Tilted Normal Distribution and Its Survival Properties

Sudhansu S. Maiti[1*] and Mithu Dey[2]

[1]*Visva-Bharati University* and [2]*Asansol Engineering College*

Abstract: To analyze skewed data, skew normal distribution is proposed by Azzalini (1985). For practical problems of estimating the skewness parameter of this distribution, Gupta and Gupta (2008) suggested power normal distribution as an alternative. We search for another alternative, named tilted normal distribution following the approach of Marshall and Olkin (1997) to add a positive parameter to a general survival function and taking survival function is of normal form. We have found out different properties of this distribution. Maximum likelihood estimate of parameters of this distribution have been found out. Comparison of tilted normal distribution with skew normal and power normal distribution have been made.

Key words: Akaike information criterion, compound distribution, failure rate, Kolmogorov discrepancy measure, maximum likelihood.

1. Introduction

The skew normal distribution, proposed by Azzalani (1985), can be a suitable model for the analysis of data exhibiting a unimodal density having some skewness present, a structure often occurring in data analysis. The proposed distribution is a generalization of the standard normal distribution and the probability density function (pdf) is given by

$$\phi(z; \lambda) = 2\phi(z)\Phi(\lambda z), \quad -\infty < z < \infty, \tag{1.1}$$

where $\phi(x)$ and $\Phi(x)$ denote the $N(0, 1)$ density and distribution function respectively. The parameter λ regulates the skewness and $\lambda = 0$ corresponds to the standard normal case. The density given by (1.1) enjoys a number of formal properties which resemble those of the normal distribution, for example, if X has the pdf, given by (1.1), then X^2 has a chi-square distribution with one degree of freedom. That is, all even moments of X are exactly the same as the

---

*Corresponding author.

corresponding even moments of the standard normal distribution. For more information, see Azzalini and Capitanio (1999) and Genton (2004). A motivation of the skew normal distribution has been elegantly exhibited by Arnold *et al.* (1993). This model can naturally arise in applications as hidden function and/or selective reporting model, see Arnold and Beaver (2002).

Gupta and Gupta (2008) observed that the estimation of the skewness parameter of model (1.1) is problematic when the sample size is not large enough. Monti (2003) noted that the estimate is $\widehat{\lambda} = \pm\infty$, even when the data are generated by a model with finite λ. In the case of a skew normal model the moment estimator of λ is the solution of

$$\sqrt{\frac{2}{\pi}} \frac{\lambda}{\sqrt{1+\lambda^2}} = \bar{X}. \tag{1.2}$$

Therefore, the solution exists if and only if $|\bar{X}| < \sqrt{2/\pi}$. The exact distribution of $\bar{X}$ is not known and it is not easy to establish. Gupta and Gupta (2008) estimated the value of $\mathrm{P}(|\bar{X}| > \sqrt{2/\pi})$ based on simulation data for different values of n and λ. They have shown that as the sample size increases, the probability of the feasibility of the moment estimator increases. On the other hand, as the skewness parameter increases, the chances of obtaining the moment estimator decreases.

To find out the maximum likelihood estimator(MLE) of λ based on a random sample $X_1, X_2, \cdots, X_n$, the likelihood equation is

$$\sum_{i=0}^{n} \frac{X_i \phi(\lambda X_i)}{\Phi(\lambda X_i)} = 0. \tag{1.3}$$

It is clear that if all X_i have the same sign then there is no solution of (1.3), see also Liseo (1990) and Azzalini and Capitanio (1999). The probability that $X > 0$, for $\lambda > 0$, is an increasing function of λ. When the absolute value of the skewness parameter is large, the probability that all the observations have the same sign is large. Therefore, for small to moderate sample sizes, the MLE may not be accurate enough for a practical use. The Centered Parametrization (CP) can solve these difficulties to some extent. Azzalini (1985) reparametrized the problem by writing $X = \mu + \sigma Z^0$, where $Z^0 = (Z - \mu_z)/\sigma_z$, $\sigma_z = (1 - \mu_z^2)^{\frac{1}{2}}$, $\mu_z = \sqrt{2/\pi} \cdot (\lambda/\sqrt{1+\lambda^2})$ and considering the centred parameters CP$= (\mu, \sigma, \gamma_1)$ instead of the direct parameters DP$= (\mu, \sigma, \lambda)$. Here γ_1 is the usual univariate index of skewness, taken with same sign as that of λ [see for details, Azzalini and Capitanio (1999)]. The set of S-PLUS routines developed for the computation are freely available at the following World Wide Web address:

http://azzalini.stat.unipd.it/SN/.

For the above reason, Gupta and Gupta (2008) considered another skewed model for which normal distribution is a special case. They proposed the power normal model whose distribution function and density function are given by

$$F_2(x; \alpha) = [\Phi(x)]^\alpha, \quad -\infty < x < \infty, \quad \alpha > 0, \tag{1.4}$$

and

$$f_2(x; \alpha) = \alpha [\Phi(x)]^{\alpha - 1} \phi(x), \quad -\infty < x < \infty, \quad \alpha > 0. \tag{1.5}$$

This is basically an exponentiated family of distributions with baseline distribution as normal. This is nothing but a system of adding one or more parameters to the existing normal distribution. Such an addition of parameters makes the resulting distribution richer and more flexible for modeling data. The power normal distribution has several nice properties. For $\alpha = 1$ it reduces to the standard normal distribution. This distribution is a unimodal density which is skewed to the right if $\alpha > 1$ and to the left if $0 < \alpha < 1$. Gupta and Gupta (2008) considered α as skewness parameter of the distribution.

