SOME CONSIDERATIONS AND DEVELOPMENTS ABOUT RECENTLY PROPOSED FAMILIES OF DISTRIBUTIONS WITH HINGE POINTS

Maurizio Brizzi

1. INTRODUCTION

The research work regarding probability density models is, necessarily, always in progress, since many researchers are interested in finding reasonably simple functions, which are able to fit a set of experimental data. Nowadays, we receive a lot of quantitative information concerning several fields, from biology to sociology, from chemistry to economy, from sport to music, and the list would be much longer; therefore, we often have to deal with "new" sets of data, and is always more important to have a large choice of probability models, discrete or continuous, in order to find the best fitting one.

Sometimes, when representing the distribution of some kind of natural or social phenomena, we realize that smaller values have a different behaviour than larger ones, so that we cannot use an unique type of function, but rather two (or more) different functions, separated by hinge points.

In this paper, two families of p.d.f.'s with hinge points are discussed and developed. The first one is the MTSP (Modified Two-Sided Power), derived from the STSP (Standard Two-Sided Power), defined and studied by Van Dorp and Kotz (2002a); the second one is the two-faced distribution (TF), recently proposed in Brizzi (2006), defined by joining a triangular distribution and an exponential one.

2. FROM STSP TO MTSP

The Standard Two-Sided Power distribution is a limited domain model, depending on two parameters: the "hinge point" θ, in which the p.d.f. changes and becomes to decrease, and the shape parameter κ. In its original proposal by Van Dorp and Kotz (2002), the domain is the interval *[0,1]*, and the p.d.f. is the following:

$$f_X(x) = \begin{cases} \kappa \left(\dfrac{x}{\theta} \right)^{\kappa-1} & 0 < x \le \theta \\[4mm] \kappa \left(\dfrac{1-x}{1-\theta} \right)^{\kappa-1} & \theta < x \le 1 \end{cases} \tag{1}$$

The effect of varying the parameters is represented in Figure 1 (where θ is fixed and κ is variable) and Figure 2 (κ fixed and θ variable).
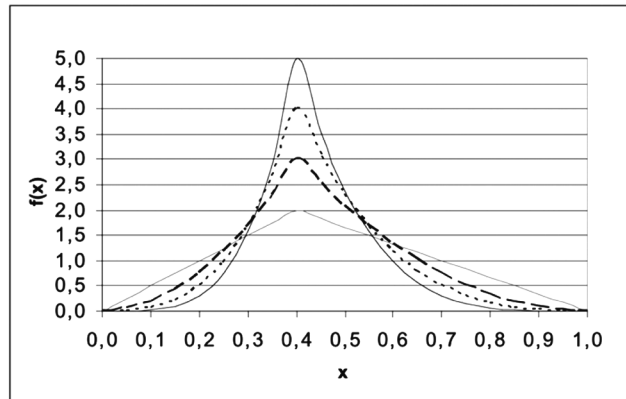

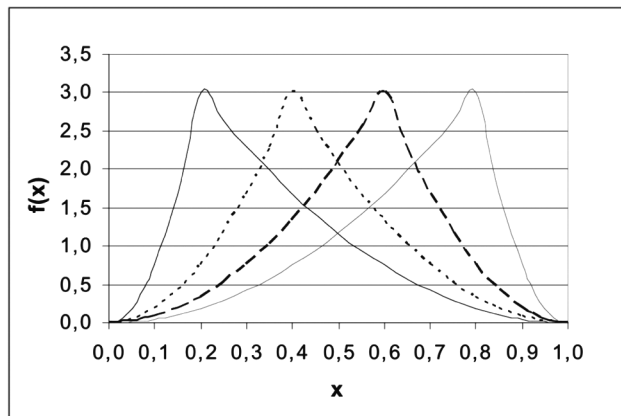*Figure 1* – Probability density function of some STSP distributions ($\theta = 0.4$, κ variable).

Figure 2 – Probability density function of some STSP distributions ($\kappa = 3$, θ variable).

