DOI: https://doi.org/10.24297/jam.v20i.8934

Comparison among Some Methods for Estimating the Parameters of Truncated Normal Distribution

Hilmi Kittani*, Mohammad Alaesa, Gharib Gharib

Abstract

The aim of this study is to investigate the effect of different truncation combinations on the estimation of the normal distribution parameters. In addition, is to study methods used to estimate these parameters, including MLE, moments, and L-moment methods. On the other hand, the study discusses methods to estimate the mean and variance of the truncated normal distribution, which includes sampling from normal distribution, sampling from truncated normal distribution and censored sampling from normal distribution. We compare these methods based on the mean square errors, and the amount of bias. It turns out that the MLE method is the best method to estimate the mean and variance in most cases and the L-moment method has a performance in some cases.

Keywords: MLE method, method of moments, L-moment method, truncated normal distribution, censored sample.

1. Introduction

The idea of truncated samples appeared on surface quite early when Sir Francis, Galton (1897) studies registered speeds of American trotting horses. He assumed that the distribution is normal, estimated the mean by the sample modes, and used semi-interquartile ranges to estimate population standard deviations. Led Karl Pearson (1902) estimated the parameters of normal distribution from truncated sample by fitting parabolas to logarithms of the sample frequencies. Then, he used the method of moments to estimate the parameters from a singly left truncated sample with Alice Lee, Pearson. Special tables were needed to estimate practical procedure, which were expanded by Lee (1914).

Since the equations of maximum likelihood estimates (MLE) from truncated and censored sample and from normal distribution could not be solved directly because of their rather complex non-linear equations, many researchers looked for methods to simplify and reduce this calculation and tried to be more accurate. For example, the tables by Cohen and Woodward 1953) for singly truncated normal samples which found the values by interpolating between table entries, figure by Cohen (1957) for doubly truncated and censored samples, tables by Gupta (1952) also gave the asymptotic variances for estimators of singly and doubly censored samples, tables and figure by Cohen (1959), Cohen (1961) tables for MLE of singly truncated and singly censored samples and tables of Cooley and Cohen after nearly 9 years. Lifsey (1965) gave an auxiliary estimating function for doubly truncated normal and expanded the table and figure that Cohen used.

Shah and Jaiswal (1966) used the first four moments to estimate parameters of doubly truncated normal distribution and compared it with MLE. Sugiura and Gomi (1985) presented Pearson's diagrams for the truncated normal distributions. Mittal and Dahiya (1987) proposed a new estimating procedure called "Mixed Estimation." In addition, by simulation study, they compared between the MLEs and the modified MLEs, and found that the modified method was better than the MLE.

Hegde and Dahiya (1989), studied the truncated normal distribution but this time just for singly truncated, by using simulation to compare between the modified maximum likelihood estimator, and the mixed estimator. Barr and Sherrill (1999) published a simulation study to estimate the mean and variance of truncated normal distributions for sample sizes (100, 50, 36, 25, and 10). They compared the results by the censored sample, and found that the full-sample estimators have greater mean square error than the censored-sample estimators in general. He proposed a new alternative expression for the variance in terms of the chi-square CDF of one-sided truncations. Barr's formula was used in standard software. However, this work is insufficient to return to take just the left truncated case for few value of truncated point too.

CC BY

Many economy applications using truncated normal distribution were dissected. For example, Johnson and Thomopoulos (2002) talked about replacing used left truncated normal distribution instead of normal distribution to determine the safety stock required to achieve a better service level since the demand (the quantity of a commodity or a service that people are willing or able to buy at a certain price) truncated from left at zero value. which was done by other researchers as Bookbinder and Lordahl (1989), and Sinha (1991). Johnson and Thomopoulos (2002) presented tables of the cumulative distribution function of the left truncated normal distribution, and later for doubly truncated case, Johnson (2002).

Khasawneh, Bowling et al. (2004) made useful tables of a truncated standard normal distribution in a singly truncated case, which was extended the next year to doubly truncated case Khasawneh, Bowling et al. (2005). Iwueze (2007) took the mean equal one and studied the relation between truncated and untruncated normal distribution. He observed that the truncated values were greater than untruncated ones and the variables of truncated and untruncated had the same mean equal one and variance σ^2.

Bebu and Mathew (2009) took three cases of truncated cases with sample size ranging between five and eighty to provide formulas for (generalized) confidence intervals around the truncated moments, the mean and the second moment. Hattaway (2010) made a summary of truncated normal distribution problem in many sides. He took many compensations of mu and sigma of normal distribution and truncates all of them at zero to compare between MLEs and method of moment using simulation study.

2. Truncated Distributions

2.1 Normal distribution

We know that the probability density function of the normal distribution is:

$$g(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad -\infty\le x\le\infty$$

(2.1.1)

Where μ is the mean or expectation and σ^2 is the variance. The simplest form of the normal distribution occurs when the mean and variance equal zero and one respectively. The density function for the standard normal is:

(2.1.2)

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{\frac{-1}{2}x^2} \qquad -\infty\le x\le\infty$$

$$G(x)=Q\left(\frac{x-\mu}{\sigma}\right)=\frac{1}{2}\left[1+\operatorname{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right]$$

The CDF for normal distribution G(x) with mean μ and standard deviation σ, is

(2.1.3)

Where

$$Q(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{\frac{-t^2}{2}}dt$$

(2.1.4)

The CDF of normal distribution in the standard form is:

$$\Phi(x)=\frac{1}{2}\left[1+\operatorname{erf}\left(\frac{x}{\sqrt{2}}\right)\right]$$

(2.1.5)

Where erf refers to the error function that is defined by:

$$\operatorname{erf}(x) = \frac{1}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

(2.1.6)

2.2 Truncated Distributions

A truncated distribution is a conditional distribution on a specific range, which restricted the full range. In fact, it is the part of an untruncated distribution that is above or below some specified value or between two values. That means if the range of untruncated distribution is all real number, the new range will be from minus infinity to a certain value or from a value to infinity, or between two values, but not separated intervals. Truncated distributions issues happened when observation outside a certain region could not be measured or recorded.

