

Moments of a Class of Internally Truncated Normal Distributions

Hea-Jung Kim¹⁾

Abstract

Moment expressions are derived for the internally truncated normal distributions commonly applied to screening and constrained problems. They are obtained from using a recursive relation between the moments of the normal distribution whose distribution is truncated in its internal part. Closed form formulae for the moments can be presented up to N^{th} order under the internally truncated case. Necessary theories and two applications are provided.

Keywords: Internally truncated normal distribution; kurtosis; moment; skewness.

1. Introduction

The internally truncated normal distribution with parameters (μ, σ) is defined as the distribution of $X = [Y|Y < \alpha, Y > \beta]$, where $Y \sim N(\mu, \sigma^2)$ and the internal truncation points α and β ($\alpha < \beta$). The probability density function of X is

$$f(x) = \sigma^{-1} \phi\left(\frac{x - \mu}{\sigma}\right) \left[1 - \Phi\left(\frac{\beta - \mu}{\sigma}\right) + \Phi\left(\frac{\alpha - \mu}{\sigma}\right)\right]^{-1}, \quad x < \alpha, x > \beta, \quad (1.1)$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ denote the *c.d.f.* and *p.d.f.* of the standard normal variable, respectively. The degree of truncation is $\Phi((\beta - \mu)/\sigma) - \Phi((\alpha - \mu)/\sigma)$ (from inside).

The effects of truncation on modeling have long been recognized (see, for example, DePriest, 1983) and are commonly referred to in the literatures of classic statistics such as Johnson *et al.* (1994). Genton (2005) also notes that the internally truncated normal distribution is useful for a selection model whose moments are directly related with those of internally truncated normal distribution (see, Section 4.1).

One of the most common techniques for characterizing a truncated distribution is the method of TMs (truncated moments). TMs are useful to providing descriptive information about the distribution. In addition to their utility as general descriptive measures,

1) Professor, Department of Statistics, Dongguk University, Pil-Dong 3Ga, Chung-Gu, Seoul 100-715, Korea.

E-mail : kim3hj@dongguk.edu

moments are also employed for estimating parameters of the distribution. Sugiura and Gomi (1985) have given Pearson diagrams (of the coefficients of skewness and kurtosis) for the doubly truncated normal distributions. Shah and Jaiswal (1966) has discussed the moments of doubly truncated normal distribution. Hall (1979) has also derived inverse moments for a class of singly truncated normal distributions. For the other distributions, Kim (2007) obtained moments for a truncated generalized- t and Jawitz (2004) derived several expressions for TMs useful to numerical computation. These include TMs of normal, lognormal, Pearson type III, log Pearson type III and extreme value (Weibull and Gumbel) distributions. See also Shah (1966) for the TMs of the binomial distribution. Although TMs for several distributions have been studied and applied in the literature, we are not aware of any detailed exposition of TMs for the internally truncated normal distribution. This lack of detailed exposition motivates the investigation described in this article.

2. Preliminaries

Prior to derive the TMs, we provide lemmas useful for calculating them. For notational convenience, we use $TN_{(\alpha,\beta)}(\mu, \sigma^2)$ to indicate a doubly truncated normal distribution with the lower and upper truncation points α and β , respectively; the degrees of truncation are $\Phi((\alpha - \mu)/\sigma)$ (from below) and $1 - \Phi((\beta - \mu)/\sigma)$ (from above). So that $[Y|\alpha < Y < \beta] \sim TN_{(\alpha,\beta)}(\mu, \sigma)$, where $Y \sim N(\mu, \sigma^2)$. We also use $TN_{\mathbb{R} \setminus (\alpha,\beta)}(\mu, \sigma^2)$ to denote a internally truncated normal distribution whose *p.d.f.* is (1.1).