But for $0 < \alpha < 1$, skewness to the left is not so clear as is to be seen from the Figure 1 and modeling of left skewed data set will be misfit. This motivates us to think for another skewed distribution for which normal distribution is a specified case and at the same time it will be fit for modeling both kind of skewed data. We take Marshall and Olkin (1997) extended model to add a positive parameter to a general survival function.
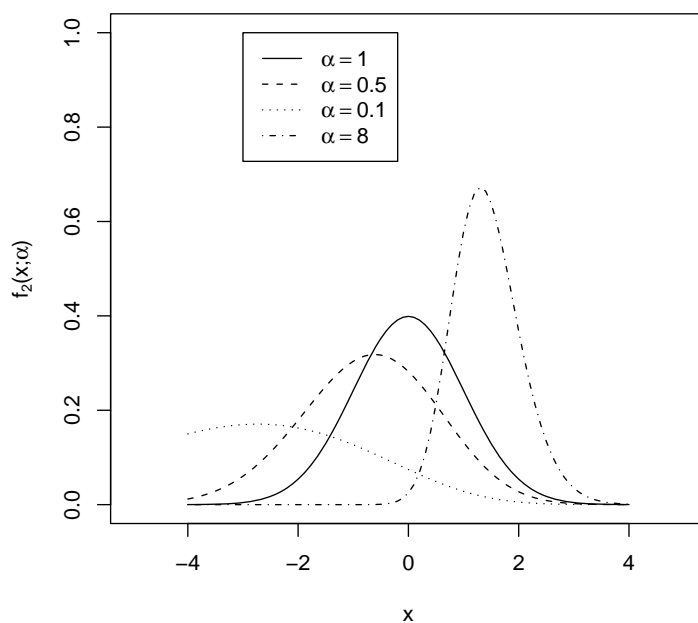

Figure 1: Plot of power normal distribution for different values of α

Let X be arandom variable with survival function and the probability density function as

$$\overline{F}_3(x; \beta) = \frac{\beta \overline{F}_0(x)}{1 - (1 - \beta)\overline{F}_0(x)}, \quad -\infty < x < \infty, \quad \beta > 0, \qquad (1.6)$$

and

$$f_3(x; \beta) = \frac{\beta f_0(x)}{[1 - (1 - \beta)\overline{F}_0(x)]^2}, \quad -\infty < x < \infty, \quad \beta > 0, \qquad (1.7)$$

respectively. Here $F_0(x)$ is the baseline distribution function and $f_0(x)$ is the corresponding probability density function. Taking $F_0(x) = \Phi(x)$ and $f_0(x) = \phi(x)$, we get the tilted normal (standard) distribution as

$$f_3(x; \beta) = \frac{\beta \phi(x)}{[1 - (1 - \beta)\{1 - \Phi(x)\}]^2}, \quad -\infty < x < \infty, \quad \beta > 0. \qquad (1.8)$$

Here $\beta = 1$ indicates a standard normal density function. The tilted normal density is a unimodal density which is skewed to the left if $\beta > 1$ and to the right if $0 < \beta < 1$. Figure 2 displays a few pdf graphs for various values of β.
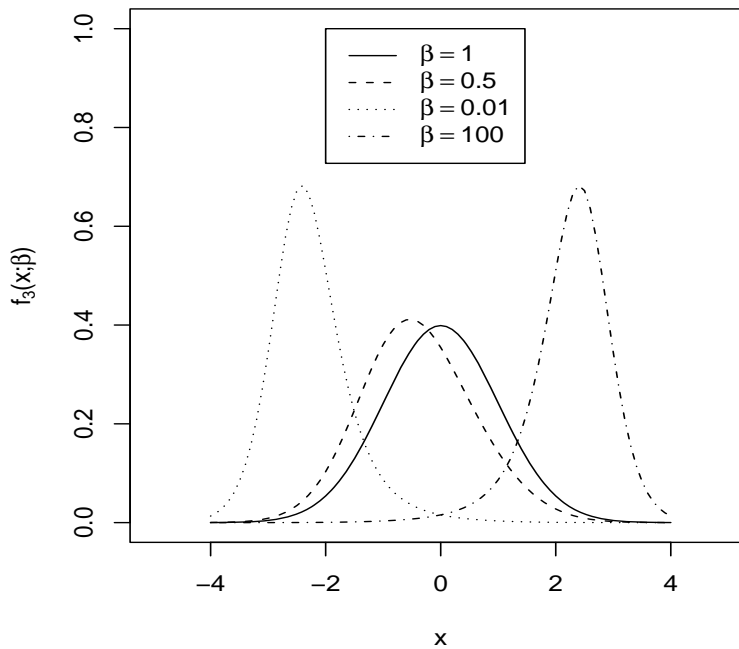

Figure 2: Plot of tilted normal distribution for different values of β

If the normal density has mean μ and variance σ^2, then the form of the tilted normal distribution will be

$$f_3(x;\beta,\mu,\sigma) = \frac{\frac{\beta}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)}{[1-(1-\beta)\{1-\Phi\left(\frac{x-\mu}{\sigma}\right)\}]^2}, \quad -\infty < x,\ \mu < \infty,\ \beta,\sigma > 0. \,(1.9)$$

We will denote this tilted normal distribution with parameters μ, σ and β as TN(μ,σ,β).