Furthermore, truncated random variables have found important applications in actuarial practice. They also can have potential applications in the context of lifetime data analysis.

There are several types of truncated distributions:

One-sided truncation: Truncated from above: all values greater than a truncated point are cut off, so the range will be from the minimum value in un truncated range to a truncated point. Or truncated from below: all values smaller than a truncated point are cut off, so the range will be from the truncated point to maximum value in un truncated range.

T**wo sided truncation** (Double truncation): all values out the truncated region are cut off, so the new range will be between the two truncated points.

Let X be a random variable with probability density function h(x) and with CDF H(x), then the probability density function of Y (the truncated version of the variable X which is truncated from the left at **a** and on the right at **b**)

$$k(y) = \frac{h(y)}{H(b) - H(a)}, \quad a \le y \le b$$

is given by :

(2.2.1)

In addition to that, the CDF of Y is:

$$K(y) = \frac{H(y) - H(a)}{H(b) - H(a)},$$

(2.2.2)

Notably, the expectation of a truncated random variable is given as:

$$E(X \mid a < x < b) = \frac{\int_a^b x g(x) dx}{H(b) - H(a)},$$

(2.2.3)

Where g(x) = h(x), for all x ϵ (a , b) and zero everywhere else.

The probability density function of double truncated normal distribution is given by:

$$f(x)=\frac{\frac{1}{\sigma}\varphi(\frac{x-\mu}{\sigma})}{\Phi(\frac{b-\mu}{\sigma})-\Phi(\frac{a-\mu}{\sigma})},\quad a\le x\le b$$

(2.2.4)

Where φ and Φ are pdf and CDF for the standard normal defined in (2.1.2) and is defined in

$$f(x)=\frac{\varphi(\xi)}{\sigma Z},\quad a\le x\le b$$

(2.2.5)

$$E(X|a<X<b)=\mu+\frac{\varphi(\frac{a-\mu}{\sigma})-\varphi(\frac{b-\mu}{\sigma})}{\Phi(\frac{b-\mu}{\sigma})-\Phi(\frac{a-\mu}{\sigma})}\sigma,\quad a\le x\le b$$

Finally, the mean and variance of this distribution are given as:

(2.2.6)

$$\alpha=\frac{a-\mu}{\sigma},\quad \beta=\frac{b-\mu}{\sigma},\quad \xi=\frac{x-\mu}{\sigma},$$

$$Var(X|a<X<b)=\sigma^2\left[1+\frac{\frac{a-\mu}{\sigma}\varphi(\frac{a-\mu}{\sigma})-\frac{b-\mu}{\sigma}\varphi(\frac{b-\mu}{\sigma})}{\Phi(\frac{b-\mu}{\sigma})-\Phi(\frac{a-\mu}{\sigma})}-\left(\frac{\varphi(\frac{a-\mu}{\sigma})-\varphi(\frac{b-\mu}{\sigma})}{\Phi(\frac{b-\mu}{\sigma})-\Phi(\frac{a-\mu}{\sigma})}\right)^2\right]$$

(2.2.7)

While in the first case of one sided truncation, truncated from above, the lowest truncated point assumes to be -∞, so that, φ(α) = Φ (α) = 0.

Then the mean and variance will be as:

$$E(X|X<b)=\mu-\sigma\frac{\varphi(\beta)}{\Phi(\beta)},\quad x\le b$$

$$Var(X|X<b)=\sigma^2\left[1-\frac{\beta\varphi(\beta)}{\Phi(\beta)}-\left(\frac{\varphi(\beta)}{\Phi(\beta)}\right)^2\right]$$

(2.2.8)

(2.2.9)

As well, the second case of one sided truncation, truncated from below, the largest truncated point will be ∞, so that, φ(β) = 0 and Φ (β) = 1.

Again, the mean and variance of this case will be, see Johnson and Kotz et. al. (1994):

(2.2.10)

$$Var(X|X>a)=\sigma^2\left[1+\frac{\alpha\varphi(\alpha)}{1-\Phi(\alpha)}-\left(\frac{\varphi(\alpha)}{1-\Phi(\alpha)}\right)^2\right]$$

$$E(X \mid X > a) = \mu + \sigma \frac{\varphi(\alpha)}{1-\Phi(\alpha)}, \qquad x \geq a$$

(2.2.11)

Figures (1), (2), and (3) show some truncated normal distributions with μ=2 and σ =1 for different truncation combinations.

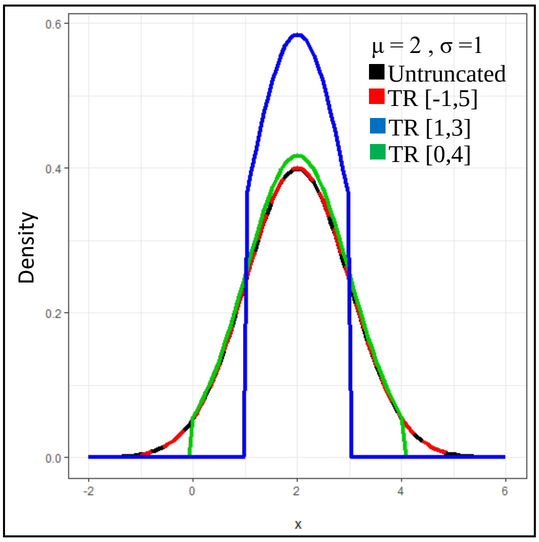

Fig. 1 : Truncated and untruncated normal distributions of μ = 2 , σ =1. Truncated by the interval : [1,3] , [0,4] , [-1,5]