Lemma 2.1 *Let $X \sim TN_{\mathbb{R} \setminus (\alpha,\beta)}(\mu, \sigma^2)$. Then the distribution of $(X - \mu)/\sigma$ is $TN_{\mathbb{R} \setminus (a,b)}(0, 1)$, where $a = (\alpha - \mu)/\sigma$ and $b = (\beta - \mu)/\sigma$.*

Proof: By definition, $X = [Y|Y < \alpha, Y > \beta]$, where $Y \sim N(\mu, \sigma^2)$. Thus

$$(X - \mu)/\sigma = [(Y - \mu)/\sigma \mid (Y - \mu)/\sigma < a, (Y - \mu)/\sigma > b], \quad (2.1)$$

where $(Y - \mu)/\sigma \sim N(0, 1)$. □

Note that the distribution $TN_{\mathbb{R} \setminus (\alpha,\beta)}(\mu, \sigma^2)$ strictly includes the singly truncated normal distributions. If $\alpha = -\infty$ (or $\beta = \infty$), the $TN_{\mathbb{R} \setminus (\alpha,\beta)}(\mu, \sigma^2)$ distribution reduces to $TN_{(\beta,\infty)}(\mu, \sigma^2)$ (or distribution $TN_{(-\infty,\alpha)}(\mu, \sigma^2)$), the left (or right) truncated normal with lower (or upper) truncation point β (or α).

Lemma 2.2 *If $X \sim TN_{\mathbb{R} \setminus (\alpha,\beta)}(\mu, \sigma^2)$, the moment generating function of X distribution is*

$$M_X(t) = \frac{e^{\mu t + \sigma^2 t^2/2} \{\Phi(-b + \sigma t) + \Phi(a - \sigma t)\}}{\Phi(-b) + \Phi(a)}, \quad t \in \mathbb{R}, \quad (2.2)$$

where $a = (\alpha - \mu)/\sigma$ and $b = (\beta - \mu)/\sigma$.

Proof:

$$\begin{aligned} M_X(t) &= \{\Phi(-b) + \Phi(a)\}^{-1} \left\{ \int_{-\infty}^{\alpha} \frac{e^{tx}}{\sigma} \phi\left(\frac{x-\mu}{\sigma}\right) dx + \int_{\beta}^{\infty} \frac{e^{tx}}{\sigma} \phi\left(\frac{x-\mu}{\sigma}\right) dx \right\} \\ &= \frac{e^{\mu t + \sigma^2 t^2 / 2}}{\Phi(-b) + \Phi(a)} \left\{ \int_{-\infty}^{\alpha} \sigma^{-1} \phi(u(x)) dx + \int_{\beta}^{\infty} \sigma^{-1} \phi(u(x)) dx \right\}, \end{aligned}$$

where $u(x) = \{x - (\mu + \sigma^2 t)\}/\sigma$. A direct integration using the change of variable, *i.e.* $z = \{x - (\mu + \sigma^2 t)\}/\sigma$, gives result. \square

3. The Moments

As implied by Lemma 2.1, it is sufficient to compute the moments of $Z = (X - \mu)/\sigma$ for those of $X \sim TN_{\mathbb{R} \setminus (\alpha, \beta)}(\mu, \sigma^2)$, where $Z \sim TN_{\mathbb{R} \setminus (a, b)}(0, 1)$, $a = (\alpha - \mu)/\sigma$, and $b = (\beta - \mu)/\sigma$. The relationship between X and Z is

$$X = \mu + \sigma Z. \quad (3.1)$$

From (1.1), we see that Z has the density

$$f_Z(z) = [\Phi(-b) + \Phi(a)]^{-1} \phi(z), \quad z < a, z > b. \quad (3.2)$$

The moment generating function of Z is

$$M_Z(t) = e^{t^2/2} \frac{\{\Phi(-b+t) + \Phi(a-t)\}}{\Phi(-b) + \Phi(a)}, \quad t \in \mathbb{R} \quad (3.3)$$

by Lemma 2.2. Naturally, the moments of Z can be obtained by using the moment generating function differentiation. For example

$$EZ = M'_Z(t)|_{t=0} = \frac{\phi(b) - \phi(a)}{\Phi(-b) + \Phi(a)}. \quad (3.4)$$

Unfortunately, for higher moments this rapidly becomes tedious.