The MTSP distribution family is generated by modifying the central part of an STSP distribution. The basic idea is to "strengthen" the neighbourhood of the hinge point θ by adding an uniform component.

$$f_Y(y) = (1-\alpha)f_Y(y) + \alpha f_U(y) \tag{2}$$

where X is a STSP (κ, θ) and U is an Uniform $(\theta-\beta/2, \theta+\beta/2)$. The domain of Y is the interval *[0,1]*, the same of STSP distribution.

The p.d.f. assumes, within its domain, four different expressions:

$$f_Y(y) = \begin{cases} (1-\alpha)\kappa\left(\dfrac{y}{\theta}\right)^{\kappa-1}, & 0 < y \leq \theta - \beta/2 \\[2ex] (1-\alpha)\kappa\left(\dfrac{y}{\theta}\right)^{\kappa-1} + \dfrac{\alpha}{\beta}, & \theta - \beta/2 < y \leq \theta \\[2ex] (1-\alpha)\kappa\left(\dfrac{1-y}{1-\theta}\right)^{\kappa-1} + \dfrac{\alpha}{\beta}, & \theta < y \leq \theta + \beta/2 \\[2ex] (1-\alpha)\kappa\left(\dfrac{1-y}{1-\theta}\right)^{\kappa-1}, & \theta + \beta/2 < y \leq 1 \end{cases} \tag{3}$$

Parameters and corresponding parametric spaces are listed below:

$\alpha =$ Weight of the uniform component   *0 < α < 1*
$\beta =$ Width of the uniform component   *0 < β < 2 min [θ, (1−θ)]*
$\kappa =$ Shape parameter   *κ > 0*
$\theta =$ Mode and hinge point   *0 < θ < 1*

In Figures 3 and 4 we have represented the p.d.f. of a MTSP distribution, as a function of the parameters α, β.
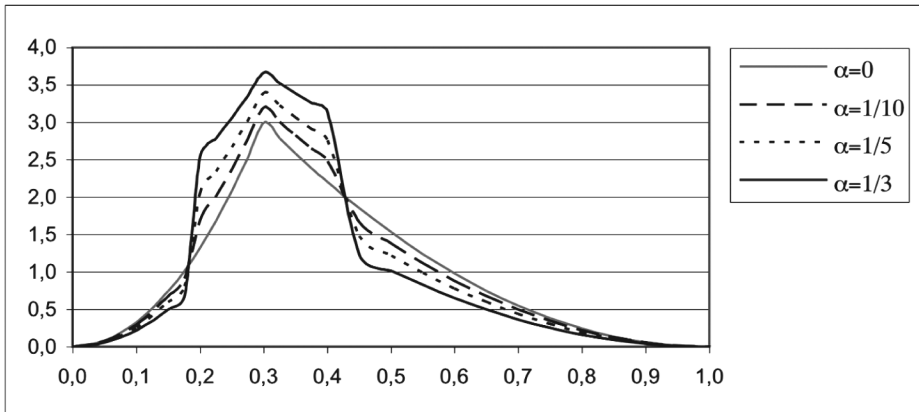

Figure 3 – Probability density function of some MTSP distributions (κ=3, θ=0.3, α variable, β=0.2).
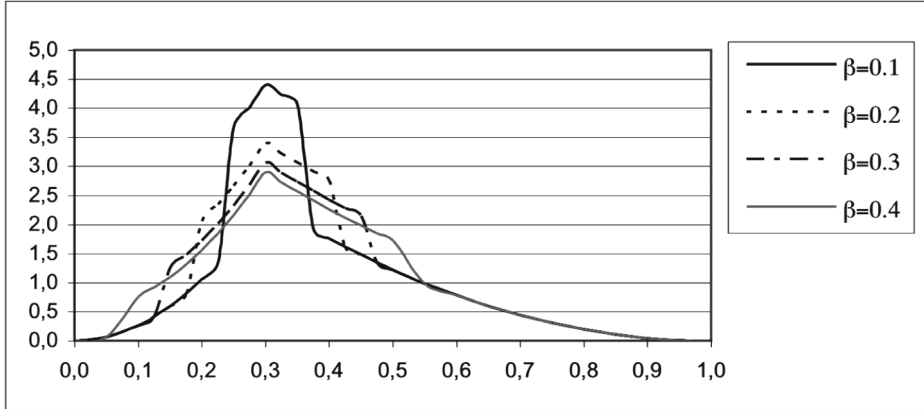
Figure 4 – Probability density function of some MTSP distributions ($\kappa=0.3,\ \theta=0.3,\ \alpha=0.2,\ \beta$ variable).