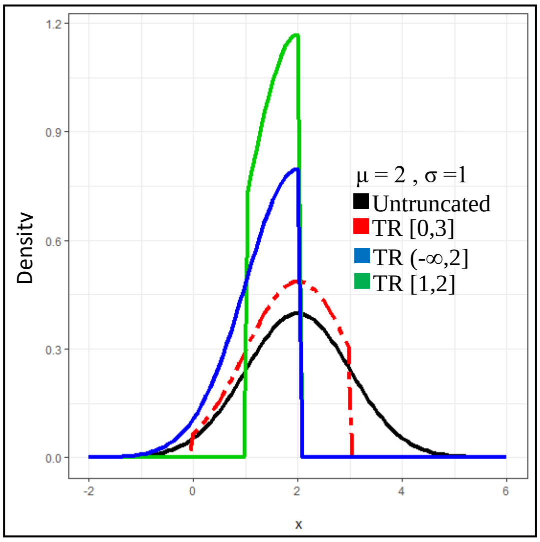

Fig. 2 : Truncated and untruncated normal distributions of μ = 2 , σ =1. Truncated by the interval : [0,3], (-∞,2] , [1,2]

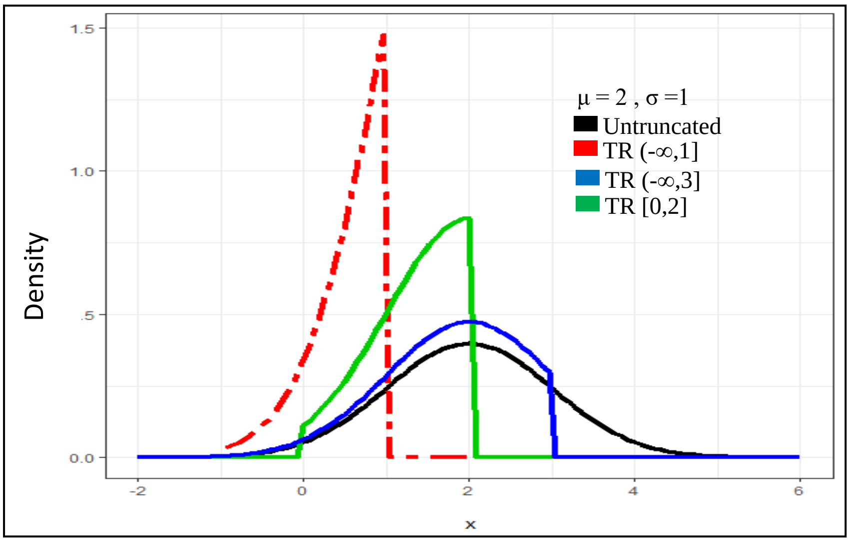

Fig. 3 : Truncated and untruncated normal distributions of μ = 2 , σ =1. Truncated by the interval : (-∞,1] , (-∞,2] , (-∞,3]

CC BY

3. Methods of Estimation

We will discuss three methods of estimation in this work:

3.1 Method of Moment (MOM)

It was introduced by Karl Pearson in 1894. The main idea here is to estimate population moments from samples that are drawn from this population. Method of Moments (MOM) has the virtue of being quite simple to use, and has many advantages not only that its calculation is usually simple, but also, sometimes, it is used to be the first approximation in other estimation methods when needed like using Newton–Raphson technique in MLE. On the other hand, estimates found by this method may be out of the parameters range sometime, especially in large sample sizes.

Definitions and Basic Properties

Let $X_1, X_2,\ldots,X_n$ be a sample from a population with pdf, $f(x, \theta_1,\ldots,\theta_k)$. Method of Moment estimators are found by equating the first k sample moments to the corresponding k population moments, and solving the resulting system of simultaneous equations. There are two forms of Method of Moments depending on the following definition:

1) $E(X^k)$ is the k^{th} population moment of the distribution (about the origin), for k= 1,2, …

2) $E[(X - \mu)^k]$ is the k^{th} (theoretical) moment of the distribution (about the mean), for k = 1, 2, …

3) $M_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$ is the k^{th} sample moment about origin, for k = 1, 2, …

4) $M_k^* = \frac{1}{n}\sum_{i=1}^{n} (X_i - \bar{X})^k$ is the k^{th} sample moment about the mean, for k = 1, 2, …

The first form is done by equating the sample moments about the origin with the corresponding population moments, until we form equations that has a size equal to the number of parameters. The second form is formed by equating sample moments about the mean with the corresponding population moments about the mean. Once again, we form equations that has a size equal to the number of parameters. Casella and Berger (2002).

Moments of Truncation Normal Distributions

In light of the above definitions and properties, equating the E(X) of the truncated normal distribution from (2.2.6) by the first sample moment form:

$$\mu + \frac{\varphi(\alpha)-\varphi(\beta)}{\Phi(\beta)-\Phi(\alpha)}\sigma = \bar{X} \quad (3.1.1)$$

Where μ, σ are the parameters of normal distribution and $\bar{X}$ is the sample mean.

Then, by the same way, equating $E(X^2)$ with the second sample moment about the origin forms:

$$\sigma^2\left[1+\frac{\alpha\varphi(\alpha)-\beta\varphi(\beta)}{\Phi(\beta)-\Phi(\alpha)}-\left(\frac{\varphi(\alpha)-\varphi(\beta)}{\Phi(\beta)-\Phi(\alpha)}\right)^2\right]+\mu^2+\left[\frac{\varphi(\alpha)-\varphi(\beta)}{\Phi(\beta)-\Phi(\alpha)}\sigma\right]^2+$$
$$2\mu\frac{\varphi(\alpha)-\varphi(\beta)}{\Phi(\beta)-\Phi(\alpha)}\sigma=\frac{\sum(x_i)^2}{n}$$

(3.1.2)

The lack of a closed-form solution means that the MOM estimators must be found and can be done by using an iterative method.

CC BY

For the one sided truncation, take truncation from below, that is by assuming the upper truncated point goes to infinity, and take the equation of one sided truncated normal distribution mean and variance (2.2.8) , (2.2.9) into consideration, the previous double truncated equations transforms to one sided equations will be , see Casella and Berger (2002) :

(3.1.3)

$$\sigma^2\left[1+\frac{\alpha\varphi(\alpha)}{1-\Phi(\alpha)}-\left(\frac{\varphi(\alpha)}{1-\Phi(\alpha)}\right)^2\right]+\mu^2+\left[\frac{\varphi(\alpha)}{1-\Phi(\alpha)}\sigma\right]^2+$$

$$2\mu\frac{\varphi(\alpha)}{1-\Phi(\alpha)}\sigma=\frac{\sum(x_i)^2}{n}$$

(3.1.4)

3.2 Maximum likelihood Estimation (MLE) Method:

Maximum likelihood estimation method (MLE) is the most popular method to derive estimators. The main idea in this method is finding the parameters values that maximizes the likelihood function form from the density function of the distribution.