The paper is organized as follows. In Section 2, we have shown that the TN distribution can be obtained as a compound distribution with mixing exponential distribution. The basic structural properties of the proposed model including the reliability properties are presented in Section 3. In Section 4, we discuss the estimation of parameters of the TN distribution. Section 5 is devoted for studying closeness of skew normal, power normal and tilted normal distributions. A data set has been analyzed using the TN model in Section 6. Section 7 concludes.

2. Compounding

Let $\bar{G}(x|\theta), -\infty < x,\ \theta < \infty$, be a survival function (SF) of a random variable X given θ. Let θ be a random variable with pdf $m(\theta)$. Then a distribution with survival function

$$\bar{G}(x) = \int_{-\infty}^{\infty} \bar{G}(x|\theta)m(\theta)d\theta, \quad -\infty < x < \infty,$$

is called a compound distribution with mixing density $m(\theta)$. Compound distributions provide a tool to obtain new parametric families of distributions from existing ones. They represent heterogeneous models where populations' individuals have different risks.

The following theorem show that the TN distribution can be obtained as a compound distribution.

Theorem 2.1 Suppose that the conditional SF of a continuous random variable X given $\Theta = \theta$ is given by

$$\bar{G}(x|\theta) = \exp\left[-\theta\frac{\Phi\left(\frac{x-\mu}{\sigma}\right)}{1-\Phi\left(\frac{x-\mu}{\sigma}\right)}\right], \quad -\infty < x,\ \mu < \infty,\ \theta,\ \sigma > 0.$$

Let θ have an exponential distribution with pdf

$$m(\theta) = \beta e^{-\beta\theta}, \quad \theta,\ \beta > 0.$$

Then the compound distribution of X is the TN$(\mu,\ \sigma,\ \beta)$ distribution.

Proof: For $-\infty < x < \infty$, $-\infty < \mu < \infty$, $\beta > 0$, the SF of X is given by

$$\bar{G}(x) = \int_0^\infty \bar{G}(x|\theta)m(\theta)d\theta$$

$$= \beta \int_0^\infty e^{-\theta \frac{\Phi\left(\frac{x-\mu}{\sigma}\right)}{1-\Phi\left(\frac{x-\mu}{\sigma}\right)}} e^{-\beta\theta} d\theta$$

$$= \frac{\beta\left[1 - \Phi\left(\frac{x-\mu}{\sigma}\right)\right]}{1 - (1-\beta)\left[1 - \Phi\left(\frac{x-\mu}{\sigma}\right)\right]},$$

which is the SF of the TN$(\mu,\ \sigma,\ \beta)$ distribution.

For fixed $\theta > 0$, the distribution with survival function $\bar{G}(x|\theta)$ gives a class of non-standard distributions. As a result of Theorem 2.1 above, compounding any of the distributions belonging this class with an exponential mixing distribution for θ will lead to a certain form of TN distribution. This theorem also provides another way to interpret the TN distribution in addition to the geometric-extreme stability property given by Marshall and Olkin (1997).

3. Basic Structural and Survival Properties

The expressions for the mean, variance and skewness of the TN(μ, σ, β) model seem to be not available in compact form. For the proposed model, we have,

$$E\left(X^r\right) = \int_{-\infty}^\infty x^r \frac{\frac{\beta}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)}{[1 - (1-\beta)\{1 - \Phi\left(\frac{x-\mu}{\sigma}\right)\}]^2} dx.$$

Using this expression, with the help of computer, we can find measures of central tendency, dispersion and skewness index of the TN$(\mu,\ \sigma,\ \beta)$. We have presented measures of central tendency (mean, median, mode), and variance and skewness index, $\beta_1 = \mu_3/\mu_2^{\frac{3}{2}}$ (where μ_2 and μ_3 are second and third order central moments) for the TN$(0,\ 1,\ \beta)$ in Figure 3 and Figure 4 respectively. The median M of the TN$(0,1,\beta)$ distribution is given by

$$M = \Phi^{-1}\left(\frac{\beta}{1+\beta}\right),$$

which is the $\beta/(1+\beta)$th quantile of standard normal distribution. The mode is a solution of the equation

