in (3.12). If $D(\lambda) < 0$ the inequality (3.12) will hold for all $\lambda > 0$. Otherwise, the inequality (3.12) is fulfilled provided there exists $\lambda_0 > 0$ such that $x_0 = \max_{\lambda \ge \lambda_0}\{z_0(\lambda)\}$, where $z_0(\lambda)$ denotes the greatest zero of the quadratic equation $q(\lambda,x) = 0$. Similarly to (3.11), we can express $z_0(\lambda)$ as function of γ only, namely:

$$z_0(\lambda) = z_0(\gamma) = \left(\frac{1-\gamma^2 M(\gamma)^2}{\gamma^2 M(\gamma)^2} \right) \cdot \left\{ \gamma + \sqrt{D(\gamma)} \right\},$$
$$D(\gamma) = \gamma^2 - 2\left(\frac{\gamma^2 M(\gamma)^2}{1-\gamma^2 M(\gamma)^2} \right) \left(\tfrac{1}{2}\ln\left\{\tfrac{\pi}{2}\right\} + \ln\{\gamma M(\gamma)\} - \tfrac{1}{2}\ln\left\{1-\gamma^2 M(\gamma)^2\right\} \right). \tag{3.13}$$

Numerical analysis shows that $\max_{D(\gamma)\geq 0 \wedge \gamma \geq \gamma_0} \{z_0(\gamma)\} = z_0(\gamma_0) = x_0 = 2.43527$ for $\gamma_0 = 1.0527$, which yields $\lambda_0 = \gamma_0 M(\gamma_0) / \sqrt{1-\gamma_0^2 M(\gamma_0)^2} = 0.90639$ by the relationship (3.6). The following result has been shown.

Theorem 3.1. Given are the skew-normal and the skew-normal-Laplace distribution functions $G_\lambda(x), H_{\phi(\lambda)}(x)$ with equal mean. Then, for given $\delta > 0$, the signed tail difference satisfies the double uniform bound:

$$0 \leq \Delta(x,\lambda) = \overline{G}_\lambda(x) - \overline{H}_{\phi(\lambda)}(x) \leq \delta, \quad \forall\, x \geq x_0 = 2.43527, \forall\, \lambda \geq \lambda_\delta, \tag{3.14}$$

provided the signed difference equation $\Delta(x_0,\lambda) = \delta$ has a solution $\lambda = \lambda_\delta \geq \lambda_0 = 0.90639$.

This result implies the following normal approximation to Owen's T function.

Corollary 3.1. Given $\delta > 0$ and the definitions of Theorem 3.1, Owen's T function satisfies the double uniform normal approximation:

$$\tfrac{1}{2} A(x,\gamma) \leq T(x,\lambda) \leq \tfrac{1}{2}\{\delta + A(x,\gamma)\}, \quad \forall\, x \geq x_0, \forall\, \lambda \geq \lambda_\delta,$$
$$A(x,\gamma) = \overline{\Phi}(x) - e^{\frac{1}{2}\gamma^2} \cdot \overline{\Phi}(x+\gamma), \quad \gamma \cdot M(\gamma) = \lambda / \sqrt{1+\lambda^2}. \tag{3.15}$$

Proof. This follows from Theorem 3.1 using the representations (3.1), (3.3) and the fact that $\gamma = \phi(\lambda)$ solves the equation (3.6).