MLE's of Truncated Normal Distribution

The likelihood for the truncated normal distribution is:

$$L(\mu,\sigma^2)=\prod_{i=1}^{n} f(x_i)=(z)^{-n}\left(\sqrt{2\pi\sigma^2}\right)^{-n}\exp\left(\frac{-\sum(x_i-\mu)^2}{2\sigma^2}\right)$$

(3.2.1)

Where z = Φ (β) - Φ (α), Now, define:

$$\psi\ (\mu,\sigma)=\int_a^b \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{-(y-\mu)^2}{2\sigma^2}\right)dy$$

(3.2.2)

The first partial derivative vector or the gradient (G) of

$$G=\begin{bmatrix}\frac{\partial\ell}{\partial\mu}\\ \frac{\partial\ell}{\partial\sigma}\end{bmatrix}=\begin{bmatrix}g_1\\ g_2\end{bmatrix}=\begin{bmatrix}-n\frac{\psi_\mu}{\psi}-\frac{1}{\sigma^2}(n\mu-\sum x_i)\\ -n\frac{\psi_\sigma}{\psi}-\frac{n}{\sigma}+\frac{\sum(x_i-\mu)^2}{\sigma^3}\end{bmatrix}$$

(3.2.3)

Since there is no closed form to the solution of the system G=0, it must be solved numerically.

Moreover,

(3.2.4)

$$H=\begin{bmatrix}\frac{\partial g_1}{\partial\mu} & \frac{\partial g_1}{\partial\sigma}\\ \frac{\partial g_2}{\partial\mu} & \frac{\partial g_2}{\partial\sigma}\end{bmatrix}=\begin{bmatrix}-n\frac{\psi\psi_{\mu\mu}-(\psi_\mu)^2}{\psi^2}-\frac{n}{\sigma^2} & -n\frac{\psi\psi_{\mu\sigma}-\psi_\mu\psi_\sigma}{\psi^2}+\frac{n\mu-\sum x_i}{(\sigma^2)^2}\\ -n\frac{\psi\psi_{\mu\sigma}-\psi_\mu\psi_\sigma}{\psi^2}+\frac{n\mu-\sum x_i}{(\sigma^2)^2} & -n\frac{\psi\psi_{\mu\mu}-(\psi_\sigma)^2}{\psi^2}+\frac{n}{\sigma^2}-\frac{3\sum(x_i-\mu)^2}{\sigma^4}\end{bmatrix}$$

$$\mu+\sigma\frac{\varphi(\alpha)}{1-\Phi(\alpha)}=\bar{X}$$

CC BY

Hattaway, (2010)

3.3 L – Moment method (LMOM)

L-moments are expectations of some linear combinations of orders statistics. This method is an alternative approach of moments method described by Hosking (1990) based on quantities, L-moments are measures of the location, scale, and shape of probability distributions or data samples. They are based on linear combinations of order statistics, and can be defined for any random variable if the mean exists. L-moments may be more efficient parameter estimates than the maximum likelihood estimates in some cases.

Definitions and Basic Properties

Let X be a random variable with cumulative distribution function F(x) and quantile function x(F) (the value at which the probability of the random variable is less than or equal to the given probability) , and let $X_{l:n} \le X_{2:n} \le \ldots \le X_{n:n}$ be the order statistics of a random sample of size n drawn from the distribution of X. The L-moments of X is defined as:

(3.3.1)

$$\lambda_r = r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} EX_{r-k:r}$$

So $\lambda_1 = E(X)$, $\lambda_2 = (EX_{2:2} - EX_{1:2})/2$ and $\lambda_3 = (EX_{3:3} - 2EX_{2:3} + EX_{1:3})/3$

Hosking (1990). Likewise, the first two L- moments can be written as:

(3.3.2.a)

(3.3.2.b)

Sample L- moments

$$L_2 = \frac{1}{2}\binom{n}{2}^{-1} \sum\sum_{i>j} (x_{i:n} - x_{j:n}) = \frac{1}{2}\binom{n}{2}^{-1} \sum_{i=1}^{n} \{\binom{i-1}{1} - \binom{n-j}{1}\} x_{(i)}$$

$$L_r = \binom{n}{r}^{-1} \sum_{1\le j_1<j_2\ldots<j_r\le n}\sum\ldots\sum r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} X_{i_{r-k}-n}$$

Let $x_l, x_2, \ldots, x_n$ be the sample and $x_{l:n} \le x_{2:n} \le \ldots \le x_{n:n}$ be the ordered sample, and define the r^{th} sample L-moment to be

(3.3.3)

Hosking and Wallis (2005). In particular, the first two sample moments are:

$$L_1 = n^{-1} \sum_i x_i = \binom{n}{1}^{-1} \sum_i^n x_{(i)}$$

(3.3.4.a)

(3.3.4.b)

where the second right hand is the simple form Hosking and Wallis (2005) , and $x_{(i)}$ denote the i^{th} order statistic. Indeed, the L-moments are derived from probability-weighted moments.

$$\lambda_1 = EX = \int_0^1 x(F)\,dF$$

$$\lambda_2 = \frac{1}{2} E(X_{2:2} - X_{1:2}) = \int_0^1 x(F)(2F-1)\,dF$$

CC BY

L-moments of Truncated Normal Distributions

(3.3.5)

$$\lambda_1 = \int_a^b x\ f(x)\ dx = E(x) = \mu + \frac{\varphi(\frac{a-\mu}{\sigma}) - \varphi(\frac{b-\mu}{\sigma})}{\Phi(\frac{b-\mu}{\sigma}) - \Phi(\frac{a-\mu}{\sigma})}\sigma$$

(3.3.6)

Where

$$W = \Phi(\sqrt{2}\beta) - \Phi(\sqrt{2}\alpha)$$

Finally, using L-moments method to estimate the parameters passes through the same steps of (conventional) moments. This means equating the first p sample L-moments to the corresponding population quantities of distribution that has p of unknown parameter to estimate them.