An alternative procedure, making use of the following theorem, gives a simple way of calculating the moments.

Theorem 3.1 *If $Z \sim TN_{\mathbb{R} \setminus (a, b)}(0, 1)$, then, for $\ell = 1, 2, \dots$,*

$$E[Z^\ell] = (\ell - 1)E[Z^{\ell-2}] + \frac{b^{\ell-1}\phi(b) - a^{\ell-1}\phi(a)}{\Phi(-b) + \Phi(a)}. \quad (3.5)$$

Proof: We see that

$$\frac{dz^{k+1}\phi(z)}{dz} = (k+1)z^k\phi(z) - z^{k+2}\phi(z) \quad \text{for } k = -1, 0, 1, 2, \dots$$

This gives

$$\begin{aligned} M_k &= \int_a^b \{(k+1)z^k - z^{k+2}\} \phi(z) dz / (\Phi(-b) + \Phi(a)) \\ &= (b^{k+1}\phi(b) - a^{k+1}\phi(a)) / (\Phi(-b) + \Phi(a)) \end{aligned}$$

for $k = -1, 0, 1, 2, \dots$. Therefore,

$$\begin{aligned} E[(k+1)Z^k] - E[Z^{k+2}] &= \int_{\mathbb{R} \setminus (a,b)} \{(k+1)z^k - z^{k+2}\} \phi(z) dz / (\Phi(-b) + \Phi(a)) \\ &= \{(k+1)E[U^k] - E[U^{k+2}]\} / (\Phi(-b) + \Phi(a)) - M_k, \end{aligned}$$

where U is a $N(0, 1)$ variable and $(k+1)E[U^k] - E[U^{k+2}] = 0$. Setting $\ell = k+2$, we have the result. \square

The following corollary is straightforward from the proof of Theorem 3.1.

Corollary 3.1 *Let $V \sim TN_{(a,b)}(0, 1)$, a doubly truncated standard normal with the lower and upper points a and b , respectively. Then*

$$E[V^\ell] = (\ell - 1)E[V^{\ell-2}] + \frac{a^{\ell-1}\phi(a) - b^{\ell-1}\phi(b)}{\Phi(b) - \Phi(a)}, \quad \ell = 1, 2, \dots \quad (3.6)$$

The expression (3.6) is different from those given in Johnson *et al.* (1994) and Jawitz (2004), and hence can be taken as an alternative method for calculating the moments of the doubly truncated normal distribution.

Corollary 3.2 *Let $U \sim N(0, 1)$, $V \sim TN_{(a,b)}(0, 1)$, and $Z \sim TN_{\mathbb{R} \setminus (a,b)}(0, 1)$. Then*

$$E[U^\ell] = \{\Phi(b) - \Phi(a)\}E[V^\ell] + \{\Phi(-b) + \Phi(a)\}E[Z^\ell], \quad \ell = 1, 2, \dots, \quad (3.7)$$

where $E[U^\ell] = \ell! 2^{-\ell/2} / (\ell/2)!$ for even ℓ and $E[U^\ell] = 0$ for odd ℓ .

Proof: The statement is equivalent to

$$\int_{-\infty}^{\infty} u^\ell \phi(u) du = \int_a^b v^\ell \phi(v) dv + \int_{\mathbb{R} \setminus (a,b)} z^\ell \phi(z) dz.$$

The moment $E[U^\ell]$ is given in Johnson *et al.* (1994, p. 89). \square

Therefore, if we obtain the moments of Z (or V) distribution, those of Z (or V) distribution can be calculated from the relation (3.7). The equation (3.5) gives a recursive method for evaluating the moments of $Z \sim TN_{\mathbb{R} \setminus (a,b)}(0, 1)$. By setting $\ell = 1, 2, 3, 4$, we obtain expressions up to the fourth moments of $Z \sim TN_{\mathbb{R} \setminus (a,b)}(0, 1)$.

