

Mean Difference of Truncated Normal Distribution

Giovanni Girone¹, Antonella Massari², Fabio Manca³, Claudia Marin^{3*}

¹Emeritus Professor, University of Bari Aldo Moro, Bari, Italy

²Department of Economics, Management and Business Law, University of Bari Aldo Moro, Bari, Italy

³Department of Education, Psychology, Communication, University of Bari Aldo Moro, Bari, Italy

Email: *claudia.marin@uniba.it

How to cite this paper: Girone, G., Massari, A., Manca, F. and Marin, C. (2020) Mean Difference of Truncated Normal Distribution. *Applied Mathematics*, 11, 1162-1166.

<https://doi.org/10.4236/am.2020.1111078>

Received: October 16, 2020

Accepted: November 17, 2020

Published: November 20, 2020

Copyright © 2020 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The purpose of this paper is to broaden the knowledge of mean difference and, in particular, of an important distribution model known as truncated normal distribution, which is widely used in applied sciences and economics. In this work, we obtained the general formula of mean difference, which is not yet reported in literature, for the aforementioned distribution model and also for particular truncated cases.

Keywords

Mean Difference, Truncated Normal Distribution, Variability Indexes, Economic Sciences

1. Introduction

Truncated normal distribution is a part of the important continuous distribution models that find wide application in various fields of scientific research (Johnson *et al.*, 1994 [1]; Hamedani, 2019 [2]). Said distribution represents normal variables observed on a section of an abscissas plan, more precisely on the right half axis $X > a$, on the left half axis $X < b$ or on the segment $a < X < b$. A number of publications have dealt with the overview of the truncated normal distributions (Barr and Sherrill, 1999 [3]; Domma, 2003 [4]; Cha *et al.*, 2013 [5]; Koutroumbas *et al.*, 2014 [6]; Thomopoulos, 2015 [7]). Said distribution has important uses in applied sciences, in particular in economic sciences. This distribution has many shapes and a measure of interest is the coefficient of variation, that helps to identify the shape of the distribution. The mean difference of the truncated normal distribution (Domma *et al.*, 2014 [8]) is not known in literature, while variability indexes as range of variation, mean deviation and stan-

dard deviation and shape indexes (asymmetry and disnormality) have been already obtained. The purpose of this short note is to fill the said gap.

2. Truncated Normal Distribution

Density function of standardized normal distribution is

$$\phi(x) = \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}}, \quad -\infty < x < \infty, -\infty < \mu < \infty, \sigma > 0. \quad (1)$$

Cumulative distribution function of normal distribution is

$$\Phi(x) = \int_{-\infty}^x \phi(u) du. \quad (2)$$

We can consider the case of $\mu = 0$ and $\sigma = 1$, namely, of standardized normal distribution. Density function of truncated normal distribution between a and b corresponds to the formula

$$\varphi(x) = \frac{\phi(x)}{\Phi(b) - \Phi(a)}, \quad a < x < b. \quad (3)$$

It is clearly obtained

$$\varphi(x) = \frac{e^{-\frac{x^2}{2}}}{\int_a^b e^{-\frac{x^2}{2}} dx}, \quad a < x < b. \quad (4)$$

Cumulative distribution function of truncated normal distribution is then

$$\Psi(x) = \frac{\int_a^x e^{-\frac{x^2}{2}} dx}{\int_a^b e^{-\frac{x^2}{2}} dx}. \quad (5)$$

3. General Formula of the Mean Difference of Truncated Normal Distribution

Various formulas are available to calculate the mean difference of a continuous distribution, the simplest one uses the cumulative distribution function $F(x)$

$$\Delta = \int_a^b 2F(x)[1 - F(x)] dx, \quad (6)$$

in which a and b are extreme values of the range of distribution. Dealing with truncated normal distribution, it is necessary to use the cumulative distribution function $\Psi(x)$:

$$\Delta = \int_a^b 2\Psi(x)[1 - \Psi(x)] dx. \quad (7)$$

By means of heavy integrations and simplifications, the following general formula of the mean difference of truncated normal distribution is obtained