4. Simulation study

The simulation is done by taking μ = 2 and σ = 1, also when we selected the combination of truncated points

$$\lambda_2 = 2\left[\frac{-\sigma\varphi(\beta)}{z} + \frac{1}{2}\mu + \frac{\sigma W}{2\sqrt{\pi}z^2}\right] - \mu - \frac{\varphi(\alpha) - \varphi(\beta)}{z}\sigma$$

$$\lambda_2 = \frac{\sigma}{z}\left[\frac{W}{\sqrt{\pi}z} - \varphi(\alpha) - \varphi(\beta)\right]$$

we consider: First, the untruncated data percentage. Second, the symmetric properties of normal distribution. Third, presence of μ inside or outside the truncated interval or on the boundary.

4.1.a. Double sided truncation case

The combinations of the truncated points which are chosen with the approximation percentage of untruncated data are: ([-1,5] 99.7% , [0,5] 97.5% , [0,4] 95.4%, [1,5] 83.9 %, [-1,2] 49.8% , [1,4] 81.8% , [1,3] 68.2% , [0,2] 47.7% , [-1,1] 15.7% , [-1,0] 2.1% , [0,1] 13.6% , [1,2] 34.1), but to minimize the no. of tables in this paper, we choose the intervals: [0,4], [-1,2].

Also, for each of previous truncated cases, we generate 3000 samples of the following sizes (10,15,25,50,75,100,250,1000) from truncated normal distribution with parameters μ = 2 , σ = 1. In addition, the MSE and Bias were found for each truncated case and each sample size. However, to minimize the no. of tables in the paper, we choose: (10, 100, and 1000) sample sizes and all those tables are kept upon request. The initial values for the NR optimization were chosen to be the mean and the standard deviation of sample with 3000 iteration.

Table (1): MSE of μ estimations by the three methods for ([0,4], [-1,2]) double sided truncation.

Truncation	Sample n	MLE	MM	L-Moments
[0,4]	10	0.078844185	0.11233926	0.062846153
	100	0.007629521	0.01063938	0.006067499
	1000	0.000774445	0.001082562	0.000615368
[-1,2]	10	0.623941272	0.528774206	0.759702526
	100	0.61898085	0.491141071	0.726105755
	1000	0.57366177	0.48751546	0.716500249

CC BY

First of all, the collection of truncated combinations which is symmetric around the parameter of un truncated normal distribution μ [0, 4] will be called the Symmetric Truncation Combinations (STC). Those are chosen from a large no. of truncations, so we have to choose some of them, the cancelled ones are available upon request.

From Table (1), it can be observed that the three methods for estimating the parameter μ of truncated normal distribution do follow a similar behavior with preference for LMOM method especially for the truncation period [-1,2] where μ falls on the boundary rather than inside the interval.

Furthermore, the MSE value is approximately constant for the three methods when the parameter μ falls outside of the truncation interval or in one of its boundaries, with preference for MOM method on the other methods. The LMOM method falls back to last rank with greater MSE values. These methods order reappear in the combinations, which contain μ, especially for large sample sizes but with a reduction of MSE value. Moreover, for this case, the limitation of truncation level made the MSE values converge to a relatively small positive value. It should be noted that the variance of all estimations converge to zero in all combinations and for each method.

Table (2): Bias of μ estimations by the three methods for ([0,4], [-1,2]) double sided truncation.

Truncation	Sample n	MLE	MM	L-Moments
[0,4]	10	0.001278088	-0.01096151	-0.00524513
	100	-0.000997266	0.000689271	0.000571732
	1000	-0.00060565	-4.03384E-05	-3.69838E-05
[-1,2]	10	-0.73442021	-0.70149726	-0.85072192
	100	-0.784437862	-0.698149474	-0.849987368
	1000	-0.789690008	-0.697965431	-0.846251449

Table (2) shows the absolute value of Bias convergence to zero for the three methods in STC case, and the changes of methods order as the sample size increases. In contrast, it has an approximately constant value in other combination. In addition, methods give large absolute bias value when μ falls outside the interval or on its boundary. Whatever are these cases, the methods (MOM,ML,LMOM) usually have increasing order with respect to bias value.

Table (3): MSE of σ estimations by the three methods for ([0,4], [-1,2]) double sided truncation.

Truncation	Sample n	MLE	MM	L-Moments
[0,4]	10	0.057075176	0.048226928	0.0487511
	100	0.020704944	0.017172733	0.017933238
	1000	0.014965969	0.014772472	0.015057324
[-1,2]	10	0.273790575	0.26170347	0.24698446
	100	0.179679301	0.232023777	0.228854371
	1000	0.170096806	0.229616386	0.237530261

On the other hand, in estimating the parameter σ, the symmetry about μ has a limited effect when compared with the truncation level, as in table (3).

Table (4): Bias of σ estimations by the three methods for ([0,4], [-1,2]) double sided truncation.

Truncation	Sample n	MLE	MM	L-Moments
[0,4]	10	-0.147323285	-0.136267508	-0.140922
	100	-0.134473412	-0.119831558	-0.123334659
	1000	-0.121224982	-0.120386622	-0.121581876
[-1,2]	10	-0.503912617	-0.496385899	-0.477871
	100	-0.421378322	-0.480236997	-0.477071613
	1000	-0.41219841	-0.479050692	-0.487203597

Moreover, the MOM, L-MOM and MLE methods give very close values usually with preference to MLE method on them in [-1,2] case, except for the small sample size 10. Furthermore, the bias values for these methods will be close as it is clear in table (4) in STC case, and small value for MLE method in other cases.

In one-sided truncation, decreasing the untruncated data (UTD) percentage makes parameters' estimation more difficult. For these cases, the methods have an increasing order; (MOM, ML, and LMOM) with smaller MSE values.

