

# A Skewed Model Combining Triangular and Exponential Features: The Two-faced Distribution and its Statistical Properties

Maurizio Brizzi  
University of Bologna, Italy

**Abstract:** A new continuous distribution model is introduced, joining triangular and exponential features, respectively on the left and right side of a hinge point. The cumulative distribution function is derived, as well as the first three moments. Expected values and the Pearson index of skewness are tabulated. A possible step-by-step approach to parameter estimation is outlined. An application to Italian geographical data is given, referring to a set of municipalities classified by population, showing a very satisfactory goodness of fit.

**Zusammenfassung:** Ein neues stetiges Verteilungsmodell wird vorgestellt, das links eines Trennpunktes Eigentümlichkeiten einer Dreiecks- und rechts davon einer Exponentialverteilung vereint. Die Verteilungsfunktion wie auch die ersten drei Momente werden hergeleitet. Erwartungswerte und der Pearson Schiefe-Index sind tabelliert. Die mögliche schrittweise Annäherung an die Parameterschätzer wird skizziert. Angewandt wird das Modell auf italienische geographische Daten, die sich auf Stadtverwaltungen beziehen, welche durch ihre Bevölkerungszahlen klassifiziert sind. Dabei zeigt sich eine sehr zufrieden stellende Anpassungsgüte.

**Keywords:** Triangular Distribution, Exponential Distribution, Skewness, Maximum Likelihood Estimation.

## 1 Introduction

The research regarding probability density models is always in progress, since many researchers are interested in finding simple (or quite simple) functions, which are able to fit a set of experimental data. Nowadays, we have a lot of quantitative information from several fields, from biology to sociology, from chemistry to economy, from sports to music, and the list would be much longer. Therefore, we often have to deal with “new” sets of data, and it is always more important to have a large choice of probability models, discrete or continuous, in order to find the best fitting one.

Recently, the triangular distribution, a very simple and well-known one, has been reconsidered and developed by Johnson (1997), Johnson and Kotz (1999), Van Dorp and Kotz (2002b). One of the most recently developed family of densities, the STSP family proposed by Van Dorp and Kotz (2002a) includes the triangular distribution as a particular case. Here we propose a mixed distribution, having some triangular and exponential features at the same time. The new distribution is then fitted to a set of Italian population data.

## 2 Two-faced Distribution: Characterizing Functions and Moments

Suppose that a random variable  $Y$  has a probability density function (pdf) which increases linearly until a modal point  $\theta$ , and then decreases exponentially. We call such a variable, having a triangular left side and an exponential right side, a two-faced (TF) variable. The original parameters are  $\theta$  (hinge point) and  $\lambda$  (exponential coefficient).

The pdf of such a TF distribution follows the following pattern

$$f_Y(y) = \begin{cases} 0, & y < 0, \\ \alpha y, & 0 \leq y < \theta, \\ \beta \exp(-\lambda y), & \theta \leq y, \end{cases} \quad (1)$$

where

$$\alpha = \frac{2\lambda}{\theta(\lambda\theta + 2)} \quad \text{and} \quad \beta = \frac{2\lambda \exp(\lambda\theta)}{\lambda\theta + 2}.$$

The value of  $\alpha$  and  $\beta$  has been determined by the two necessary conditions:

1. the pdf integral has to be equal to 1,
2. the value at the hinge point has to be unique, i.e., the pdf has to be continuous at  $\theta$ .

For better simplicity, it seems to be useful to change the second parameter. Let  $\kappa = \lambda\theta$  denote a shape parameter. With this notation we rewrite (1) as

$$f_Y(y) = \begin{cases} 0, & y < 0, \\ \frac{2\kappa}{\theta^2(\kappa + 2)}y, & 0 \leq y < \theta, \\ \frac{2\kappa}{\theta(\kappa + 2)} \exp\left(-\frac{\kappa}{\theta}(y - \theta)\right), & \theta \leq y. \end{cases} \quad (2)$$

In Figure 1 we represent the pdf (2) for various values of  $\kappa$  and in Figure 2 for different values of  $\theta$ .