$$\Delta = \frac{2 \left\{ 2 \left[\operatorname{erf}(b) - \operatorname{erf}(a) \right] - \sqrt{2} \left[e^{-\frac{a^2}{2}} + e^{-\frac{b^2}{2}} \right] \left[\operatorname{erf}\left(\frac{b}{\sqrt{2}}\right) - \operatorname{erf}\left(\frac{a}{\sqrt{2}}\right) \right] \right\}}{\sqrt{\pi} \left[\operatorname{erf}\left(\frac{b}{\sqrt{2}}\right) - \operatorname{erf}\left(\frac{a}{\sqrt{2}}\right) \right]}, \quad (8)$$

in which the error function occurs

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt. \quad (9)$$

The complement of error function is

$$\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt = 1 - \operatorname{erf}(x). \quad (10)$$

In the following **Figure 1**, the mean difference Δ is shown on vertical axis depending on the minimum value a and the maximum value b indicated on base axes.

As can easily be seen, starting from 0, when a and b are both zeros, the mean difference grows with a decreasing and b increasing up to the maximum, equal to $\frac{2}{\sqrt{\pi}}$, which is obtained when $a = -\infty$ and $b = \infty$, that is the case of complete normal distribution.

4. Special Cases

Let us consider mean difference formulas in some interesting special cases of truncated normal distribution.

4.1. Symmetric Truncated Distribution

When $b = -a$ in the general formula, after a few steps it leads to

$$\Delta = \frac{2\sqrt{2}e^{-\frac{a^2}{2}} \operatorname{erf}\left(\frac{a}{\sqrt{2}}\right) - 2\operatorname{erf}(a)}{2\sqrt{\pi} \operatorname{erf}\left(\frac{a}{\pi}\right)^2}. \quad (11)$$

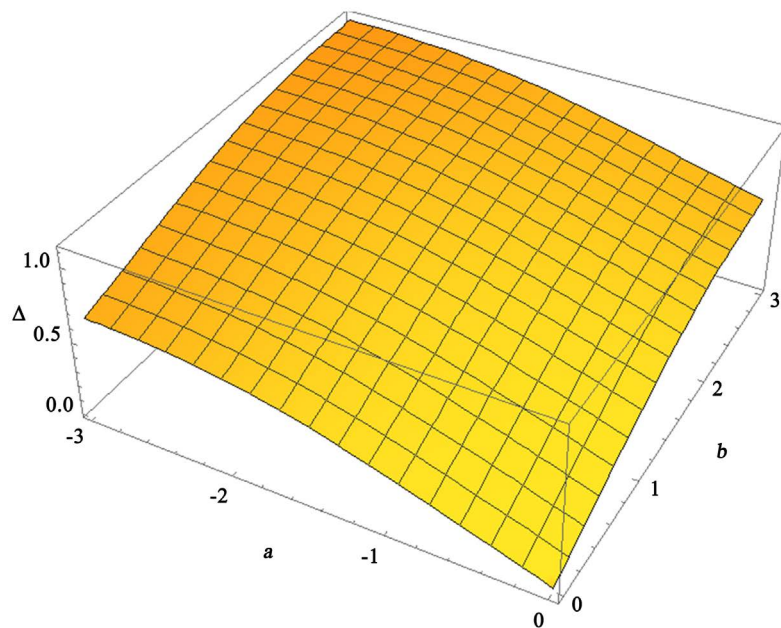


Figure 1. Mean difference of truncated normal distribution.

