Journal of Progressive Research in Mathematics(JPRM)
ISSN: 2395-0218

SCITECH
RESEARCH ORGANISATION

Volume 15, Issue 1
Published online: April 23, 2019|

Journal of Progressive Research in Mathematics
www.scitecresearch.com/journals

An alternative one parameter triangular distribution

Eric Batista FERREIRA[1], Rafael Lemos BASTOS[2], Marcela Costa ROCHA[3]

[1]Statistics Department, Federal University of Alfenas, Unifal-MG, Brazil
[2]Federal Institute of Education, Science and Technology of South of Minas Gerais – IFSULDEMINAS campus Passos
[3]Federal Institute of Education, Science and Technology of South of Minas Gerais – IFSULDEMINAS campus Machado

Abstract

Triangular distributions are well known but not so widely studied in the specialized literature. Despite of having few applications, triangular distributions are useful in finances, econometrics, risk analysis and hydraulics. However, triangular distributions are usually conceived using two or more parameters, what gives them flexibility but turns them less parsimonious, mainly for small samples. For that reason, this paper brings a new alternative of parsimonious triangular distribution.

Key words: Triangular model, parsimony, probability density function, moments.

1.Introduction

Triangular distributions are well known but not so widely studied in the specialized literature ([7], [8], [9], [3]). Although it has few applications, triangular distributions are useful in finances, econometrics, risk analysis ([1], [6]) and hydraulics ([10], [2]).

In econometrics, [6] states that when a time series is not available for variables that make up the cash ow, one can use the minimum, the average and the maximum values for associating those variables to a triangular distribution.

On the other hand, [8] say that many forms of risk analysis incorporate appropriate probability distributions to represent the uncertainty surrounding key parameters in the analysis. As many uncertain quantities have identifiable minimum and maximum values, a distribution with finite limits is intuitively plausible to non-statistically minded decision makers. The beta distribution is often seen as a suitable model in this context as long as it provides a wide variety of distributional shapes over a finite interval. It is therefore one of the most versatile of all the standard distributions used in risk analysis. Other commonly used distribution with similar properties is the triangular distribution, which in the case of right-triangular distributions, is a special case of the beta distribution.

In hydraulics, [10] states that variables like discharge coefficient and width reduction of pipes caused by debris follow triangular distributions, and that their minimum, modal and maximum observed values are needed to estimate the suitable parameters. On the other hand, [2] simulated evapotranspiration data from Gaussian and triangular distributions for coffee growing.

However, triangular distributions are usually modeled using two or more parameters, what gives then flexibility but turns them less parsimonious, requiring large samples. For that reason, this paper aims to propose an alternative one parameter triangular distribution.

In section 2 we present both the triangular model proposed in this paper and the standard triangular one. We present also several moments and other properties of the new model.

An application to real data is presented in section 3, highlighting the important of a model, mainly for being parsimonious.

2 The proposed triangular model

In this paper we propose a parsimonious triangular model following the probability density function:

$$f(X) = (1 - \beta + X)I_{[\beta-1;\beta]}(X) + (1 + \beta - X)I_{(\beta;\beta+1]}(X) \tag{1}$$

where $-\infty < \beta < \infty$ is the location parameter.

The proposed model can be written in the standard triangular form, and is a particular case of the general triangular distribution,

$$f(X) = \frac{2(X-a)}{(c-a)(b-a)}I_{[a;b]}(X) + \frac{2(c-X)}{(c-a)(c-b)}I_{(b,c]}(X) \tag{2}$$

setting $a = \beta - 1$, $b = \beta$ and $c = \beta + 1$. Figure 1 shows the probability density function for four possible values of the location parameter, $\beta \in \{-1, 0, 1, 2\}$.

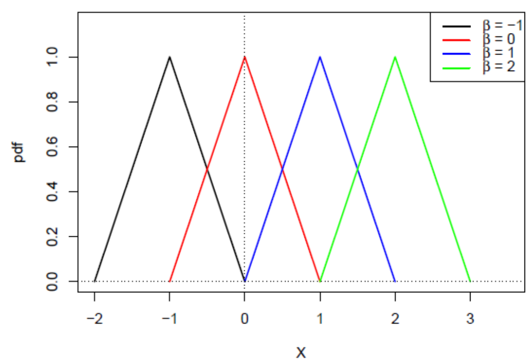

Figure 1: Proposed triangular probability density function.