Since Owen's T function has a probabilistic meaning, and noting that $A(x,\gamma) = \overline{H}_\gamma(x) - \overline{\Phi}(x)$, we obtain the following probabilistic interpretation of Corollary 3.1:

$$\tfrac{1}{2}\{P(SNL(\gamma) > x) - P(N > x)\} \leq P(N_1 > x, 0 < N_2 < \lambda N_2)$$
$$\leq \tfrac{1}{2}\{\delta + P(SNL(\gamma) > x) - P(N > x)\}, \quad \forall\, x \geq x_0, \forall\, \lambda \geq \lambda_\delta, \tag{3.16}$$

where N_1, N_2, N are independent standard normal random variables and $SNL(\gamma)$ is a skew-normal-Laplace random variable. In other words, as $x \to \infty$ the following asymptotic equivalence holds:

$$P(N_1 > x, 0 < N_2 < \lambda N_2) \sim \tfrac{1}{2}\{P(SNL(\gamma) > x) - P(N > x)\}, \quad \forall\, \lambda \geq \lambda_\delta. \tag{3.17}$$

[17] mention that some approximations to Owen's T function cannot be used for accurate calculations [19], [8], and especially emphasize their inaccuracy when $x \to \infty$ [6]. Corollary 3.1 yields a useful accurate analytical asymptotic approximation, which in

contrast to all previous numerical approximations, has the advantage of a stochastic interpretation.
From a numerical elementary perspective it is possible to simplify further the obtained approximation using the exponential function only. For this purpose we propose to use the simple analytical approximation by [11] of the standard normal tail probability function given by:

$$\overline{\Phi}(x) \approx \frac{1}{2}\exp\left(-\alpha x - \beta x^2\right), \quad x \geq 0, \quad \alpha = 0.717, \quad \beta = 0.416. \tag{3.18}$$

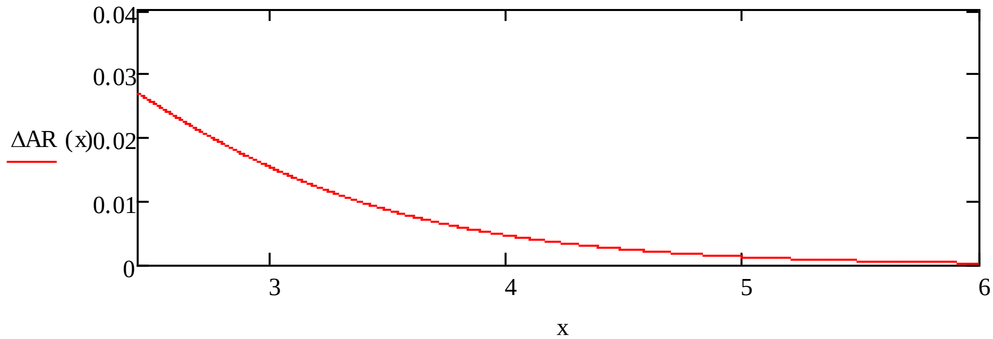
Corollary 3.2. Given the definitions of Theorem 3.1, Lin's exponential approximation $E(x,\gamma)$ to the signed difference $A(x,\gamma) = \overline{H}_\gamma(x) - \overline{\Phi}(x)$ yields the exponential approximation:

$$\begin{aligned} &T(x,\lambda) \sim \tfrac{1}{2}E(x,\gamma) = \tfrac{1}{4}\exp\left\{-\alpha x - \beta x^2\right\}\cdot\left(1-\exp\left\{-(\alpha+2\beta x)\gamma + \tfrac{1}{2}(1-2\beta)\gamma^2\right\}\right), \\ &\gamma \cdot M(\gamma) = \lambda / \sqrt{1+\lambda^2}, \quad \forall\, \lambda \geq \lambda_\delta, \end{aligned} \tag{3.19}$$