$$x = -\frac{2(1-\beta)\phi(x)}{1 - (1-\beta)[1 - \Phi(x)]}.$$

From Figure 3 and Figure 4, it is found that, as expected, median is less than mean for $\beta < 1$. Variance is increasing for $\beta < 1$ and is decreasing for $\beta > 1$, and skewness index is lying between -1 and 1 for the TN$(0,1,\beta)$ variate.
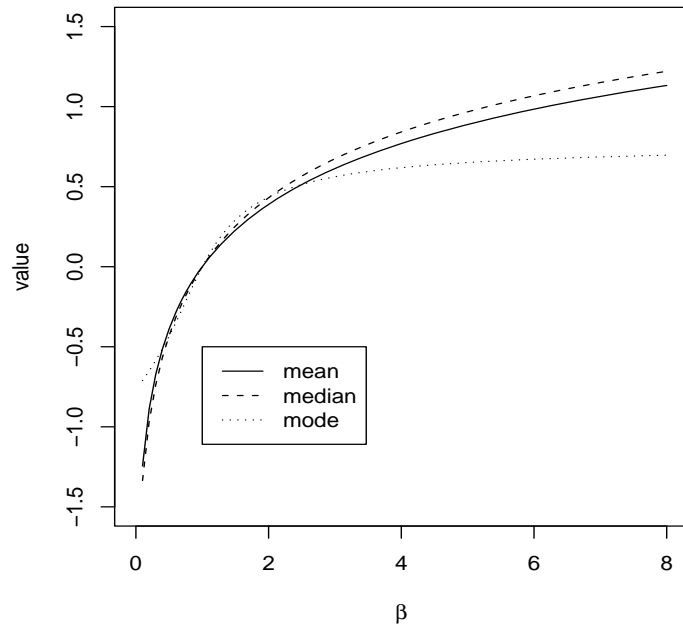
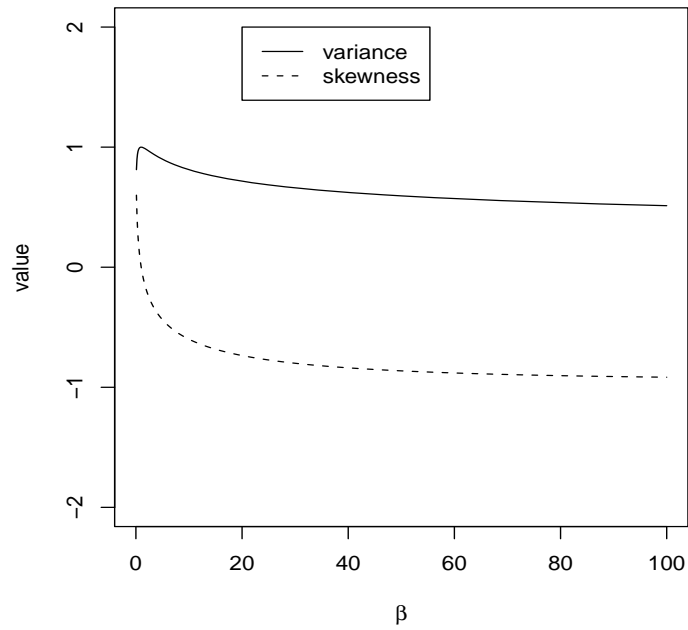Figure 3: Plot of mean, median and mode of tilted normal distribution for different values of β

Figure 4: Plot of variance and skewness index of tilted normal distribution for different values of β

The closed form expressions for basic structural properties are not available. Therefore, we now concentrate to get survival properties of the TN distribution. The failure rate function of X is given by

$$h_{TN}(x) = \frac{1}{1 - (1 - \beta)[1 - \Phi(x)]} \cdot \frac{\phi(x)}{1 - \Phi(x)}$$

$$= \frac{1}{1 - (1 - \beta)[1 - \Phi(x)]} \cdot h_N(x),$$

with the integrated failure rate function

$$H_{TN}(x) = -\ln \frac{\beta [1 - \Phi(x)]}{1 - (1 - \beta)[1 - \Phi(x)]}.$$

The reversed failure rate function of X is given by

$$r_{TN}(x) = \frac{\beta}{1 - (1 - \beta)[1 - \Phi(x)]} \cdot \frac{\phi(x)}{\Phi(x)}$$

$$= \frac{\beta}{1 - (1 - \beta)[1 - \Phi(x)]} \cdot r_N(x).$$

Here $h_N(x)$ and $r_N(x)$ are failure rate and reversed failure rate function of the $N(0, 1)$ distribution respectively. Figure 5 and Figure 6 show failure rates and reversed failure rates of the tilted normal distribution for different values of β respectively.
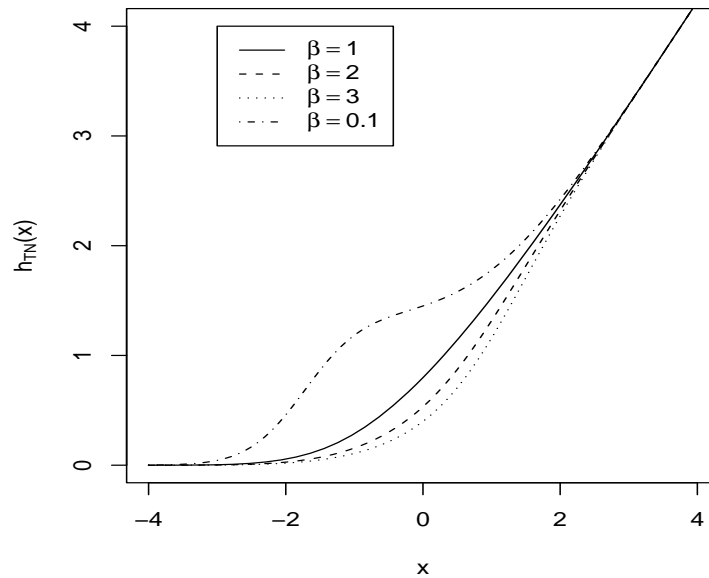