[11] presented a particular case ($\beta = 1$) of triangular distribution proposed in this work. Note that when $\beta = 1$, the triangular distribution is given by the sum of two independent uniform random variables in the interval [0,1].

Theorem 1 *The cumulative distribution function of (1) is given by*

$$F(x) = \left(\frac{x^2}{2} + x(1-\beta) + \frac{\beta^2}{2} - \beta + \frac{1}{2}\right)I_{[\beta-1;\beta]}(X) + \left(-\frac{x^2}{2} + x(1+\beta) - \frac{\beta^2}{2} - \beta + \frac{1}{2}\right)I_{(\beta;\beta+1]}(X).$$

Proof. If $x \leq \beta$, the cumulative function becomes

$$\begin{aligned} F(X) &= \int_{\beta-1}^{x}(1-\beta+x)dx \\ &= \frac{x^2}{2} + x(1-\beta) + \frac{\beta^2}{2} - \beta + \frac{1}{2} \end{aligned} \tag{3}$$

If $x > \beta$, the cumulative function becomes

$$\begin{aligned} &= \int_{\beta-1}^{\beta}(1-\beta+x)dx + \int_{\beta}^{x}(1+\beta-x)dx \\ &= -\frac{x^2}{2} + x(1+\beta) - \frac{\beta^2}{2} - \beta + \frac{1}{2} \end{aligned} \tag{4}$$

From (3) and (4) gives

$$F(x) = \left(\frac{x^2}{2} + x(1-\beta) + \frac{\beta^2}{2} - \beta + \frac{1}{2}\right) I_{[\beta-1;\beta]}(X) + \left(-\frac{x^2}{2} + x(1+\beta) - \frac{\beta^2}{2} - \beta + \frac{1}{2}\right) I_{(\beta;\beta+1]}(X).$$

□

Since the segmented density function is linear it is straightforward to derive its cumulative density function (cdf). Figure 2 shows the cumulative density function for four possible values of $\beta \in \{-1, 0, 1, 2\}$.

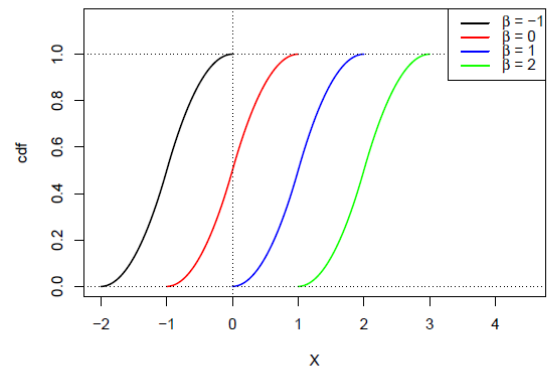

Figure 2: Proposed triangular probability density function.

2.1 Moments and other properties

Its population moments (mean, variance, skewness and kurtosis) were derived as well as its coefficients of skewness and kurtosis, moment generating function, cumulative density function, mean deviation the ration between the mean deviation and the standard deviation.

Theorem 2 *The expectation of (1) is given by*

$$E[X] = \beta. \tag{5}$$

Proof.

$$\begin{aligned} E[X] &= \int_D x f(x) dx \\ &= \int_{\beta-1}^{\beta} x(1-\beta+x)dx + \int_{\beta}^{\beta+1} x(1+\beta-x)dx \\ &= \beta. \end{aligned}$$

□

Theorem 3 *The variance of (1) is given by*

$$Var[X] = \frac{1}{6}.$$

Proof.

$$\begin{aligned}
Var[X] &= \int_D (x - E[X])^2 f(x)dx \\
&= \int_{\beta-1}^{\beta} (x-\beta)^2(1-\beta+x)dx + \int_{\beta}^{\beta+1} (x-\beta)^2(1+\beta-x)dx \\
&= \int_{\beta-1}^{\beta} (x^2 - 3\beta x^2 + x^3 - 2\beta x + 3\beta^2 x + \beta^2 - \beta^3)dx \\
&+ \int_{\beta}^{\beta+1} (x^2 + 3\beta x^2 - x^3 - 2\beta x - 3\beta^2 x + \beta^2 + \beta^3)dx \\
&= \frac{1}{6}.
\end{aligned}$$