4.2. Truncated Semi-Normal Distribution

When $a = 0$ in the general formula, after a few steps it leads to

$$\Delta = \frac{2 \left[2\operatorname{erf}(b) - \sqrt{2}\operatorname{erf}\left(\frac{b}{\sqrt{2}}\right) \left(1 + e^{-\frac{b^2}{2}}\right) \right]}{2\sqrt{\pi}\operatorname{erf}\left(\frac{b}{\sqrt{2}}\right)}. \quad (12)$$

When $b = 0$, it leads to a similar result

$$\Delta = \frac{2 \left[\sqrt{2}\operatorname{erf}\left(\frac{a}{\sqrt{2}}\right) \left(1 + e^{-\frac{a^2}{2}}\right) - 2\operatorname{erf}(a) \right]}{2\sqrt{\pi}\operatorname{erf}\left(\frac{a}{\sqrt{2}}\right)}. \quad (13)$$

4.3. Semi-Normal Distribution

When $a = 0$ and $b = \infty$ in the general formula, after a few steps it leads to

$$\Delta = \frac{2(2 - \sqrt{2})}{\sqrt{\pi}}. \quad (14)$$

The same result is obtained when $a = -\infty$ and $b = 0$.

4.4. Tail Normal Distribution

When $a > 0$ and $b = \infty$ in the general formula, after a few steps it leads to

$$\Delta = \frac{4 \left[1 - \operatorname{erf}(a) \right] - 2\sqrt{2}e^{-\frac{a^2}{2}} \left[1 - \operatorname{erf}\left(\frac{a}{\sqrt{2}}\right) \right]}{\sqrt{\pi} \left[1 - \operatorname{erf}\left(\frac{a}{\sqrt{2}}\right) \right]^2}. \quad (15)$$

The same result is obtained with $a = -\infty$ and $b < 0$ in the general formula:

$$\Delta = \frac{4 \left[1 + \operatorname{erf}(b) \right] - \sqrt{2}e^{-\frac{b^2}{2}} \left[1 + \operatorname{erf}\left(\frac{b}{\sqrt{2}}\right) \right]}{\sqrt{\pi} \left[1 + \operatorname{erf}\left(\frac{b}{\sqrt{2}}\right) \right]^2}. \quad (16)$$

4.5. Complete Normal Distribution

When $a = -\infty$ and $b = \infty$, it leads to the well-known result of mean difference of complete normal distribution.

$$\Delta = \frac{2}{\sqrt{\pi}}. \quad (17)$$

In this work, we do not deal with other special cases.

5. Conclusion

Truncated normal distribution is an important distribution model that can be

applied in many experimental and observational fields, particularly in quantitative economic sciences. Variability indexes as range of variation, mean deviation and standard deviation and shape indexes (asymmetry and disnormality) have been already obtained, but the mean difference of the truncated normal distribution is not known in literature. This note fills the gap regarding the lack of knowledge of mean difference general formula of said model. Moreover, the obtained formula is proposed for any truncation and for some particular truncated cases.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1994) Continuous Univariate Distributions. Wiley, New York, Vol. 1, Cap. 10, Par. 1.
- [2] Hamedani, G.G. (2019) Characterizations of Recently Introduced Univariate Continuous Distributions II. Nova Science Publishers, New York.
- [3] Barr, D.R. and Sherrill, E.T. (1999) Mean and Variance of Truncated Normal Distributions. *The American Statistician*, **53**, 357-361.
<https://doi.org/10.1080/00031305.1999.10474490>
- [4] Domma, F. (2003) Informazione di Fisher e modellitroncati. *Statistica*, **63**, 267-284.
- [5] Cha, J., Cho, B.R. and Sharp, J.L. (2013) Rethinking the Truncated Normal Distribution. *International Journal of Experimental Design and Process Optimisation*, **3**, 27-63. <https://doi.org/10.1504/IJEDPO.2013.059667>
- [6] Koutroumbas, K.D., Themelis, K.E. and Rontogiannis, A.A. (2014) Approximating the Mean of a Truncated Normal Distribution, Cornell University.
- [7] Thomopoulos, N.T. (2015) Standard Normal and Truncated Normal Distributions. In: *Demand Forecasting for Inventory Control*, Springer International Publishing, New York, 137-148.
https://doi.org/10.1007/978-3-319-11976-2_10
- [8] Domma, F. and Hamedani, G.G. (2014) Characterizations of a Class of Distributions by Dual Generalized Order Statistics and Truncated Moments. *Journal of Statistical Theory and Applications*, **13**, 222-234.
<https://doi.org/10.2991/jsta.2014.13.3.4>