4.1.b. One sided truncation case

For the one sided truncated case, we work with the lower truncation because the other case will be the same since the normal distribution is symmetric. Like the double truncated case, our samples size will be (10, 100, and 1000) with 3000 repetitions. However the truncated combinations with the approximating percentages of un truncated data are ([2,∞) 50% , [1,∞) 84.2%).

Table (5): MSE of μ estimations by the three methods for ([2,∞), [1,∞)) one sided truncation

Truncation	Sample n	MLE	MM	L-Moments
[2,∞)	10	0.583415579	0.537408987	0.780712691
	100	0.630024312	0.498812394	0.743207637
	1000	0.633533494	0.492565162	0.73261041
[1,∞)	10	0.141020762	0.114075008	0.190164886
	100	0.089841624	0.047042249	0.125844083
	1000	0.082912822	0.040955215	0.119175433

In addition, the absolute bias values are approximately constant over all sample sizes. However, this constant value is smaller in high UTD case, as in table (5).

Table (6): Bias of μ estimations by the three methods for ([2,∞), [1,∞)) one sided truncation

Truncation	Sample n	MLE	MM	L-Moments
[2,∞)	10	0.739916128	0.707948551	0.86149513
	100	0.791447354	0.703598041	0.859818588
	1000	0.795723727	0.701565899	0.855696723
[1,∞)	10	0.271834577	0.208471835	0.35769478
	100	0.289630027	0.199336171	0.346086442
	1000	0.286866604	0.200553827	0.344323792

For table (6), it appears that the bias is smaller for the STC case when mean falls inside the truncation interval than when it is outside the truncation interval.

CC BY

Table (7): MSE of σ estimations by the three methods for ([2,∞), [1,∞)) one sided truncation

Truncation	Sample n	MLE	MM	L-Moments
[2,∞)	10	0.26987647	0.258785744	0.240781353
	100	0.169574879	0.225438819	0.221127593
	1000	0.160645048	0.221746462	0.228416219
[1,∞)	10	0.11246286	0.100816719	0.097054063
	100	0.044533877	0.067359264	0.069161244
	1000	0.043590135	0.064265161	0.070422031

For table (7), the smallest MSE and bias values for σ came from MLE method for large sample sizes.

Table (8): Bias of σ estimations by the three methods for ([2,∞), [1,∞)) one sided truncation

Truncation	Sample n	MLE	MM	L-Moments
[2,∞)	10	-0.4956612	-0.494020884	-0.469875102
	100	-0.408459519	-0.473218289	-0.468368923
	1000	-0.400497612	-0.470745948	-0.4776699
[1,∞)	10	-0.2622055	-0.280147584	-0.26463474
	100	-0.205189025	-0.255679538	-0.258023333
	1000	-0.207969291	-0.253103366	-0.264843845

For table (8), the smallest MSE and bias values for σ came from MLE method for large sample sizes.

4.2 Estimating the mean and variance of truncated normal distribution

The mean and variance of truncated normal distribution are functions of the parameters. We know that the sample mean and the sample variance are MLEs of parameters of full normal distribution, the sample mean and the samples expressions are:

$$E_{Tr}(x) = \mu + \frac{\varphi(\frac{a-\overline{X}}{S}) - \varphi(\frac{b-\overline{X}}{S})}{\Phi(\frac{b-\overline{X}}{S}) - \Phi(\frac{a-\overline{X}}{S})} S, \quad a \le x \le b$$

(4.2.1)

$$Var_{Tr}(x) = S^2 \left[1 + \frac{\frac{a-\overline{X}}{S}\varphi(\frac{a-\overline{X}}{S}) - \frac{b-\overline{X}}{S}\varphi(\frac{b-\overline{X}}{S})}{\Phi(\frac{b-\overline{X}}{S}) - \Phi(\frac{a-\overline{X}}{S})} - \left(\frac{\varphi(\frac{a-\overline{X}}{\sigma}) - \varphi(\frac{b-\overline{X}}{\sigma})}{\Phi(\frac{b-\overline{X}}{S}) - \Phi(\frac{a-\overline{X}}{S})} \right)^2 \right]$$

(4.2.2)

Where the $E_{tr}(x)$ and Var_{Tr} (x) denote the mean and variance of truncated normal distribution respectively. Now, the simulation study is done by first, generating 3,000 samples from full normal distribution then substituting the samples mean and variance in (4.2.1) and (4.2.2) formulas to compute the MLEs. Second, it is done by censoring each sample by the truncated point and calculating the new sample mean and variance. The third stage is generating 3000 samples from truncated normal distribution with same parameters of the previous full distribution and calculating the sample mean and variance again.

CC BY

Table (9): MSE of mean estimations by the three methods for ([0,4], [-1,2]) double sided truncation

Truncation	Sample n	MLE	CS	TS
[0,4]	10	0.064894668	0.091367252	0.076431222
	100	0.005951892	0.009040314	0.008094043
	1000	0.00059569	0.000910166	0.000747738
[-1,2]	10	0.039609673	0.18800652	0.661327994
	100	0.003803323	0.156834895	0.627883698
	1000	0.000377413	0.154431563	0.626433735

In fact, the MLE method is almost perfect to estimate the mean of truncated normal distribution as shown in table (9), since the MSE values converge to zero in all combinations, while the TS method gives smaller MSE values more than the censored sample (CS) method in the STC cases.

Table (10): Bias of mean estimations by the three methods for ([0,4], [-1,2]) double sided truncation.

Truncation	Sample n	MLE	CS	TS
[0,4]	10	0.000984042	0.001823657	0.003896737
	100	-0.000669542	-0.000860404	-0.000210676
	1000	-0.000208852	-0.000233462	0.000244985
[-1,2]	10	0.055102736	0.392842199	-0.00068271
	100	0.004999822	0.391737378	0.000937487
	1000	0.000374846	0.392547498	-0.000101799

From table (10), it was observed that the absolute bias of CS method converges to zero only when the combination belongs to STC. For the interval [-1,2], the MLE and TS have smaller bias than the CS case. In fact, it must be stressed here that this increase in the values of the absolute bias accompanied by the steep decline in variance values may actually improve the CS method with small MSE values. This may give an unrealistic preference to this method.