□

Theorem 4 *The third population central moment of (1) is given by*

$$\mu_3 = 0.$$

Proof.

$$\begin{aligned}
\mu_3 &= \int_D (x - E[X])^3 f(x)dx \\
&= \int_{\beta-1}^{\beta} (x-\beta)^3(1-\beta+x)dx + \int_{\beta}^{\beta+1} (x-\beta)^3(1+\beta-x)dx \\
&= \int_{\beta-1}^{\beta} \left[x^4 + (1-4\beta)x^3 + (6\beta^2 - 3\beta)x^2 + (3\beta^2 - 4\beta^3)x - \beta^3 + \beta^4\right] dx \\
&+ \int_{\beta}^{\beta+1} \left[-x^4 + (1+4\beta)x^3 - (6\beta^2 + 3\beta)x^2 + (3\beta^2 + 4\beta^3)x - \beta^3 - \beta^4\right] dx \\
&= 0.
\end{aligned}$$

□

Theorem 5 *The fourth population central moment of (1) is given by*

$$\mu_4 = \frac{1}{15}.$$

Proof.

$$\begin{aligned}
\mu_4 &= \int_D (x - E[X])^4 f(x)dx \\
&= \int_{\beta-1}^{\beta} (x-\beta)^4(1-\beta+x)dx + \int_{\beta}^{\beta+1} (x-\beta)^4(1+\beta-x)dx \\
&= \int_{\beta-1}^{\beta} \left[x^5 + (1-5\beta)x^4 + (-4\beta + 10\beta^2)x^3 + (6\beta^2 - 10\beta^3)x^2 + (5\beta^4 - 4\beta^3)x + \beta^4\right] dx \\
&+ \int_{\beta}^{\beta+1} \left[-x^5 + (1+5\beta)x^4 - (4\beta + 10\beta^2)x^3 + (6\beta^2 + 10\beta^3)x^2 - (4\beta^3 + 5\beta^4)x + \beta^4\right] dx \\
&= \frac{1}{15}.
\end{aligned}$$

□

Theorem 6 *The coefficient of asymmetry of (1) is given by*

$$\sqrt{\beta_1} = 0.$$

Proof.

$$\sqrt{\beta_1} = \frac{\mu_3}{(\mu_2)^{\frac{3}{2}}} = 0$$

□

As the value $\sqrt{\beta_1}$ equal zero, such distribution is obviously symmetric.

Theorem 7 *The kurtosis coefficient of (1) is given by*

$$\beta_2 = 2.4$$

As $\beta_2 < 3$, such distribution is placticurtic.

Proof.

$$\beta_2 = \frac{\mu_4}{(\mu_2)^2} = 2.4$$

□

Theorem 8 *The moments generating function of (1) is given by*

$$M_x(t) = \frac{e^{t\beta}(-2 + e^{-t} + e^t)}{t^2}.$$

Proof.

$$\begin{aligned}
M_x(t) &= \int_D e^{tx} f(x)dx \\
&= \int_{\beta-1}^{\beta} e^{tx}(1-\beta+x)dx + \int_{\beta}^{\beta+1} e^{tx}(1+\beta-x)dx \\
&= \left[\frac{e^{tx}}{t}\right]_{\beta-1}^{\beta} - \left[\beta\frac{e^{tx}}{t}\right]_{\beta-1}^{\beta} + \left[\frac{xe^{tx}}{t} - \frac{e^{tx}}{t^2}\right]_{\beta-1}^{\beta} + \left[\frac{e^{tx}}{t}\right]_{\beta}^{\beta+1} + \left[\beta\frac{e^{tx}}{t}\right]_{\beta}^{\beta+1} - \left[\frac{xe^{tx}}{t} - \frac{e^{tx}}{t^2}\right]_{\beta}^{\beta+1} \\
&= \frac{e^{t\beta}(-2 + e^{-t} + e^t)}{t^2}.
\end{aligned}$$

□

Theorem 9 *The mean deviation of (1) is given by*

$$MD = \frac{1}{3}.$$

Proof.