The cumulative distribution function $F_Y(y)$ assumes, within its domain $[0, 1]$, four different expressions (outside the domain, the c.d.f. is obviously constant, 0 on the left and 1 on the right):

$$F_Y(y) = \begin{cases} (1-\alpha)\,\theta\left(\dfrac{y}{\theta}\right)^{\kappa},\ \ 0 < y \le \theta-\beta/2 \\[2mm] (1-\alpha)\,\theta\left(\dfrac{y}{\theta}\right)^{\kappa}+\alpha\,\dfrac{y-(\theta-\beta/2)}{\beta},\ \ \theta-\beta/2 < y \le \theta \\[2mm] (1-\alpha)\left[1-(1-\theta)\left(\dfrac{1-y}{1-\theta}\right)^{\kappa}\right]+\alpha\,\dfrac{y-(\theta-\beta/2)}{\beta},\ \ \theta < y \le \theta+\beta/2 \\[2mm] (1-\alpha)\left[1-(1-\theta)\left(\dfrac{1-y}{1-\theta}\right)^{\kappa}\right],\ \ \theta+\beta/2 < y \le 1 \end{cases} \qquad (4)$$

In particular, we have, at point θ:

$$F_Y(\theta) \;=\; (1-\alpha)\,F_X(\theta)+\alpha\,F_U(\theta) = (1-\alpha)\theta + \alpha/2 = \theta + \alpha(1/2-\theta) \qquad (5)$$

In Figure 5 and Figure 6 the c.d.f. of MTSP distribution is represented, as a function of the parameters α and β, respectively.
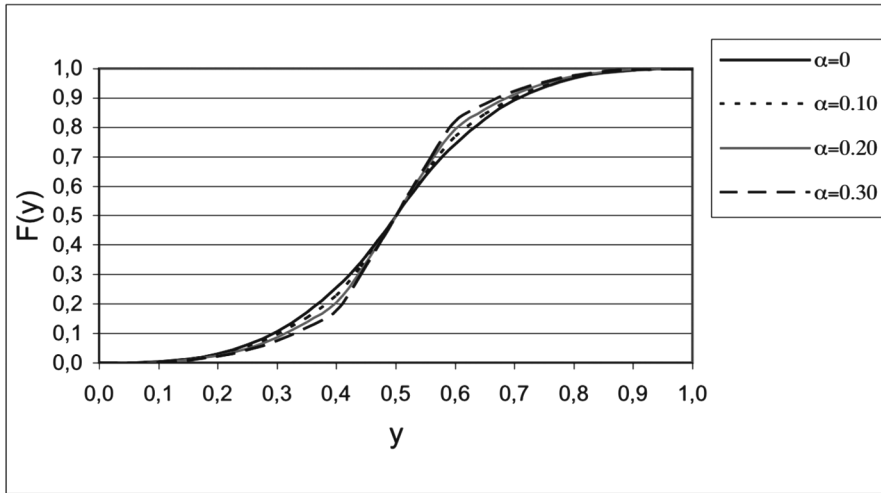
Figure 5 – Cumulative distribution function of MTSP (α variable, β=*0.20*).
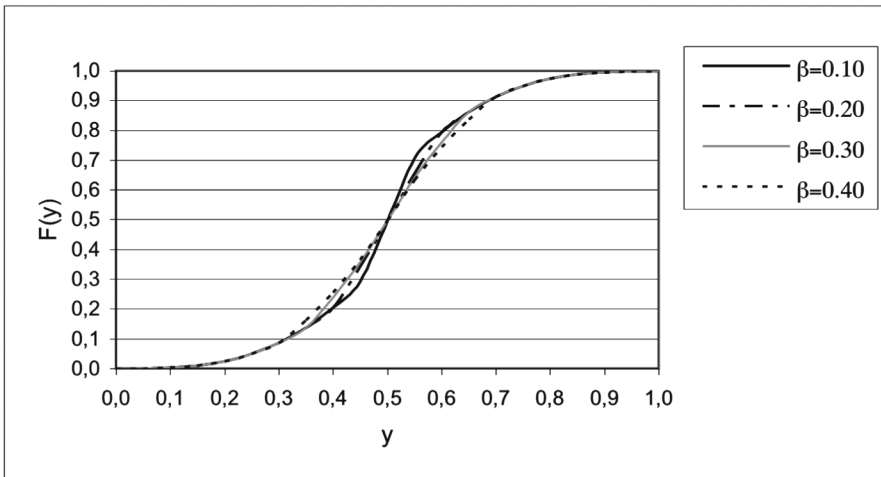
Figure 6 – MTSP cumulative distribution function (α=*0.20*, β variable).