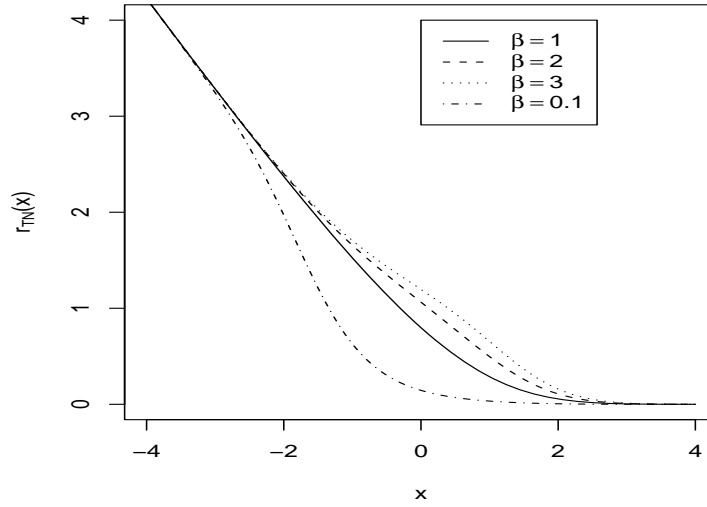Figure 5: Plot of failure rate of tilted normal distribution for different values of β

Figure 6: Plot of reversed failure rate of tilted normal distribution for different values of β

Theorem 3.1 If $X \sim TN(0, 1, \beta_1), Y \sim TN(0, 1, \beta_2)$, and $\beta_1 > \beta_2$, then $Y <_{LR} X \Rightarrow Y <_{HAZ} X \Rightarrow Y <_{ST} X$ and $Y <_{HAZ} X \Rightarrow Y <_{MRL} X$. Here $Y <_W X$ means Y is smaller than X in W ordering.

Proof: The logarithm of the likelihood ratio

$$v(x) = \ln \frac{f_3(x, \beta_1)}{f_3(x, \beta_2)}$$
$$= \ln \frac{\beta_1}{\beta_2} + 2 \left[\ln\{1 - (1 - \beta_2)(1 - \Phi(x))\} - \ln\{1 - (1 - \beta_1)(1 - \Phi(x))\} \right]$$

is an increasing function of x if $\beta_1 > \beta_2$, since

$$v'(x) = \frac{2\phi(x)(\beta_1 - \beta_2)}{[1 - (1 - \beta_1)(1 - \Phi(x))][1 - (1 - \beta_2)(1 - \Phi(x))]} > 0,$$

for all x. Therefore, the tilted normal distribution has the likelihood ratio ordering, which implies it has the failure rate ordering as well as the stochastic ordering and the mean residual life ordering, see Gupta and Kirmani (1987).

4. Maximum Likelihood Estimation

Let $X_i, i = 1, 2, \cdots, n$, be the ith observation and let Y_i be the corresponding censoring point. The X_i's and Y_i's are assumed independent. The X_i's are assumed to have the $TN(\mu, \sigma, \beta)$ distribution and the Y_i's are assumed to have a

non-informative distribution i.e. a distribution that does not involve the parameters μ, σ and β. One observes only the pair (T_i, δ_i), where $T_i = \min(X_i, Y_i)$ and $\delta_i = I(X_i \leq Y_i)$ is the censoring indicator.

Given the data $\{(t_i, \delta_i), i = 1, 2, \cdots, n\}$, the likelihood function is given by

$$
\begin{aligned}
L_n(\mu, \sigma, \beta) &= \prod_{i=1}^{n} \{f_3(t_i; \mu, \sigma, \beta)\}^{\delta_i} \{1 - F_3(t_i; \mu, \sigma, \beta)\}^{1-\delta_i} \\
&= \prod_{i=1}^{n} \left[\frac{\frac{1}{\sigma}\phi\left(\frac{t_i-\mu}{\sigma}\right)}{1 - \Phi\left(\frac{t_i-\mu}{\sigma}\right)} \cdot \frac{1}{1 - (1-\beta)\left\{1 - \Phi\left(\frac{t_i-\mu}{\sigma}\right)\right\}} \right]^{\delta_i} \\
&\quad \times \left[\frac{\beta\left\{1 - \Phi\left(\frac{t_i-\mu}{\sigma}\right)\right\}}{1 - (1-\beta)\left\{1 - \Phi\left(\frac{t_i-\mu}{\sigma}\right)\right\}} \right].
\end{aligned}
$$

Maximizing the likelihood function is equivalent to maximizing its logarithm. Hence the log-likelihood $l(\mu, \sigma, \beta) = \ln L_n(\mu, \sigma, \beta)$ is given by

$$
\begin{aligned}
l(\mu, \; \sigma, \; \beta) = &- \ln\sigma \sum_{i=1}^{n} \delta_i + n\ln\beta + \sum_{i=1}^{n} \delta_i \ln\phi\left(\frac{t_i-\mu}{\sigma}\right) \\
&+ \sum_{i=1}^{n} (1-\delta_i) \ln\left\{1 - \Phi\left(\frac{t_i-\mu}{\sigma}\right)\right\} \\
&- \sum_{i=1}^{n} (1+\delta_i) \ln\left[1 - (1-\beta)\left\{1 - \Phi\left(\frac{t_i-\mu}{\sigma}\right)\right\}\right].
\end{aligned}
$$