$$\begin{aligned}
MD &= \int_D |x - E[x]| f(x)dx \\
&= \int_{\beta-1}^{\beta} |x-\beta|(1-\beta+x)dx + \int_{\beta}^{\beta+1} |x-\beta|(1+\beta-x)dx \\
&= \int_{\beta-1}^{\beta} -(x-\beta)(1-\beta+x)dx + \int_{\beta}^{\beta+1} (x-\beta)(1+\beta-x)dx \\
&= \int_{\beta-1}^{\beta} (-x + 2x\beta - x^2 + \beta - \beta^2)dx + \int_{\beta}^{\beta+1} (x + 2x\beta - x^2 - \beta - \beta^2)dx \\
&= \frac{1}{3}
\end{aligned}$$

□

3 Estimators of β

[4] states that, for standard triangular distribution (eq. 2) "maximum likelihood estimators for a, b, c do not have closed form. But we can maximize the log-likelihood numerically. Furthermore, moment based estimators have to be computed numerically solving the system of sample moments and theoretical ones."

Indeed, we could not solve the equation derived from the maximum likelihood (section 3.3). On the other hand, the moment based estimator is straightforward obtaining for the proposed distribution (section 3.1).

Let $X_1, X_2, ..., X_n$ be a random sample of the distribution (1) with parameter β. Next, the estimators for β will be obtained.

3.1 Method of moments

Theorem 10 *The moments based estimator for β is given by*

$$\hat{\beta} = \bar{x}$$

Proof. Since the first non-central population moment was derived in (5)

$$\mu_1' = E[X] = \beta$$

and the correspondent sample moment is

$$m_1' = \frac{\sum_{i=1}^{n} X_i}{n} = \bar{X},$$

we automatically obtain the estimator of β setting them to each other. Then,

$$\mu_1' = m_1' \Rightarrow \hat{\beta} = \bar{X}.$$

□

3.2 Least squares method

Theorem 11 *The least squares estimator for β is given by*

$$\hat{\beta} = \bar{x}$$

Proof. Using again the first population moment (5), we first state the residual sum of squares (Z) for the proposed model:

$$Z = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (X - E[X])^2 = \sum_{i=1}^{n} (X - \beta)^2$$

Then we find the value of $\hat{\beta}$ that minimizes Z:

$$\begin{aligned} \frac{dZ}{d\beta} &= 0 \\ -2\sum_{i=1}^{n}(X_i - \hat{\beta}) &= 0 \\ \hat{\beta} &= \frac{\sum_{i=1}^{n} X_i}{n} = \bar{X} \end{aligned}$$

□

3.3 Maximum likelihood method

Let $L(X)$ be the likelihood function for a sample size $n = n_1 + n_2$ of the proposed model. Then,

$$\begin{aligned} L &= \prod_{i=1}^{n}[(1-\beta+X_i)I_{[\beta-1;\beta]}(X_i) + (1+\beta-X_i)I_{(\beta;\beta+1]}(X_i)] \\ &= \prod_{i=1}^{n_1}(1-\beta+X_i) \times \prod_{i=1}^{n_2}(1+\beta-X_i) \end{aligned}$$

Thus, the support function $S(X)$ is given by

$$\begin{aligned} S(X) &= \ln L(X) \\ &= \sum_{i=1}^{n_1}\ln(1-\beta+X_i) + \sum_{i=1}^{n_2}\ln(1+\beta-X_i) \end{aligned}$$

Then, the score function becomes

$$\begin{aligned} \frac{d}{d\beta}S(X) &= 0 \\ \sum_{i=1}^{n_1}\frac{1}{1-\hat{\beta}+X_i} &= \sum_{i=1}^{n_2}\frac{1}{1+\hat{\beta}-X_i}, \end{aligned}$$

what is not straightforward solving.

4 Rnyi entropy

Theorem 12 *The Rnyi entropy of order α of (1) is given by*

$$H_\alpha(X) = \frac{1}{1-\alpha}\ln\left(\frac{[1-\beta+X]^{\alpha+1}}{\alpha+1}I_{[\beta-1;\beta]}(X) - \frac{[1+\beta-X]^{\alpha+1}}{\alpha+1}I_{(\beta;\beta+1]}(X)\right)$$

where $\alpha \geq 0$ and $\alpha \neq 1$.