Table (11): MSE of variance estimations by the three methods for ([0,4], [-1,2]) double sided truncation

Truncation	Sample n	MLE	CS	TS
[0,4]	10	0.048869436	0.066314036	0.04994314
	100	0.003380908	0.037043022	0.017454641
	1000	0.000336549	0.034833755	0.014717248
[-1,2]	10	0.017046482	0.077678673	0.209119558
	100	0.001894573	0.056798243	0.172252346
	1000	0.000191317	0.055271715	0.16882872

For table (11) as expected, the ML method showed a very high effectiveness in estimating the variance, especially for large samples size. On the other hand, as in estimating mean case, the TS method has a poor performance to estimate the variance as UTD reduced, unlike the CS method.

CC BY

Table (12) Bias of variance estimations by the three methods for ([0,4], [-1,2]) double sided truncation.

Truncation	Sample n	MLE	CS	TS
[0,4]	10	-0.110134421	0.163243709	0.08791553
	100	-0.011442747	0.183834468	0.105000797
	1000	-0.000835708	0.185729231	0.106001741
[-1,2]	10	-0.043075096	0.201278743	0.21996201
	100	-0.004311321	0.230875408	0.239982338
	1000	-0.000329394	0.234368896	0.241955167

For Table (12), we have smaller bias for the MLE method that converging to zero while constant for TS and CS cases.

Table (13): MSE of mean estimations by the three methods for (-∞,3], (-∞,2]) one sided truncation.

Truncation	Sample n	MLE	CS	TS
(-∞,3]	10	0.064042929	0.11589587	0.146913094
	100	0.006193276	0.048933165	0.089472169
	1000	0.000600189	0.042478075	0.083405473
(-∞,2]	10	0.046853001	0.194081351	0.678700041
	100	0.004507811	0.162668824	0.643801233
	1000	0.000461222	0.159227624	0.637276802

The MSE values in table (13), showed that the three methods follow a similar behavior in the one sided estimation for the mean and variance. Definitely, with a preference for the method of MLE in the first rank then followed by CS case, finally the method of TS with very high MSE values, especially when the values of UTD are low, but as mentioned before the performance of CS method may be not actual.

Table (14): Bias of mean estimations by the three methods for (-∞,3], (-∞,2]) one sided truncation.

Truncation	Sample n	MLE	CS	TS
(-∞,3]	10	0.02329851	0.204473858	-0.003088
	100	0.002195019	0.204067468	-0.00048805
	1000	-5.59471E-05	0.204310292	-0.000121911
(-∞,2]	10	0.053815821	0.40017565	-0.00362762
	100	0.005588975	0.399144802	-0.0022248
	1000	0.000186358	0.398607621	-0.000185143

For Table (14), it reveals that the absolute mean bias for the MLE methods always converges to zero, as well as in the TS method. However, the CS method has a constant value.

Table (15): MSE of variance estimations by the three methods for (-∞,3], (-∞,2]) one sided truncation.

Truncation	Sample n	MLE	CS	TS
(-∞,3]	10	0.05596121	0.084428179	0.086238321
	100	0.005666458	0.058663281	0.046617141
	1000	0.000549254	0.056700119	0.042937556
(-∞,2]	10	0.02886035	0.07152637	0.200291426
	100	0.003123734	0.050771211	0.160670881
	1000	0.000320742	0.048850456	0.157863616

Finally, the MSE values in table (15), showed that the three methods follow a similar behavior in the one sided estimation for the mean and variance. Definitely, with a preference for the method of MLE in the first rank then followed by CS case, finally the method of TS with very high MSE values, especially when the values of UTD are low, but as mentioned before the performance of CS method may be not actual.

Table (16): Bias of variance estimations by the three methods for (-∞,3], (-∞,2]) one sided truncation.

Truncation	Sample n	MLE	CS	TS
(-∞,3]	10	-0.0566172	0.2181383	0.140063004
	100	-0.004401812	0.235078148	0.161860936
	1000	-3.52212E-05	0.23742005	0.163859673
(-∞,2]	10	-0.0352145	0.18067277	0.219032508
	100	-0.003704517	0.216936927	0.239086217
	1000	-5.11753E-05	0.220152651	0.239622248

For Tables (16), we have smaller bias for the MLE method that converging to zero while constant for TS and CS cases.

5. Conclusion

The classical assumption is that a population may be a normal distribution, that is, it takes values between negative and positive infinity, which cannot actually happen. This is partially motivated by the statistical community to work on the truncated normal distribution.

Looking at the tables for one and double sided estimators, the best mean estimator is the MLE-estimator. In fact, it is better than the TS case itself, particularly, when the UTD is small. In this case, the TS case performs poorly. Although it maintains low bias values, it is accompanied by very high variance values. Quite the contrary, the CS method shows a performance that seems relatively good, but it cannot be trusted because it is accompanied by high bias values. Excluded from this the cases when the combinations belong to STC where the MSE and bias of the mean in all methods converge to zero. The three methods take an ascending order from MLE method to CS case. It is noted also, that we can say that the three methods (MLE, MM and L-MM) perform well when the mean μ falls inside truncation period (STC) than when the mean falls on the boundary or outside the truncation period. Actually, it is recommended to use the MLE method to estimate the variance of truncated normal distribution in the first order, followed by the TS case only in cases where μ belongs to UTD. That is not only for one-sided truncated cases, but also for double truncated cases.

On the other hand, it is preferred to use LMOM method to estimate μ in STC cases, followed by MLE method at the second rank. Moreover, the poor performance of the LMOMT appears not only in cases, which do not belong to STC but also in cases of one-sided truncation.

MOM method gains the advantage of being used in estimating μ not only in the cases that do not belong to STC, but also in one-sided truncation cases. While MLE method is the second best in use. It is in forefront at large UTD and sample sizes.