The expected value of a MTSP distribution is:

$$\mathrm{E}(Y) = (1-\alpha)\,\mathrm{E}(X) + \alpha\,\mathrm{E}(U) = (1-\alpha)\frac{(\kappa-1)\theta+1}{\kappa+1} + \alpha\theta = \frac{(2\alpha+\kappa-1)\theta+1-\alpha}{\kappa+1} \quad (6)$$

Therefore, $E(Y)$ does not depend on β. When α increases, $E(Y)$ moves linearly towards θ. For example, fixing some values of θ and κ, we may derive, by changing the value of α, the expected values represented in Table 1:

TABLE 1

Expected value of MTSP as a function of α

α	κ=3, θ=1/4	κ=4, θ=1/4	κ=5, θ=1/4	κ=4, θ=1/3	κ=4, θ=1/2	κ=4, θ=2/3
0.05	0.3688	0.3450	0.3292	0.3967	0.5000	0.6033
0.10	0.3625	0.3400	0.3250	0.3933	0.5000	0.6067
0.20	0.3500	0.3300	0.3167	0.3867	0.5000	0.6133
0,30	0.3375	0.3200	0.3083	0.3800	0.5000	0.6200
0,40	0.3250	0.3100	0.3000	0.3733	0.5000	0.6267
0,50	0.3125	0.3000	0.2917	0.3667	0.5000	0.6333

The effect of increasing α is to move the expected value towards the value of θ, which is the central point of the uniform window; when $\theta = 1/2$, there is no effect, since we have already $E(X) = \theta$. The effect of increasing α is to move the expected value towards the value of θ, which is the central point of the uniform window; when $\theta = 1/2$, there is no effect, since we have already $E(X) = \theta$.

The second moment of *Y*, i.e. the expected value of Y^2 , is:

$$
E(Y^2) = (1-\alpha)\ E(X^2) + \alpha\ E(U^2) =
$$
$$
= (1-\alpha)\left[\frac{\kappa\theta^3}{(\kappa+2)} + (1-\theta) - 2\frac{\kappa}{\kappa+1}(1-\theta)^2 + \frac{\kappa}{\kappa+2}(1-\theta)^3\right] + \alpha\left(\theta^2 + \frac{\beta^2}{12}\right)\ (7)
$$

Consequently, the variance $V(Y)=E(Y^2)-[E(Y)]^2$ assumes a troubling complicated general expression. Nevertheless, we can study the effect of varying α and β by fixing the values of the STSP parameters, κ and θ. We have reported, in Table 2, the variance corresponding to some values of α and β for $\kappa = 4$ and $\theta = 1/3$.

TABLE 2

Variance of MTSP for κ = 4 , θ = 1/3

	β =	0.1	0.2	0.3	0.4	0.5
α =						
0.1		0.0165	0.0167	0.0172	0.0177	0.0185
0.2		0.0151	0.0156	0.0164	0.0176	0.0191
0.3		0.0136	0.0144	0.0156	0.0174	0.0196
0.4		0.0121	0.0131	0.0147	0.0171	0.0201
0.5		0.0104	0.0117	0.0138	0.0167	0.0204

The effect of varying β becomes stronger as well as α increases; on the other side, the effect of varying α depends on the difference between the variance of STSP component and uniform component: when $\beta = 0.43$ the two components have approximately the same variance, and the value of α has virtually no effect on it.