One finds the moments,

$$\begin{aligned} E[Z] &= \frac{\phi(b) - \phi(a)}{\Phi(-b) + \Phi(a)}, \\ \text{Var}(Z) &= 1 + \frac{b\phi(b) - a\phi(a)}{\Phi(-b) + \Phi(a)} - \left(\frac{\phi(b) - \phi(a)}{\Phi(-b) + \Phi(a)} \right)^2, \\ \text{Skewness}(Z) &= \gamma^{-3/2} \{ \beta_2 - \beta_0(1 + 3\beta_1 - 2\beta_0^2) \}, \\ \text{Kurtosis}(Z) &= \gamma^{-2} \{ 3(1 + \beta_1) + \beta_3 - \beta_0(4\beta_2 - 6\beta_1\beta_0 + 2\beta_0 + 3\beta_0^3) \}, \end{aligned}$$

where $\beta_\ell = \{b^\ell \phi(b) - a^\ell \phi(a)\} / \{\Phi(-b) + \Phi(a)\}$, $\ell = 0, 1, 2, 3$, and $\gamma = 1 + \beta_1 - \beta_0^2$.

Using the relationship (3.1), we have the following result. Denoting EZ^ℓ by $\lambda_\ell(a, b)$, one then has the general formula of the moments of $X \sim TN_{(\alpha, \beta)}(\mu, \sigma^2)$ given by

$$EX^\ell = \sum_{j=0}^{\ell} \binom{\ell}{j} \mu^{\ell-j} \sigma^j \lambda_j(a, b) \quad (3.8)$$

and $\text{Var}(X) = \sigma^2 \text{Var}(Z)$, where $a = (\alpha - \mu)/\sigma$, $b = (\beta - \mu)/\sigma$. Further note that the skewness (the third standardized central moment) and kurtosis of X and Z distributions are the same.

As for an example, we use the half-normal distribution, written by $TN_{\mathbb{R} \setminus (-\infty, 0)}(0, 1)$. If $Z \sim TN_{\mathbb{R} \setminus (-\infty, 0)}(0, 1)$, the *p.d.f.* in (1.1) reduces to

$$f_Z(z) = 2\phi(z), \quad z > 0. \quad (3.9)$$

One finds the moments

$$\begin{aligned} E[Z] &= \sqrt{2/\pi}, \\ \text{Var}(Z) &= 1 - 2/\pi, \\ \text{Skewness}(Z) &= (4/\pi - 1)(2/\pi)^{1/2}(1 - 2/\pi)^{-3/2}, \\ \text{Kurtosis}(Z) &= (1 - 2/\pi)^{-2}(3 - 4/\pi - 12/\pi^2). \end{aligned}$$

These four values of Z agree with those of the half-normal distribution given in Sugiura and Gomi (1985).

4. Applications

4.1. Sum of the normal and internally truncated normals

Let $\mathbf{Y} = (Y_1, Y_2)^\top$ be a bivariate standard normal with correlation coefficient ρ . Under the distribution of \mathbf{Y} , suppose that the distribution of Y_2 is internally truncated. So that the random variable Y_2 is truncated in the form of $Y_2 < a$ and $Y_2 > b$, where a and b ($a < b$) are truncation points. Then it is straightforward to see that the marginal

density of Y_1 is the same as that of $W = [Y_1|Y_2 < a, Y_2 > b]$. The density given by Genton (2005) is

$$f(w) = \phi(w) \frac{\Phi\left(-(b - \rho w)/\sqrt{1 - \rho^2}\right) + \Phi\left((a - \rho w)/\sqrt{1 - \rho^2}\right)}{\Phi(-b) + \Phi(a)}, \quad w \in \mathbb{R}. \quad (4.1)$$

Our problem is to obtain the moments of the conditional distribution, i.e. $E[W^k]$, $k = 1, 2, \dots$. Naturally, the moments of W can be obtained by using the moment generating function differentiation. Unfortunately this needs tedious analytic calculation. We describe a simple alternative procedure as follows.