In estimating the parameter σ, the three methods are very similar to STC. Otherwise, it is preferred to use the MLE method approximately in all truncation cases, followed by the LMOM and MOM respectively, except for some special combinations. Finally, it becomes more difficult to estimate the parameters whenever the UTD decreases.

REFERENCES

1. Barr, D. R. and E. T. Sherrill (1999). "Mean and variance of truncated normal distributions." The American Statistician53(4): 357-361. https://doi.org/10.1080/00031305.1999.10474490

2. Bebu, I. and T. Mathew (2009). "Confidence intervals for limited moments and truncated moments in normal and lognormal models." Statistics & Probability Letters 79(3): 375-380. https://doi.org/10.1016/j.spl.2008.09.006

3. Bookbinder, J. H. and A. E. Lordahl (1989). "Estimation of inventory re-order levels using the bootstrap atistical procedure." IIE transactions 21(4): 302-312. https://doi.org/10.1080/07408178908966236

4. Casella, G. and R. L. Berger (2002). "Statistical inference." Duxbury Pacific Grove, CA.

5. Cohen, A. C. (1957). "On the solution of estimating equations for truncated and censored samples from normal populations." Biometrika 44(1/2): 225-236. https://doi.org/10.1093/biomet/44.1-2.225

6. Cohen, A. C. (1959). "Simplified estimators for the normal distribution when samples are singly censored or truncated." Technometrics 1(3): 217-237. https://doi.org/10.1080/00401706.1959.10489859

7. Cohen, A. C. (1961). "Tables for maximum likelihood estimates: singly truncated and singly censored samples." Technometrics 3(4): 535-541. https://doi.org/10.1080/00401706.1961.10489973

8. Cohen, A. C. and J. Woodward (1953). "Tables of Pearson-Lee-Fisher functions of singly truncated normal distributions." Biometrics 9(4): 489-497. https://doi.org/10.2307/3001440

9. Galton, F. (1897). "An examination into the registered speeds of American trotting horses, with remarks on their value as hereditary data." Proceedings of the Royal Society of London 62(379-387): 310-315. https://doi.org/10.1098/rspl.1897.0115

10. Gupta, A. K. (1952). "Estimation of the mean and standard deviation of a normal population from a censored sample." Biometrika 39(3/4): 260-273. https://doi.org/10.1093/biomet/39.3-4.260

11. Hattaway, J. T. (2010). "Parameter estimation and hypothesis testing for the truncated normal distribution with applications to introductory statistics grades."

12. Hegde, L. M. and R. C. Dahiya (1989). "Estimation of the parameters in a truncated normal distribution." Communications in Statistics-Theory and Methods 18(11): 4177-4195. https://doi.org/10.1080/03610928908830149

13. Hosking, J. R. M. (1990). "L-moments: analysis and estimation of distributions using linear combinations of order statistics." Journal of the Royal Statistical Society. Series B (Methodological): 105-124. https://doi.org/10.1111/j.2517-6161.1990.tb01775.x

14. Hosking, J. R. M. and J. R. Wallis (2005). Regional frequency analysis: an approach based on L-moments, Cambridge University Press.

15. Iwueze, I. S. (2007). "Some Implications of Truncating the N (1,σ2) Distribution to the Left at Zero." Journal of Applied Sciences 7(2): 189-195.https://doi.org/10.3923/jas.2007.189.195

16. Johnson, N. L., S. Kotz, et al. (1994). "Continuous Univariate Probability Distributions.",(Vol. 1), John Wiley & Sons Inc., NY.

17. Johnson, A. C. (2002). "Characteristics and tables of the doubly-truncated normal distribution." Proc. of POM High Tech., 2002. Johnson, A. C. and N. T. Thomopoulos (2002). "Characteristics and tables of the left-truncated normal distribution." Proceedings of the Midwest Decision Sciences: 133-139.

18. Johnson, A. C. and N. T. Thomopoulos (2002). Use of the left-truncated normal distribution for improving achieved service levels. Proceedings of the 2002 Annual Meeting of the Decision Sciences Institute.

19. Khasawneh, M. T., S. R. Bowling, et al. (2004). "Tables of a truncated standard normal distribution: a singly truncated case." Quality Engineering 17(1): 33-50.https://doi.org/10.1081/QEN-200028681

20. Khasawneh, M. T., S. R. Bowling, et al. (2005). "Tables of a truncated standard normal distribution: A doubly truncated case." Quality Engineering 17(2): 227-241. https://doi.org/10.1081/QEN-200057321

21. Lee, A. (1914). "Table of the Gaussian" Tail" Functions; When the" Tail" is Larger than the Body." Biometrika 10(2/3): 208-214. https://doi.org/10.1093/biomet/10.2-3.208

22. Lifsey, J. D. (1965). "Auxiliary estimating functions for doubly truncated normal samples."

23. Mittal, M. M. and R. C. Dahiya (1987). "Estimating the parameters of a doubly truncated normal distribution: Estimating the parameters." Communications in Statistics-Simulation and Computation 16(1): 141-159.
https://doi.org/10.1080/03610918708812582

24. Pearson, K. (1902). "On the systematic fitting of curves to observations and measurements." Biometrika 1(3): 265-303. https://doi.org/10.1093/biomet/1.3.265

25. Pearson, K. and A. Lee (1908). "On the generalised probable error in multiple normal correlation." Biometrika 6(1): 59-68. https://doi.org/10.1093/biomet/6.1.59

26. Shah, S. M. and M. C. Jaiswal (1966). "Estimation of parameters of doubly truncated normal distribution from first four sample moments." Annals of the Institute of Statistical Mathematics 18(1): 107-111. https://doi.org/10.1007/BF02869520

27. Sinha, D. (1991). "Optimal policy estimation for continuous review inventory systems." Computers & operations research 18(5): 487-496. https://doi.org/10.1016/0305-0548(91)90025-M

28. Sugiura, N. and A. Gomi (1985). "Pearson diagrams for truncated normal and truncated Weibull distributions." Biometrika 72(1): 219-222. https://doi.org/10.1093/biomet/72.1.219