Analysing the third moment of MTSP distribution, we can check its skewness. The third moment of MTSP is equal to:

$$E(Y^3) = (1-\alpha)E(X^3) + \alpha E(U^3) =$$

$$= (1-\alpha)\left[\frac{\kappa\theta^4}{(\kappa+3)} + (1-\theta) - 3\frac{\kappa}{\kappa+1}(1-\theta)^2 + 3\frac{\kappa}{\kappa+2}(1-\theta)^3 - \frac{\kappa}{\kappa+3}(1-\theta)^4\right] + \alpha\theta\left(\theta^2 + \frac{\beta^2}{4}\right) \quad (8)$$

The third central moment $\mu_3(Y)$ is determined by a polynomial function of simple moments of order *1, 2, 3*:

$$\mu_3(Y) = E[Y - E(Y)]^3 = E(Y^3) - 3E(Y^2)E(Y) + 2\ E(Y)^3 \tag{9}$$

We can then directly calculate the Pearson index of skewness: $\gamma_1(Y) = \mu_3(Y)\ /\ \sigma^3(Y)$, whose final expression is rather cumbersome.

Likewise, we can evaluate the kurtosis of MTSP distribution by analysing the fourth moment. As done before, we can calculate the fourth moment by linear combination:

$$E(Y^4) = (1-\alpha)\ E(X^4) + \alpha\ E(U^4) \tag{10}$$

where:

$$E(X^4) = \left[\frac{\kappa\,\theta^5}{(\kappa+4)} + (1-\theta) - 4\frac{\kappa}{\kappa+1}(1-\theta)^2 + 6\frac{\kappa}{\kappa+2}(1-\theta)^3 - 4\frac{\kappa}{\kappa+3}(1-\theta)^4 + \frac{\kappa}{\kappa+4}(1-\theta)^5\right]$$

$$E(U^4) = \left(\theta^4 + \frac{\theta^2\beta^2}{2} + \frac{\beta^4}{80}\right)$$

The fourth central moment may be calculated as a polynomial function of simple moments:

$$\mu_4(Y) = E[Y - E(Y)]^4 = E(Y^4) - 4E(Y^3)E(Y) + 6\ E(Y^2)[E(Y)]^2 - 3[E(Y)]^4 \tag{11}$$

From (11) we can derive the Fisher index of kurtosis: $\gamma_2(Y) = \mu_4(Y)\ /\ \sigma^4\ (Y)$. We prefer not to write here the final expression of: $\gamma_2(Y)$, since it is results quite long and complicated; nevertheless, it is not so difficult to calculate such index, once fixed the values of κ and θ.

Example. Consider once more the parameter values represented in Table 2 *($\kappa=4$, $\theta = 1/3$)*. The first four moments have been calculated and reported below:

$$E(Y) = \frac{2}{5} - \frac{1}{15}\alpha$$

$$E(Y^2) = \frac{8}{45} + \left(\frac{\beta^2}{12} - \frac{1}{15} \right) \alpha$$

$$E(Y^3) = \frac{82}{945} + \left(\frac{\beta^2}{12} - \frac{47}{945} \right) \alpha$$

$$E(Y^4) = \frac{29}{630} - \frac{191}{5670} \alpha + \frac{1}{18} \beta^2 + \frac{1}{80} \beta^4$$

Varying the parameter values, we can derive some primary statistics (moments, dispersion, skewness and kurtosis), as reported in Table 3:

TABLE 3

Effects of the parameter values on moments, variability, skewness and kurtosis

	$\alpha = 0$ $\beta = 0$	$\alpha = 0.2$ $\beta = 0.2$	$\alpha = 0.3$ $\beta = 0.2$	$\alpha = 0.4$ $\beta = 0.2$	$\alpha = 0.1$ $\beta == 0.1$	$\alpha = 0.1$ $\beta = 0.2$	$\alpha = 0.1$ $\beta = 0.4$
$E(Y)$	0.4000	0.3867	0.3800	0.3733	0.3933	0.3933	0.3933
$E(Y^2)$	0.1778	0.1651	0.1588	0.1524	0.1712	0.1714	0.1724
$E(Y^3)$	0.0868	0.0775	0.0729	0.0682	0.0819	0.0821	0.0831
$E(Y^4)$	0.0460	0.0397	0.0366	0.0335	0.0427	0.0429	0.0436
$\sigma(Y)$	0.1333	0.1249	0.1199	0.1143	0.1284	0.1294	0.1332
$C.V.(Y)$	0.3333	0.3230	0.3155	0.3062	0.3264	0.3289	0.3386
$\mu_3(Y)$	0.0014	0.0016	0.0016	0.0015	0.0016	0.0015	0.0014
$\mu_4(Y)$	0.0011	0.0009	0.0009	0.0008	0.0010	0.0010	0.0011
$\gamma_1(Y)$	0.6071	0.8135	0.9218	1.0328	0.7461	0.7086	0.5733
$\gamma_2(Y)$	3.3616	3.8771	4.2316	4.6692	3.6740	3.5910	3.3615