Maximum likelihood estimates (MLEs) of the parameters β, μ, σ are obtained by solving the non-linear equations $\partial l/\partial\beta = 0, \partial l/\partial\mu = 0, \partial l/\partial\sigma = 0$. Hence, the iterative solutions are

$$
\beta = \frac{n}{\sum_{i=1}^{n}\left[\frac{(1+\delta_i)\left\{1-\Phi\left(\frac{t_i-\mu}{\sigma}\right)\right\}}{1-(1-\beta)\left\{1-\Phi\left(\frac{t_i-\mu}{\sigma}\right)\right\}}\right]},
$$

$$
\mu = \frac{\sigma}{\sum_{i=1}^{n}\delta_i}\left[\frac{1}{\sigma}\sum_{i=1}^{n}\delta_i t_i + \sum_{i=1}^{n}\frac{(1-\delta_i)\phi\left(\frac{t_i-\mu}{\sigma}\right)}{1-\Phi\left(\frac{t_i-\mu}{\sigma}\right)} + \sum_{i=1}^{n}\frac{(1+\delta_i)(1-\beta)\phi\left(\frac{t_i-\mu}{\sigma}\right)}{1-(1-\beta)\left\{1-\Phi\left(\frac{t_i-\mu}{\sigma}\right)\right\}}\right],
$$

and

$$
\begin{aligned}
\sigma^2 = \frac{1}{\sum_{i=1}^{n}\delta_i}&\left[\sum_{i=1}^{n}\delta_i(t_i-\mu)^2 + \sigma\sum_{i=1}^{n}\frac{(1-\delta_i)(t_i-\mu)\phi\left(\frac{t_i-\mu}{\sigma}\right)}{1-\Phi\left(\frac{t_i-\mu}{\sigma}\right)}\right. \\
&\left.+ \sigma\sum_{i=1}^{n}\frac{(1+\delta_i)(t_i-\mu)(1-\beta)\phi\left(\frac{t_i-\mu}{\sigma}\right)}{1-(1-\beta)\left\{1-\Phi\left(\frac{t_i-\mu}{\sigma}\right)\right\}}\right],
\end{aligned}
$$

respectively. For the complete sample, $\delta_i = 1$, for all i, and for the type-I censored sample, $Y_i = t_0$, for all i.

The large-sample variance-covariance matrix of the MLEs is obtained by inverting the matrix of minus second order derivatives of the log-likelihood evaluated at the MLEs.

The likelihood ratio test will be used to test the null hypothesis $H_0 : \beta = 1$ (Normal distribution). When H_0 is true, the deviance test statistic $d_n = -2[l(\hat{\mu}, \hat{\sigma}, 1) - l(\hat{\mu}, \hat{\sigma}, \hat{\beta})]$, where $\hat{\mu}, \hat{\sigma}$ are the MLEs of μ, σ under $H_0 : \beta = 1$, has approximately a chi-square distribution with 1 degree of freedom. In addition, for model selection, we use the Akaike Information Criterion (AIC) defined as:

$$AIC = \text{-2 log-likelihood} + 2k,$$

where k is the number of parameters in the model and n is the sample size. For more details about the AIC, see Akaike (1969). The model with minimum AIC is the one that better fits the data.

5. Closeness of Skew Normal, Power Normal and Tilted Normal Distribution

It should be noted that the cumulative distribution function of the skew normal, power normal and tilted normal distributions are

$$F_1(x, \lambda) = 2 \int_{-\infty}^{x} \Phi(\lambda t)\phi(t)dt, \quad -\infty < \lambda < \infty,$$

$$F_2(x, \alpha) = \Phi^{\alpha}(x), \quad \alpha > 0,$$

and

$$F_3(x, \beta) = \frac{\Phi(x)}{1 - (1 - \beta)\{1 - \Phi(x)\}}, \quad \beta > 0,$$

respectively. $F_1(x, \lambda)$, $F_2(x, \alpha)$ and $F_3(x, \beta)$ are exactly equal to normal distribution when $\lambda = 0$, $\alpha = 1$ and $\beta = 1$. $F_1(x, \lambda)$ and $F_2(x, \alpha)$ are also exactly equal when $\lambda = 1$ and $\alpha = 2$. Since all the models can be used to model positively or negatively skewed data, it might be possible to approximate $F_1(x, \lambda)$ by $F_2(x, \alpha)$ or $F_3(x, \beta)$. To investigate this, we apply the following procedure:

1. Set a value of λ.

2. Find a and b such that $a = \max \left\{ x : 2 \int_{-\infty}^{x} \Phi(\lambda t)\phi(t)dt \leq 0.001 \right\}$ and
   $b = \min \left\{ x : 2 \int_{-\infty}^{x} \Phi(\lambda t)\phi(t)dt \geq 0.999 \right\}$.

3. Divide the range $[a, b]$ by 100 parts, say, $a = x_0 \leq x_1 \leq \cdots \leq x_{100} = b$, and set $y_i = F_1(x_i, \lambda)$; $i = 0, 1, 2, \cdots, 100$.