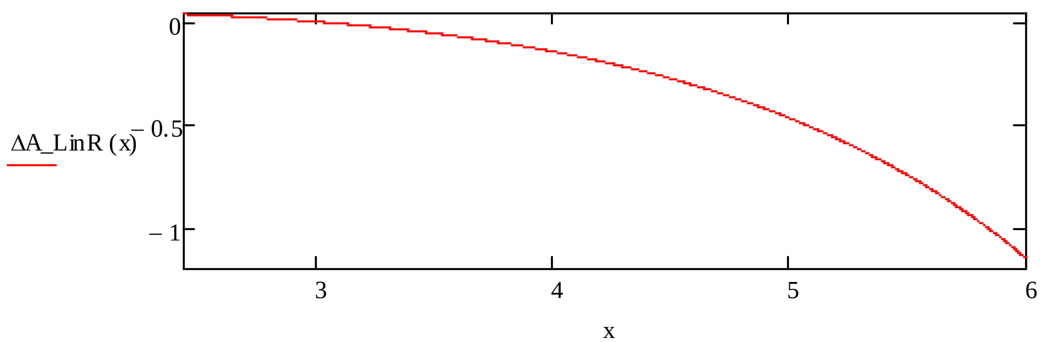
with fixed $\alpha = 0.717$, $\beta = 0.416$.
Table 3.1 illustrates numerically the obtained elementary approximations. We note that the absolute accuracy of the normal approximation over the range $\{x \geq x_0,\, \lambda \geq \lambda_\delta\}$ is worst at $x = x_0, \lambda = \lambda_\delta$ for the given level of accuracy $\delta > 0$. By increasing $x \geq x_0$ the relative accuracy of the normal approximation increases very rapidly as illustrated in the Graph 3.1. In contrast, the relative accuracy of the exponential approximation remains small over a finite interval only, for example $x \in [x_0, 3]$ in Graph 3.2. Therefore, the exponential approximation is not a useful asymptotic approximation. Owen's T function is calculated using numerical integration (as implemented in MATHCAD for example).

Table 3.1: Double uniform approximation bounds to Owen's T function

$\frac{1}{2}\delta$	λ_δ	$\gamma_\delta = \gamma(\lambda_\delta)$	$T(x_0,\lambda_\delta)\cdot 10^3$	$\frac{1}{2}A(x_0,\lambda_\delta)\cdot 10^3$	$\frac{1}{2}E(x_0,\lambda_\delta)\cdot 10^3$
10^{-4}	1.06400	1.27465	3.70338	3.60338	3.57126
10^{-5}	1.70203	2.19954	3.72014	3.71014	3.68645
10^{-6}	2.30404	3.07979	3.72017	3.71917	3.69800
10^{-7}	2.91364	3.96712	3.72017	3.72007	3.69950
10^{-8}	3.53131	4.86111	3.72017	3.72016	3.69972



Graph 3.1: Relative signed error $1-\frac{1}{2}A(x,\gamma_\delta)/T(x,\lambda_\delta)$ for $x \ge x_0$ by fixed $\lambda_\delta = 1.064$



Graph 3.2: Relative signed error $1-\frac{1}{2}E(x,\gamma_\delta)/T(x,\lambda_\delta)$ for $x \ge x_0$ by fixed $\lambda_\delta = 1.064$

4 Application to the Bivariate Normal Distribution

Owen's T function has originally been introduced to simplify the calculation of bivariate normal probabilities. The bivariate standard normal distribution with correlation coefficient ρ is defined and denoted by:

$$\Phi_2(x,y;\rho) = \int_{-\infty}^{x}\int_{-\infty}^{y} \varphi_2(u,v;\rho)dudv, \quad -\infty < x,y < \infty, \quad -1 \le \rho \le 1,$$
$$\varphi_2(x,y;\rho) = \frac{1}{\sqrt{2\pi(1-\rho^2)}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(x^2 - 2\rho xy + y^2\right)\right\}. \tag{4.1}$$