It is well known that the bivariate standard normal variables can be expressed in terms of two independent standard normals U and V :

$$Y_2 = V \quad \text{and} \quad Y_1 = \rho V + \sqrt{1 - \rho^2}U. \quad (4.2)$$

This in turn gives following result. Since U and V are independent, by using (4.2), we can express $W = [Y_1|Y_2 < a, Y_2 > b]$ as

$$W = \rho V_{\mathbb{R} \setminus (a, b)} + \sqrt{1 - \rho^2}U, \quad (4.3)$$

where $V_{\mathbb{R} \setminus (a, b)} \sim TN_{\mathbb{R} \setminus (a, b)}(0, 1)$. This representation of W immediately gives

$$E[W^k] = \sum_{\ell=0}^k \binom{k}{\ell} \rho^\ell (1 - \rho^2)^{(k-\ell)/2} E[Z^\ell] E[U^{k-\ell}], \quad (4.4)$$

where $Z = V_{\mathbb{R} \setminus (a, b)} \sim TN_{\mathbb{R} \setminus (a, b)}(0, 1)$. Denoting $E[Z^\ell]$ and $E[U^{k-\ell}]$ by respective values in Theorem 3.1 and Corollary 3.2, one then evaluate $E[W^k]$, $k = 1, 2, \dots$. For example, one finds moments,

$$\begin{aligned} E[W] &= \rho \frac{\phi(b) - \phi(a)}{\Phi(-b) + \Phi(a)}, \\ \text{Var}(W) &= 1 + \rho^2 \left\{ \frac{b\phi(b) - a\phi(a)}{\Phi(-b) + \Phi(a)} - \left(\frac{\phi(b) - \phi(a)}{\Phi(-b) + \Phi(a)} \right)^2 \right\}. \end{aligned}$$

4.2. An inequality constrained regression

With the normal linear regression model, we assume that an $n \times 1$ vector of observations \mathbf{y} on our dependent variable satisfies

$$\mathbf{y} = \mathbf{X}\boldsymbol{\theta} + \boldsymbol{\varepsilon}, \quad (4.5)$$

where $\boldsymbol{\varepsilon} \sim N_n(\mathbf{0}, \sigma^2 I)$ with known σ^2 , $\mathbf{X} : n \times k$, and $\text{rank}(\mathbf{X}) = k$ for the regression model. Under the model, let $\hat{\boldsymbol{\theta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$ and $M = (\mathbf{y} - \mathbf{X}^T \hat{\boldsymbol{\theta}})^T (\mathbf{y} - \mathbf{X}^T \hat{\boldsymbol{\theta}})$. Then the joint *p.d.f.* for the elements of \mathbf{y} given \mathbf{X} and $\boldsymbol{\theta}$ is

$$f(\mathbf{y} | \mathbf{X}, \boldsymbol{\theta}) \propto \sigma^{-n} \exp \left\{ -\frac{1}{2\sigma^2} [M + (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^T \mathbf{X}^T \mathbf{X} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})] \right\}. \quad (4.6)$$

For a possibly improper diffuse prior $p(\boldsymbol{\theta}) \propto q(\boldsymbol{\theta})$, the posterior distribution posterior $p.d.f.$ for the elements of $\boldsymbol{\theta}$ is proportional to the product of a multivariate normal $p.d.f.$ and $q(\boldsymbol{\theta})$:

$$p(\boldsymbol{\theta} | \mathbf{y}, \mathbf{X}) \propto \exp \left\{ -\frac{(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^T \mathbf{X}^T \mathbf{X} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})}{2\sigma^2} \right\} q(\boldsymbol{\theta}). \quad (4.7)$$