These are the effects of increasing the value of α:
– expected value decreases slowly, and moves towards the value θ;
– standard deviation and the C.V. decrease more than linearly;
– skewness and kurtosis increase quickly and markedly.

On the other hand, these are the effects of increasing the value of β:
– expected value does not change (as said before, it does not depend on β);
– standard deviation and C.V. increase slowly;
– skewness decreases quickly and markedly;
– kurtosis decreases slowly and slightly.

3. SYMMETRIC STSP AND MTSP DISTRIBUTION

Under the condition $\theta = 1/2$, the STSP distribution is symmetric, and its p.d.f. assumes a simpler form:

$$f_X(x) = \begin{cases} k \cdot 2^{k-1} \cdot x^{k-1} \,, & \text{if } 0 < x \le 1/2 \\ k \cdot 2^{k-1} \cdot (1-x)^{k-1} \,, & \text{if } 1/2 < x \le 1 \end{cases} \tag{12}$$

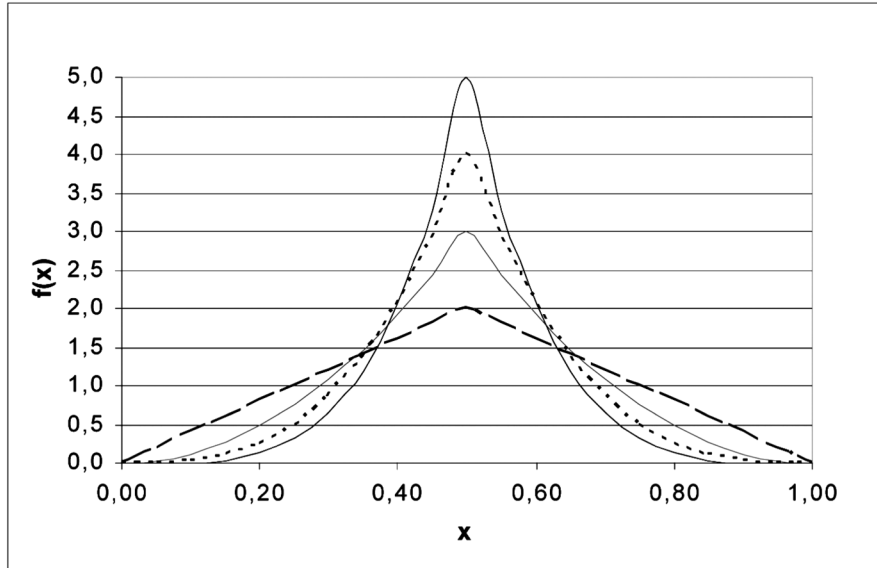In Figure 7 the p.d.f. of a symmetrical STSP distribution is reported, for some values of κ.

Figure 7 – Symmetrical STSP distribution ($\theta=0.5$, κ variable).

The corresponding c.d.f. is, as well, simplified:

$$F_X(x) = \begin{cases} 0 & x < 0 \\ \dfrac{(2x)^k}{2} & 0 \le x < 1/2 \\ 1 - \dfrac{[2(1-x)]^k}{2} & 1/2 \le x < 1 \\ 1 & x \ge 1 \end{cases} \tag{13}$$

The expected value of a symmetric STSP variable is, obviously, always equal to *1/2*, which is the axis of symmetry. The second moment and the variance are, respectively:

$$E(X^2) = \frac{1}{4} \frac{\kappa^2 + 3\kappa + 4}{\kappa^2 + 3\kappa + 2} \tag{14}$$

$$V(X) = \frac{\kappa - \dfrac{\kappa - 1}{2}}{(\kappa + 2)(\kappa + 1)^2} = \frac{1}{2(\kappa + 1)(\kappa + 2)} \tag{15}$$

In Table 4, we have represented the value of $E(X)$, $V(X)$ and $S(X)$ as a function of κ.