Recall the skew-normal distribution (3.1), which in suggestive changed notation reads:

$$\Phi_s(x,\lambda) = \Phi(x) - 2\cdot T(x,\lambda). \tag{4.2}$$

In terms of (4.2) the formula (2.1) in [16] can be rewritten as:

$$\Phi_2(x,y;\rho)=\tfrac{1}{2}\Phi_s\left(x,\tfrac{y-\rho x}{x\sqrt{1-\rho^2}}\right)+\tfrac{1}{2}\Phi_s\left(y,\tfrac{x-\rho y}{y\sqrt{1-\rho^2}}\right), \\ -\tfrac{1}{2}\left(1-1\{xy>0\vee(xy=0\wedge x+y\geq 0\}\right) \quad (4.3)$$

with $1\{\cdot\}$ the indicator function. Since $\Phi_2(x,0;\rho)=\tfrac{1}{2}\Phi_s\left(x,\tfrac{-\rho}{\sqrt{1-\rho^2}}\right)$ the formula (4.3) is equivalent to the reduction formula:

$$\Phi_2(x,y;\rho)=\Phi_2(x,0;\operatorname{sgn}(x)\cdot\tfrac{\rho x-y}{\sqrt{x^2-2\rho xy+y^2}})+\Phi_2(y,0;\operatorname{sgn}(y)\cdot\tfrac{\rho y-x}{\sqrt{x^2-2\rho xy+y^2}}) \\ -\tfrac{1}{2}\left(1-1\{xy>0\vee(xy=0\wedge x+y\geq 0\}\right), \quad (4.4)$$

whose special case $xy\neq 0$ is given in [22]. It is also useful to note that (4.4) is equivalent to the normal copula formula (3.16) in [12]. Usually, the numerical implementation of the bivariate normal distribution is done with the tetrachoric series, an infinite bivariate expansion based on the Hermite polynomials (consult [10] for a comprehensive review). However, this expansion converges only slightly faster than a geometric series with quotient ρ, which is not very practical to use when ρ is large in absolute value. The author [22] improves on this numerical evaluation and derives for $\Phi_2(x,0;\rho)$ another series that converges approximately as a geometric series with quotient $1-\rho^2$.

Alternatively to the latter, the use of Theorem 3.1 yields a simple normal analytical approximation (via the skew-normal Laplace), which is valid for sufficiently large values of the arguments.

Theorem 4.1. Assume that $x\geq x_0=2.43527$ and $1>|\rho|\geq\lambda_\delta/\sqrt{1+\lambda_\delta^2}$, where $\Delta(x_0,\lambda_\delta)=\delta$ and $\lambda_\delta\geq\lambda_0=0.90639$ as in Theorem 3.1. Then, the bivariate standard normal distribution satisfies the following inequalities:

$$\overline{H}_\gamma(-x)\leq\Phi_2(x,0;\rho)\leq\delta+\overline{H}_\gamma(-x),\quad \rho>0, \\ H_\gamma(x)-\delta\leq\Phi_2(x,0;\rho)\leq H_\gamma(x),\quad \rho<0, \quad (4.5)$$

where $H_\gamma(x)$ denotes the skew-normal-Laplace distribution (3.3) with parameter determined by the equation $\gamma\cdot M(\gamma)=|\rho|$.

Proof. In the notations of Section 3 set the parameter of the skew-normal equal to $\lambda=|\rho|/\sqrt{1-\rho^2}$.

We have $\Phi_2(x,0;\rho)=G_{-\lambda}(x)=\overline{G}_\lambda(-x)$ if $\rho>0$ and $\Phi_2(x,0;\rho)=G_\lambda(x)$ if $\rho<0$. In the first case, we use the fact that the difference $G_\lambda(x)-H_\gamma(x)$ is an even function ([21], p.663) to see that $\overline{G}_\lambda(-x)-\overline{H}_\gamma(-x)=\overline{G}_\lambda(x)-\overline{H}_\gamma(x)$. The first inequality in (4.5) follows from (3.12) by noting that the condition $\lambda\geq\lambda_\delta$ is equivalent to

$\rho \geq \lambda_\delta / \sqrt{1+\lambda_\delta^2}$. If $\rho < 0$ then $G_\lambda(x) - H_\gamma(x) = \overline{H}_\gamma(x) - \overline{G}_\lambda(x) = -\Delta(x,\lambda)$. The second inequality in (4.5) follows again from (3.12).