The cumulative distribution function (cdf) is directly derived from (2) as

$$F_Y(y) = \begin{cases} 0, & y < 0 \\ \frac{\kappa}{\theta^2(\kappa + 2)}y^2, & 0 \leq y < \theta \\ 1 - \frac{2}{\kappa + 2} \exp\left(-\frac{\kappa}{\theta}(y - \theta)\right), & \theta \leq y. \end{cases} \quad (3)$$

The pdf and the cdf at the hinge point  $\theta$  are

$$f_Y(\theta) = \frac{2\kappa}{\theta(\kappa + 2)}, \quad F_Y(\theta) = \frac{\kappa}{\kappa + 2}. \quad (4)$$

In particular:

if  $\kappa = 2$ , then  $F_Y(\theta) = 1/2$ , so  $\theta$  is the median,

if  $\kappa < 2$ , then  $F_Y(\theta) < 1/2$ , so  $\theta$  is less than the median,

if  $\kappa > 2$ , then  $F_Y(\theta) > 1/2$ , so  $\theta$  is greater than the median.

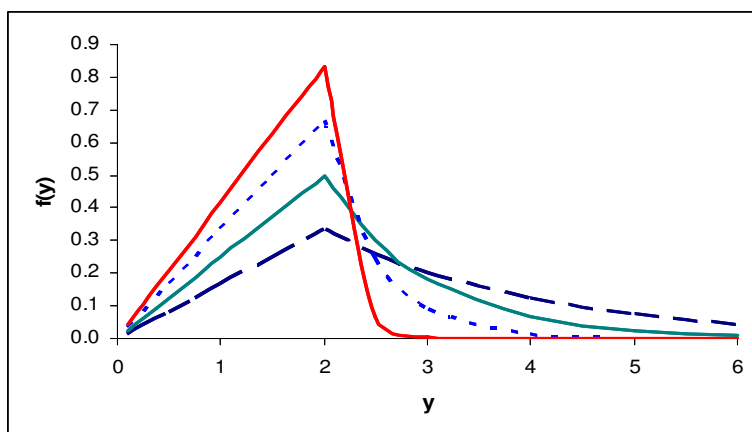


Figure 1: Two-faced probability density function with  $\theta = 2$  and  $\kappa = 1, 2, 4, 10$ .

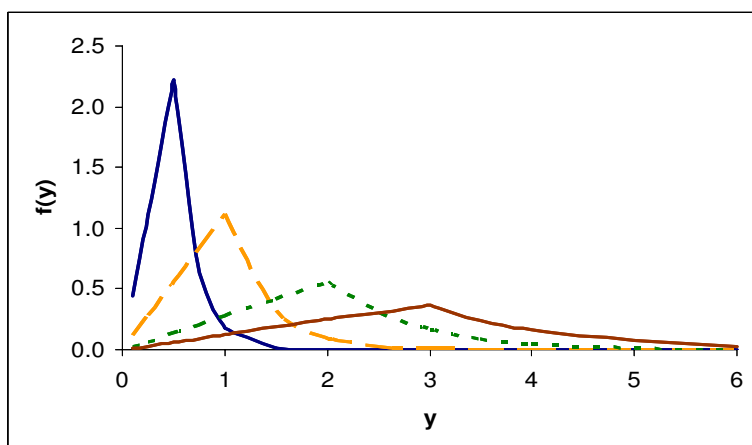


Figure 2: Two-faced probability density function with  $\kappa = 2.5$  and  $\theta = 0.5, 1, 2, 3$

The expected value  $E(Y)$  is to be calculated by considering separately (and then adding) the two parts (two “faces”) of the TF distribution:

$$\begin{aligned}
 E(Y) &= \int_0^\theta \alpha y^2 dy + \int_\theta^{+\infty} \beta y \exp(-\lambda y) dy \\
 &= \frac{2\kappa\theta}{3(\kappa+2)} + \frac{2\theta(\kappa+1)}{\kappa(\kappa+2)} \\
 &= \frac{2(\kappa^2 + 3\kappa + 3)}{3\kappa(\kappa+2)} \theta.
 \end{aligned} \tag{5}$$