TABLE 4

Expected value and variance of a symmetric STSP for some values of κ

κ	$E(X)$	$V(X)$	$S(X)$
1	0.3333	0.0833	0.2887
2	0.2917	0.0417	0.2041
3	0.2750	0.0250	0.1581
4	0.2667	0.0167	0.1291
6	0.2589	0.0089	0.0945
10	0.2538	0.0038	0.0615

If we introduce the MTSP adjustment, described above, we get a symmetric MTSP distribution, whose p.d.f. is represented in Figure 8 and Figure 9, as a function of α and β, respectively.
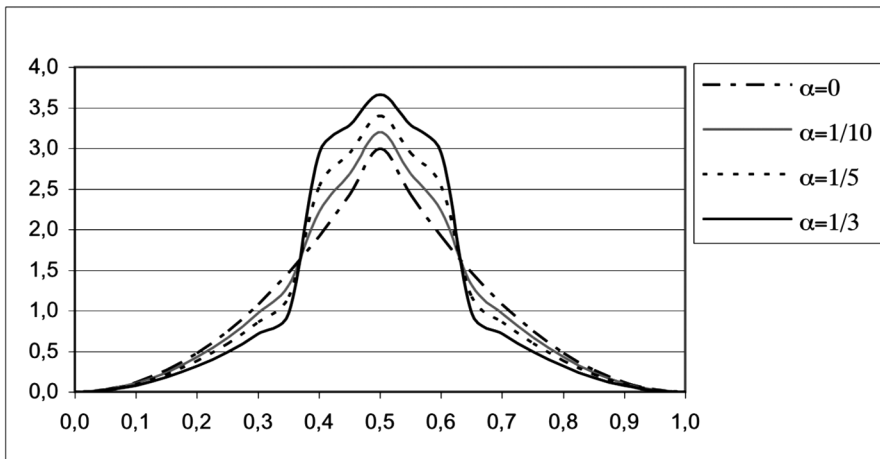

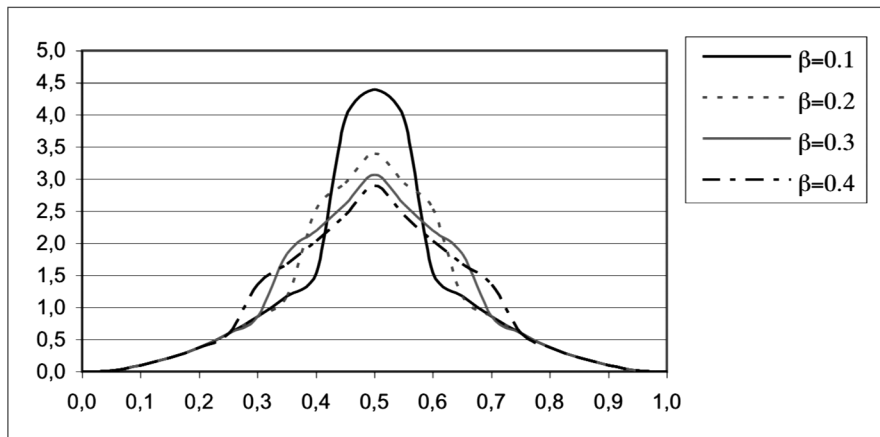*Figure 8* – Symmetric MTSP density ($\kappa=3$, $\beta=0.2$, α variable).

Figure 9 – Symmetric MTSP density ($\kappa=3$, $\alpha=0.2$, β variable).