Finally, we note that Owen's T function has further applications in mathematical statistics, including non-central t-probabilities (e.g. [7], [23]) and multivariate normal probabilities [18] among others.

References

[1] M. Ali, M. Pal and J. Woo, Skewed reflected distributions generated by reflected gamma kernel, *Pakistan J. Statist.*, **24**(1), (2008), 77-86.

[2] M. Ali, M. Pal and J. Woo, Skewed reflected distributions generated by the Laplace kernel, *Austrian J. Statist.*, **38**(1), (2009), 45-58.

[3] A. Azzalini, A class of distributions which includes the normal ones, *Scand. J. Statist.*, **12**, (1985), 171-178.

[4] A. Azzalini, Further results on a class of distributions which includes the normal ones, *Statistica (Bologna)*, **46**(2), (1986), 199-208.

[5] Z.W. Birnbaum, An inequality for Mill's ratio, *Ann. Math. Statist.*, **13**, (1942), 245-246.

[6] B.E. Cooper, Algorithm AS4, An auxiliary function for distribution integrals, *Appl. Statist.*, **17**, (1968), 190-192, Corrigenda, **18**, (1969), 118, **19** (1970), 204.

[7] B.E. Cooper, Algorithm AS5, The integral of non-central t-distribution, *Appl. Statist.*, **17**, (1968), 193-194.

[8] T.G. Donelly, Algorithm 462, Bivariate normal distribution, *Comm. Assoc. Comput. Mach*, **16**, (1973), 638.

[9] R.D. Gordon, Values of Mill's ratio of area to bounding ordinate of the normal probability integral for large values of the argument, *Ann. Math. Statist.*, **12**, (1941), 364-366.

[10] S.S. Gupta, Probability integrals of multivariate normal and multivariate t, *Ann. Math. Statist.*, **34**, (1963), 792-828.

[11] J.-T. Lin, Approximating the normal tail probability and its inverse for use on a pocket calculator, *Appl. Statist.*, **38**, (1989), 69-70.

[12] C. Meyer, The bivariate normal copula, *Preprint*, (2009), URL: arXiv:0912.2816v1 [math.PR]

[13] S. Nadarajah and S. Kotz, Skewed distributions generated by the normal kernel, *Statist. Probab. Lett.*, **65**, (2003), 269-277.

[14] S. Nadarajah and S. Kotz, Skewed distributions generated by the Cauchy kernel, *Braz. J. Probab. Statist.*, **19**(1), (2005), 39-51.

[15] A. O'Hagan and T. Leonard, Bayes estimation subject to uncertainty about parameter constraints, *Biometrika*, **63**(1), (1976), 201-203.

[16] D.B. Owen, Tables for computing bivariate normal probabilities, *Ann. Math. Statist.*, **27**, (1956), 1075-1090.

[17] M. Patefield and D. Tandy, Fast and accurate calculation of Owen's T function, *J. Statist. Software*, **5**(5), (2000), 1-25.

[18] M.H. Schervish, Multivariate normal probabilities with error bound, *Appl. Statist.*, **33**, (1984), 81-94.

[19] R.R. Sowden and J.R. Ashford, Computation of the bivariate normal integral, *Appl. Statist.*, **18**, (1969), 169-180.

[20] D. Umbach, Some moment relationships for skew-symmetric distributions, *Statist. Probab. Lett.*, **76**, (2006), 507-512.

[21] D. Umbach, The effect of the skewing distribution on skew-symmetric families, *Soochow J. Math.*, **33**(4), (2007), 657-668.

[22] O.F. Vasicek, A series expansion for the bivariate normal integral, *The J. of Comput. Finance*, **1**, (1998), 5-10.

[23] R.A. Wijsman, New algorithms for the function T(h,a) of Owen, with application to bivariate normal and noncentral t-probabilities, *Comput. Statist. Data Anal.*, **21**(3), (1996), 261-271.