Proof.

$$\begin{aligned} H_\alpha(x) &= \frac{1}{1-\alpha}\ln\left(\int f(x)^\alpha dx\right) \\ &= \frac{1}{1-\alpha}\ln\left(\int \left[(1-\beta+X)I_{[\beta-1;\beta]}(X) + (1+\beta-X)I_{(\beta;\beta+1]}(X)\right]^\alpha dx\right) \\ &= \frac{1}{1-\alpha}\ln\left(\int_{x\in[\beta-1;\beta)}[1-\beta+X]^\alpha dx + \int_{x\in(\beta;\beta-1]}[1+\beta-X]^\alpha dx\right) \\ &= \frac{1}{1-\alpha}\ln\left(\frac{[1-\beta+X]^{\alpha+1}}{\alpha+1}I_{[\beta-1;\beta]}(X) - \frac{[1+\beta-X]^{\alpha+1}}{\alpha+1}I_{(\beta;\beta+1]}(X)\right) \end{aligned}$$

□

5 Application

An application example is presented in this section. Real Cash flow data, adapted from [5], were used to illustrate the model adjustment. Fifteen observations of cash flow were normalized ($\beta = 0$) and the proposed model adjusted (Figure 2).

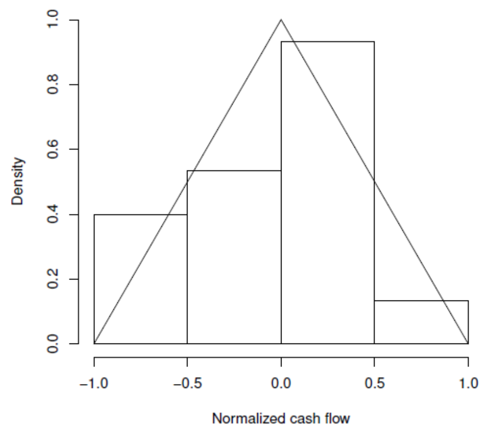

Figure 3: Adjustment of the proposed triangular model for normalized cash flow data.

The model revealed to fit the data suitably since chi-squared test did not indicate to reject the null hypothesis of adherence (p-value=0.2615).

Such small sample situation is an example of interesting case for using the proposed parsimonious model.

6 Conclusions

The parsimonious triangular distribution proposed in this paper is a considerable alternative for the case of small samples (or large samples in situations of symmetric populations).

7 Acknowledgments

Warm acknowledgments to FAPEMIG, CNPq and CAPES for the support.

References

[1] K. J. Cohen, S. F. Maier, R. A. Schwartz, and D. K. Whitcomb. Transaction costs, order placement strategy, and existence of the bid-ask spread. *The Journal of Political Economy*, 89:287–305, 1981.

[2] J. L. M. de Souza, J. A. Frizzone, and E. S. Gervásio. Simulação do balanço hídrico climatológico decendial para a cultura do cafeeiro, na região de lavras, minas gerais. In *V Simpósio Brasileiro de Pesquisa em Cafeicultura Irrigada*, 2002.

[3] R. Van Dorp and S. Kotz. A novel extension of the triangular distribution and its parameter estimation. *Journal of the Royal Statistical Society D*, 51:63–79, 2002.

[4] Christophe Dutang. A guide on probability distributions, 2013. Available at: http://dutangc.free.fr/pub/prob/probdistr-main.pdf.

[5] R. J. Ferreira. Demonstração dos fluxos de caixa., Acess on: 15 Sep 2010. Available at: www.vemconcursos.com.

[6] A. M. Figueiredo, P. A. dos Santos, R. Santolin, and B. S. Reis. Integração na criação de frangos de corte na microrregião de viçosa-mg: viabilidade econômica e análise de risco. *Revista de Economia e Sociologia Rural*, 44:713–730, 2006.

[7] R. Hesse. Triangle distribution: Mathematica link for excel. *Decision Line*, May:12–14, 2000.

[8] D. Johnson. The triangular distribution as a proxy for the beta distribution in risk analysis. *Journal of the Royal Statistical Society D*, 46:387–398, 1997.

[9] N. L. Johnson and S. Kotz. Non-smooth sailing or triangular distributions revisited after some 50 years. *Journal of the Royal Statistical Society D*, 48:179–187, 1999.

[10] R. A. Menescal and V. P. P. B. Vieira. Manutenção de sangradouro de açude e risco de ineficiência hidráulica. In *XXIII Seminário Nacional de Grandes Barragens*, 1999.

[11] S. M. Ross. *A first course in probability. 8th ed.* New Jersey: Pearson Prentice Hall. 530 p., 2010.