Therefore, moving the hinge point to the right (i.e., increasing the value of  $\theta$ ) the expected value  $E(Y)$  becomes larger, while the parameter  $\kappa$  is inversely linked to the expected value. In particular we notice that  $E(Y) = \theta$  as  $\kappa = \sqrt{6}$ , since

$$\frac{2\theta(6 + 3\sqrt{6} + 3)}{3\sqrt{6}(\sqrt{6} + 2)} = \theta.$$

Table 1: Expected value  $E(Y)$  for some parameter values

$\theta$	0.5	1	2	$\frac{\kappa}{\sqrt{6}}$	3	5	10
0.5	1.267	0.778	0.542	0.500	0.467	0.410	0.369
1	2.533	1.556	1.083	1.000	0.933	0.819	0.739
2	5.067	3.111	2.167	2.000	1.867	1.638	1.478
3	7.600	4.667	3.250	3.000	2.800	2.457	2.217
4	10.133	6.222	4.333	4.000	3.733	3.276	2.956
6	15.200	9.333	6.500	6.000	5.600	4.914	4.433
10	25.333	15.556	10.833	10.000	9.333	8.190	7.389

In the same way we derive the second moment

$$E(Y^2) = \frac{\kappa^3 + 4\kappa^2 + 8\kappa + 8}{2\kappa^2(\kappa + 2)}\theta^2. \quad (6)$$

and determine the variance as

$$\text{var}(Y) = \frac{\kappa^4 + 6\kappa^3 + 24\kappa^2 + 72\kappa + 72}{18\kappa^2(\kappa + 2)^2}\theta^2. \quad (7)$$

The variance, as also the mean, increases directly with  $\theta$  and decreases as  $\kappa$  increases.

We now compute the third moment, useful for evaluating the skewness, and get

$$E(Y^3) = \frac{2(2\kappa^4 + 5\kappa^3 + 15\kappa^2 + 30\kappa + 30)}{5\kappa^3(\kappa + 2)}\theta^3. \quad (8)$$

The well-known Pearson index of skewness is defined as  $\gamma = E[Y - E(Y)]^3 \text{var}(Y)^{-3/2}$ . The resulting formula for  $\gamma$  is quite complicated and not comfortable to use. Thus, we report in Table 2 the first three relative moments as functions of  $\kappa$ . Looking at Table 2, we realize that the moments directly depend on the parameter  $\theta$  and inversely depend on  $\kappa$ . The skewness index depends only on  $\kappa$  and it is positive for small values of  $\kappa$  and becomes negative for larger values. The value of  $\kappa$  corresponding to  $\gamma = 0$  is, approximately,  $\kappa = 6.15$ . Obviously, even in this context, the pdf will not be perfectly symmetric, having a limited left tail and an unlimited right one.

### 3 Parameter Estimation

Let  $y_1, \dots, y_n$  be an iid sample from a two-faced distribution. There are two parameters to be estimated. If  $\theta$  is known, the maximum likelihood (ML) estimator of  $\kappa$  is

$$\hat{\kappa} = \sqrt{1 + 2n\theta / \sum_{\theta \leq y_j} (y_j - \theta)} - 1 \quad (9)$$

Table 2: First three relative moments and some related indices, where  $\bar{\mu}_3 = E[Y - E(Y)]^3$

$\kappa$	$E(Y)/\theta$	$E(Y^2)/\theta^2$	$E(Y^3)/\theta^3$	$\sqrt{\text{var}}(Y)/\theta$	$\bar{\mu}_3/\theta^3$	$\gamma$
0.5	2.533	10.500	63.280	2.020	15.997	1.939
1	1.556	3.500	10.800	1.039	1.995	1.777
2	1.083	1.500	2.575	0.571	0.243	1.302
$\sqrt{6}$	1.000	1.242	1.853	0.492	0.127	1.070
3	0.933	1.056	1.396	0.429	0.066	0.834
4	0.861	0.875	1.006	0.365	0.023	0.469
5	0.819	0.780	0.825	0.330	0.007	0.207
6	0.792	0.722	0.724	0.309	0.001	0.023
8	0.758	0.656	0.616	0.285	-0.005	-0.203
10	0.739	0.620	0.561	0.272	-0.007	-0.324

and it depends, essentially, on the right tail values of the sample.