4. Find α by minimizing $S_1^2 = \sum_{i=1}^{100} [y_i - F_2(x_i, \alpha)]^2$, to fit the model $y \sim F_2(x, \alpha)$.

5. Find the Kolmogorov discrepancy measure $D_{12} = \max_{x_i} |F_2(x_i, \hat{\alpha}) - F_1(x_i, \lambda)|$.

6. Find β by minimizing $S_2^2 = \sum_{i=1}^{100} [y_i - F_3(x_i, \beta)]^2$, to fit the model $y \sim F_3(x, \beta)$.

7. Find the Kolmogorov discrepancy measure $D_{13} = \max_{x_i} |F_3(x_i, \hat{\beta}) - F_1(x_i, \lambda)|$.

In Table 1, we represent the values of $\hat{\alpha}$ and D_{12} corresponding to various values of λ, and the values of $\tilde{\beta}$ and D_{13} corresponding to various values of α (assuming $\hat{\alpha}$ as α in this case).

Table 1: Kolmogorov discrepancy measure for various values of λ and corresponding values of α and β

λ	$\hat{\alpha}$	D_{12}	$\hat{\beta}$	D_{13}
-3.0	0.523379	0.2088197	0.286948	0.1310713
-2.5	0.526644	0.1931135	0.292527	0.1180007
-2.0	0.534567	0.1723528	0.303881	0.1013205
-1.5	0.552582	0.1438773	0.328224	0.07972262
-1.0	0.595259	0.1037106	0.386520	0.05227650
-0.5	0.707068	0.0501869	0.547937	0.02133689
0.0	1.000000	0.0000000	1.000000	0.00000000
0.5	1.51708	0.01424926	1.82503	0.02133466
1.0	2.000000	0.0000000	2.58718	0.05227601
1.5	2.30154	0.01932597	3.04666	0.07973357
2.0	2.47086	0.03698127	3.29064	0.1013342
2.5	2.56769	0.05140222	3.41826	0.1179983
3.0	2.62544	0.06307444	3.48459	0.1310684

From the table it is clear that the discrepancy among the skew normal, power normal and tilted normal is minimal when $\lambda = 0$, $\alpha = 1$ and $\beta = 1$ and the three models coincide with the standard normal distribution and there is, obviously, no discrepancy among the three. The skew normal and power normal distributions are also identical when $\lambda = 1$ and $\alpha = 2$. The power normal distribution provides a better approximation to the skew normal when λ is positive as compared to the negative values of λ and the Kolmogorov discrepancy measure does not exhibit symmetric behaviour for values of λ moving away from zero in both directions. On the other hand, the tilted normal distribution approximates with equal efficiency for values of λ moving away from zero in both directions as the

Kolmogorov discrepancy measure exhibits symmetric behaviour. It provides a better approximation than the power normal to the skew normal when λ is negative. The relative Kolmogorov discrepancy of tilted normal and power normal is less for the positive values of λ as compared to that for the negative values of λ.

6. Data Analysis

In this section, we consider a data set to fit with the tilted normal distribution. The data related to failure time of bus motor is taken from Davis (1952).

We consider initial bus motor failures data of Davis (1952, p. 145). Data contain distance intervals (in thousands of miles) and corresponding observed number of failures and have been shown in Table 2. Histogram shows that the data set is negatively (left-) skewed with coefficient of skewness, $b_1 = m_3/\sqrt{m_2^3} = -0.3141923$ (where m_2 and m_3 are sample second and third order central moments). Davis fitted this data to the normal distribution with observed $\chi^2 = 9.93$ having p-value 0.08. We have fitted this data with the normal, the skew normal, the power normal and the tilted normal distributions. The results are summarized in Table 3. Histogram and fitted tilted normal, skew normal, power normal and normal curves to data have been shown in Figure 7.

Table 2: Initial bus motor failures

Distance Interval, Thousands of Miles	Observed Number of Failures
Less than 20	6
$20 - 40$	11
$40 - 60$	16
$60 - 80$	25
$80 - 100$	34
$100 - 120$	46
$120 - 140$	33
$140 - 160$	16
$160 - 180$	2
$180-$ up	2
Total	191

Table 3: Summarized results of fitting different distributions to data set of Davis (1952)

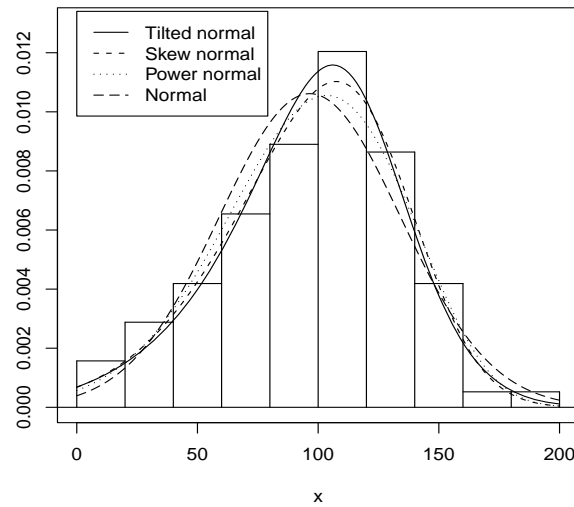| Distribution	$\hat{\mu}$	$\hat{\sigma}$	Estimate of tilt parameter	log-likelihood	AIC
Normal	96.5688	37.5829	-	-388.244	772.488
Power normal	142.48011	19.4888	$\hat{\alpha} = 0.156967$	-385.957	765.914
Skew normal	136.226843	55.04693	$\hat{\lambda} = -2.17259$	-385.068	764.136
Tilted normal	58.7097	40.1341	$\hat{\beta} = 5.59596$	-384.616	763.232