When, in priori, an inequality constraint for a regression coefficient θ_j is given in the form of $(\theta_j < \alpha, \theta_j > \beta)$, where α and β are known. We can define $q(\boldsymbol{\theta})$ as $I(\theta_j < \alpha, \theta_j > \beta)$, where $I(\cdot)$ is an indicator function. By integrating (4.7) with respect to $\theta_1, \theta_2, \dots, \theta_{j-1}, \theta_{j+1}, \dots, \theta_k$, we have the marginal posterior $p.d.f.$ of θ_j

$$p(\theta_j | \mathbf{y}, \mathbf{X}) \propto \exp \left\{ -\frac{(\theta_j - \hat{\theta}_j)^2}{2\sigma^2 h^{jj}} \right\} I(\theta_j < \alpha, \theta_j > \beta), \quad (4.8)$$

the kernel of an internally truncated normal $p.d.f.$, where h^{jj} is the (j, j) element of $(\mathbf{X}^T \mathbf{X})^{-1}$. Thus the marginal posterior distribution of θ_j is

$$\theta_j | \mathbf{y}, \mathbf{X} \sim TN_{\mathbb{R} \setminus (\alpha, \beta)}(\hat{\theta}_j, \sigma^2 h^{jj}). \quad (4.9)$$

Under the quadratic loss function, the Bayes estimator of θ_j is

$$\hat{\theta}_{j, Bayes} = \hat{\theta}_j + \frac{\sigma \sqrt{h^{jj}} \{\phi(b) - \phi(a)\}}{\Phi(-b) + \Phi(a)} \quad (4.10)$$

by (3.8), where $a = (\alpha - \hat{\theta}_j)/(\sigma \sqrt{h^{jj}})$ and $b = (\beta - \hat{\theta}_j)/(\sigma \sqrt{h^{jj}})$. Note that the estimator has following properties: (i) when $|a| > |b|$ ($a < b$), $\hat{\theta}_{j, Bayes} > \hat{\theta}_j$; (ii) when $|a| = |b|$ ($a < b$), $\hat{\theta}_{j, Bayes} = \hat{\theta}_j$; (iii) when $|a| < |b|$ ($a < b$), $\hat{\theta}_{j, Bayes} < \hat{\theta}_j$.

5. Concluding Remarks

Sometimes, observed data sets are almost exclusively truncated, because of analytical detection limits or spatial and temporal limitations on data collection. In order to analyze the data sets, some broadly related proposals and results have appeared in the literature under the concept of the truncated distribution. In the same line, the present paper has considered the moments of an internally truncated normal distribution, providing descriptive information about the distribution of internally truncated data sets. We see that the interest of TMs comes from both theoretical and applied directions. On the theoretical side, they have utility as general descriptive measures of the internally truncated normal distribution. In the applied view point, the moments can be employed for solving statistically constrained problems as given in Section 4.

References

- DePriest, D. J. (1983). Using the singly truncated normal distribution to analyze satellite data. *Communications in Statistics-Theory and Methods*, **12**, 263–272.

- Genton, M. G. (2005). Discussion of "the skew-normal". *Scandinavian Journal of Statistics*, **32**, 189–198.
- Hall, R. L. (1979). Inverse moments for a class of truncated normal distributions. *Sankhyā*, Ser. B, **41**, 66–76.
- Jawitz, J. W. (2004). Moments of truncated continuous univariate distributions. *Advances in Water Resources*, **27**, 269–281.
- Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994). *Continuous Univariate Distributions*, Vol. 1, 2nd ed., John Wiley & Sons, New York.
- Kim, H. J. (2007). Moments of truncated Student-*t* distribution. *Journal of the Korean Statistical Society*, accepted.
- Shah, S. M. (1966). On estimating the parameter of a doubly truncated binomial distribution. *Journal of the American Statistical Association*, **61**, 259–263.
- Shah, S. M. and Jaiswal, M. C. (1966). Estimation of parameters of doubly truncated normal distribution from first four sample moments. *Annals of the Institute of Statistical Mathematics*, **18**, 107–111.
- Sugiura, N. and Gomi, A. (1985). Pearson diagrams for truncated normal and truncated Weibull distributions. *Biometrika* **72**, 219–222.

[Received September 2007, Accepted November 2007]