The more complicated problem is to estimate  $\theta$ , since the pdf depends directly on its value. Since  $\theta$  denotes the modal point, a possible estimating procedure may be like:

1. Find the part of the sample, where the points are more dense, thus identifying a subset of say  $n^*$  elements with values not too far from each other.
2. Let  $\hat{\theta}$  be equal to each of the  $n^*$  points in the subset and calculate the corresponding value of  $\hat{\kappa}$  by applying (9). We find  $n^*$  parameter estimates  $(\hat{\theta}_j, \hat{\kappa}_j), j = 1, \dots, n^*$ .
3. Calculate the log-likelihood for each  $(\hat{\theta}_j, \hat{\kappa}_j)$ , and consider  $(\hat{\theta}^*, \hat{\kappa}^*)$  with the largest log-likelihood.
4. Change the value of  $\hat{\theta}^*$  slightly in both directions. If the resulting likelihood does not decrease, we keep  $(\hat{\theta}^*, \hat{\kappa}^*)$  as the ML estimate.

*Example:* Consider the sample (2.5, 3.4, 4.1, 4.8, 5.2, 5.3, 5.9, 6.8, 7.2, 8.9, 10.5, 13.4) with  $n = 12$  values. The central part of it (five sample points, values ranging from 4.1 to 5.9) is the “dense kernel”, and we consider for each kernel point the corresponding estimate  $\hat{\kappa}$  and calculate the log-likelihood. Results are shown in Table 3. The log-likelihood increases until reaching its maximum for  $\hat{\theta} = 5.3$ , then decreases again. It is easy to verify that neighboring values all show log-likelihood values less than  $-29.299$ . Therefore, we can consider  $\hat{\theta} = 5.3$  and  $\hat{\kappa} = 1.662$  as the ML estimates of the parameters  $\theta$  and  $\kappa$ .

Table 3: ML Estimates of both TF parameters

$\hat{\theta}$	$\hat{\kappa}$	-Log-likelihood
4.1	1.041	30.056
4.8	1.376	29.540
5.2	1.603	29.318
5.3	1.662	29.299
5.9	2.031	29.376

## 4 Application to the Distribution of Municipal Population of an Italian Region

We found a nice application of the two-faced model to the distribution of 328 (out of 341) municipalities of the Emilia Romagna region. We excluded 13 cities with more than 50000 resident people, since the model does not seem to fit well with big cities. The region is located between the Po river and the Apennine ridge. The biggest city of the region is Bologna with about 375000 inhabitants, and there are many middle-sized towns. The total population is around 4 million people, corresponding to 7% of the Italians. The working variable is  $Y$  which denotes the population at Census 2001 (in thousands of inhabitants). We tried to find the hinge point by analyzing the modal range of the population. Evidently, the most dense interval corresponds to small municipalities between  $Y = 1$  and  $Y = 3$ . We divided this subset of small municipalities in classes of width 0.2 (200 people), and identified two main classes: “2.0 to 2.2” with 14 units and “2.2 to 2.4” with 12 units. These classes are sensibly denser than the neighboring classes. We then chose the intermediate value as hinge point, i.e.,  $\hat{\theta}^* = 2.2$ . Applying (9) we derive the corresponding ML estimate  $\hat{\kappa} = 0.384$ . The estimated model (see formula 2) is

$$\hat{f}_Y(y) = \begin{cases} 0.0665y, & 0 < y < 2.2, \\ 0.1464 \exp(-0.349(y - 2.2)), & 2.2 \leq y. \end{cases}$$

The resulting cdf and the empirical cdf are compared in Figure 3. Considering the several factors that may affect the distribution of the population, the similarity between the two curves is quite impressive. The maximum absolute deviation between the model and the data, i.e., the Kolmogorov distance, is 0.0394.

In order to give another visual representation of the good fit of the model, a quantile-quantile plot is shown in Figure 4, putting in abscissa the observed centiles and in ordinate the corresponding centiles of the fitted two-faced model. A perfectly fitting model would give a straight line. Looking at the plot, the correspondence is evident, and some slight deviation from linearity may be found only in the right tail (upper centiles). This gives a further evidence of a satisfactory goodness-of-fit of the proposed model.

## 5 Concluding Remarks

The two-faced distribution may be suitable for fitting some natural or social phenomena in which there is a “hinge point” which is the limit point between “small” and “large” values: the density reaches its maximum at this point, and then decreases exponentially, with a very long tail. The application to Emilia Romagna municipal population data marks a hinge point of 2.2, having some observations which are 15 to 20 times larger than the modal value. It could be interesting to fit the same distribution to other geographical data, in which there may be a great majority of small and middle-sized observations and a little minority of very large ones. Probably, a similar uneven distribution may be found when analyzing some economic and social variables, and it would be interesting to check the adequacy of the present model.