Figure 7: Plot of histogram and fitted tilted normal, skew normal, power normal and normal curves to data of Davis (1952)

Here for tilted normal fitting, the deviance test statistic $d_n = 7.256$ with p-value 0.0070665. These information justify better fitting of the tilted normal distribution to the data. Since the data are related to failure times, we have presented the estimated failure rate functions for the normal, the power normal, the skew normal and the tilted normal model in Figure 8. The estimated failure rate shows an increasing pattern.
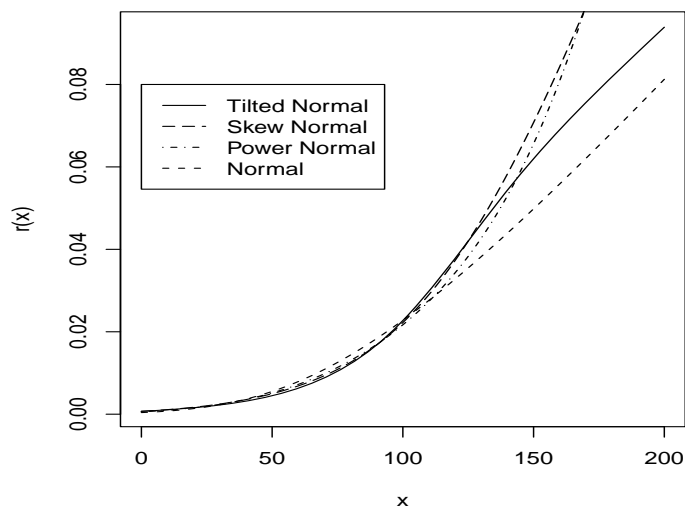

Figure 8: Plot of estimated failure rate functions of fitted tilted normal, skew normal, power normal and normal curves to data of Davis (1952)

7. Concluding Remarks

In this article, the tilted normal distribution has been studied which could be an alternative model to skew normal distribution for left skewed data. The practical problems in estimating the skewness parameter for small to moderate sample sizes of skew normal distribution has been discussed and to alleviate these problems, an alternative model called the 'Power normal model' has been proposed by Gupta and Gupta (2008). It has been observed that the 'Power normal model' is not a good fit for data skewed to the left. As an alternative this tilted normal distribution has been proposed. The normal distribution is widely and comfortably used by the practitioners as well as the theoreticians for analyzing symmetric data. Since skewed data could be analyzed by the tilted normal distribution that is obtained by adding a parameter, it is to be easy and comprehensive to the users. The structural and survival properties of this distribution have been studied and inference on parameters have also been mentioned. The appropriateness of fitting the tilted normal distribution has been established by analyzing a data set.

References

Akaike, H. (1969). Fitting autoregressive models for prediction. *Annals of the Institute of Statistical Mathematics* **21**, 243-247.

Arnold, B. C. and Beaver, R. J. (2002). Skewed multivariate models related to hidden truncation and/or selective reporting. *Test* **11**, 7-54.

Arnold, B. C., Beaver, R. J., Groeneveld, R. A. and Meeker, W. Q. (1993). The nontruncated marginal of a truncated bivariate normal distribution. *Psychometrika* **58**, 471-488.

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics* **12**, 171-178.

Azzalini, A. and Capitanio, A. (1999). Statistical applications of the multivariate skew-normal distribution. *Journal of the Royal Statistical Society, Series B* **61**, 579-602.

Davis, D. J. (1952). An analysis of some failure data. *Journal of the American Statistical Association* **47**, 113-150.

Genton, M. G. (2004). *Skew-Elliptical Distributions and Their Applications*: *A Journey Beyond Normality.* Chapman and Hall, New York.

Gupta, R. C. (2006). Variance residual life function in reliability studies. *Metron - International Journal of Statistics* **LXIV**, 343-355.

Gupta, R. C. and Kirmani, S. N. U. A. (1987). On the order relations between reliability measures of two distributions. *Stochastic Models* **3**, 149-156.

Gupta, R. D. and Gupta, R. C. (2008). Analyzing skewed data by power normal model. *Test* **17**, 197-210.

Liseo, B. (1990). La classe delle densità normali sghembe: aspetti inferenziali da un punto di vista bayesiano (in Italian). *Statistica* **L**, 59-70.

Marshall, A. W. and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika* **84**, 641-652.

Monti, A. C. (2003). A note on the estimation of the skew normal and the skew exponential power distributions. *Metron - International Journal of Statistics* **LXI**, 205-219.

Sudhansu S. Maiti
Department of Statistics
Visva-Bharati University
Santiniketan-731 235, India
dssm1@rediffmail.com

Mithu Dey
Department of Basic Sciences (Mathematics)
Asansol Engineering College
Vivekananda Sarani, Kanyapur-713 305, India
mithu.dey@gmail.com