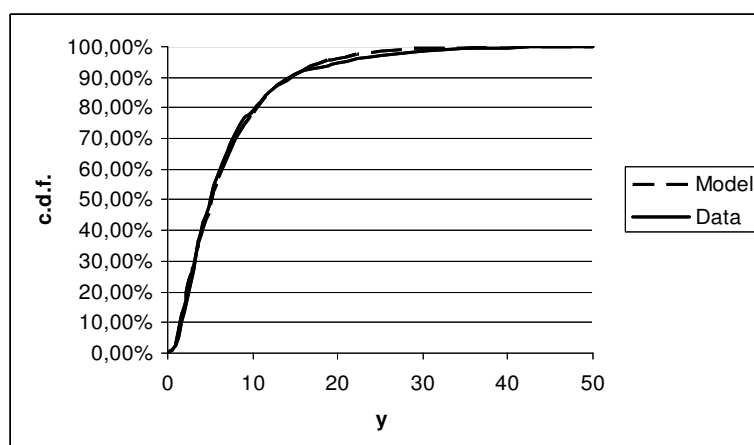


Figure 3: Population of municipalities in Emilia Romagna: comparison of the theoretical and the empirical cumulative distribution functions

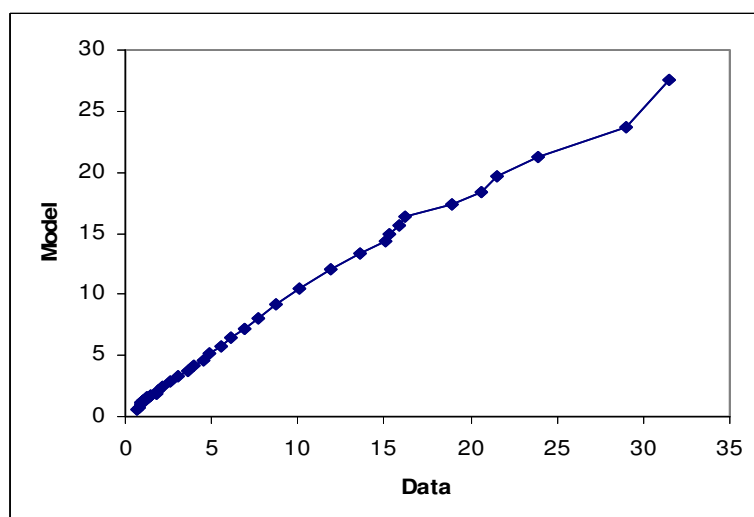


Figure 4: Population of municipalities in Emilia Romagna and two-faced theoretical model: comparison by quantile-quantile plot

Probably, the TF distribution pattern may be proposed to other “classical” models, in which the left and right tails have a different shape. The triangular distribution may be “faced” with other models, like Gaussian or Gamma (a triangular-Gamma model would include the TF distribution). Another challenging problem would be to find real phenomena behaving like that. The scientific research, as always, has a primary role, either in developing (and improving) techniques and models or in finding connections with real life. Both aspects have a great importance and are strictly connected to each other.

### Acknowledgements

The Author is grateful to Prof. Samuel Kotz (George Washington University) for a lot of useful hints and suggestions about this paper.

## References

- Johnson, N. L. (1997). The triangular distribution as a proxy for the beta distribution in risk analysis. *The Statistician*, 46, 387-398.
- Johnson, N. L., and Kotz, S. (1999). Non-smooth sailing or triangular distributions revisited after some 50 years. *The Statistician*, 48, 179-187.
- Van Dorp, J. R., and Kotz, S. (2002a). The standard two-sided power distribution and its properties; with applications in financial engineering. *The American Statistician*, 56, 90-99.
- Van Dorp, J. R., and Kotz, S. (2002b). A novel extension of triangular distribution and its parameter estimation. *The Statistician*, 51, 63-79.

Author's address:

Maurizio Brizzi

Department of Statistical Sciences "Paolo Fortunati"

University of Bologna

Via delle Belle Arti 41

40126 Bologna, Italy

E-mail: [brizzi@stat.unibo.it](mailto:brizzi@stat.unibo.it)