In this context, the p.d.f. assumes, within the interval *[0, 1]*, these expressions:

$$f_Y(y) = \begin{cases} (1-\alpha)\ \kappa\ (2y)^{\kappa-1}, & 0 < y \le \dfrac{1-\beta}{2} \\[2ex] (1-\alpha)\ \kappa\ (2y)^{\kappa-1} + \dfrac{\alpha}{\beta}, & \dfrac{1-\beta}{2} < y \le \dfrac{1}{2} \\[2ex] (1-\alpha)\ [\kappa 2^{\kappa-1}(1-y)^{\kappa-1}] + \dfrac{\alpha}{\beta}, & \dfrac{1}{2} < y \le \dfrac{1+\beta}{2} \\[2ex] (1-\alpha)\ [\kappa 2^{\kappa-1}(1-y)^{\kappa-1}], & \dfrac{1+\beta}{2} < y \le 1 \end{cases} \tag{16}$$

The corresponding cumulative distribution function $F_Y(y)$ becomes then:

$$F_Y(y) = \begin{cases} (1-\alpha)\ 2^{\kappa-1}y^{\kappa}, & 0 < y \le \dfrac{1-\beta}{2} \\[2ex] (1-\alpha)\ 2^{\kappa-1}y^{\kappa} + \dfrac{\alpha}{\beta}\left(y - \dfrac{1-\beta}{2}\right), & \dfrac{1-\beta}{2} < y \le \dfrac{1}{2} \\[2ex] (1-\alpha)\ [1 - 2^{\kappa-1}(1-y)^{\kappa}] + \dfrac{\alpha}{\beta}\left(y - \dfrac{1-\beta}{2}\right), & \dfrac{1}{2} < y \le \dfrac{1+\beta}{2} \\[2ex] (1-\alpha)\ [1 - 2^{\kappa-1}(1-y)^{\kappa}], & \dfrac{1+\beta}{2} < y \le 1 \end{cases} \tag{17}$$

There is a relevant regularity in this model: the expected value of a symmetric MTSP, for any values of the parameters κ, α, β, is always equal to *1/2*.

On the other side, the second moment (and consequently the variance) is a function of all of three parameters κ, α, β, and specifically:

$$E(Y^2) \ = \ (1-\alpha)^2\ E(X^2) + \alpha^2\ E(\mathbf{Y}^2) = (1-a)\frac{\kappa^2 + 3\kappa + 4}{4(\kappa+1)(\kappa+2)} + a\left(\theta^2 + \frac{\beta^2}{12}\right) \tag{18}$$

$$V(Y) = E(Y^2) - 1/4 \tag{19}$$

4. PARAMETER ESTIMATION FOR MTSP DISTRIBUTION

In a MTSP variable we have to estimate four parameters (α, β, κ, θ). A possible approach may be an iterative procedure, in which we consider the two couples (κ, θ) and (α, β) separately. Suppose we have a sample of size *n*, say $y_1, y_2, ..., y_n$, and denote with $y(i)$ the *i*-th order statistic; we may estimate, first of all, the STSP parameters with the ML estimator proposed by Van Dorp and Kotz (2002), thus getting the estimators:

$$\begin{cases} \kappa^* = -\dfrac{n}{\ln M(r^*)} \\[3em] \theta^* = y(r^*) \end{cases} \tag{20}$$

where:

$$M(r) = \prod_{i=1}^{r-1} \frac{y_{(i)}}{y_{(r)}} \prod_{j=r+1}^{s} \frac{1-y_{(j)}}{1-y_{(r)}} \quad \text{and} \quad r^* = \text{argmax } M(r)$$

Once estimated κ and θ, we have reduced the likelihood function $L(a, \beta, \hat{\kappa}, \hat{\theta})$ to a bi-dimensional one. Moreover, we have estimated the point at which the pdf reaches its maximum and changes its expression. We have to detect the maximum of $L(a, \beta, \hat{\kappa}, \hat{\theta})$

One possible way may be an iterative procedure like this:

1) We fix an initial "reasonable" value for β, say β_0, and then "move" α from zero towards one, until $L(\alpha,\beta_0)$ reaches its maximum.

2) Suppose that $L_1=L(\alpha_1, \beta_0)$ is the maximal value. Then we can fix $\alpha =\alpha_1$ and move β to find (possibly) a "new" maximum $L_2=L(\alpha_1, \beta_1)$.

3) Then we can fix $\beta =\beta_1$ and move α again. If we get a new maximum we can go on the same way, varying alternatively one or another parameter.

Otherwise, we can start with α, instead of β. In this case, we have to consider an initial value $\alpha =\alpha_0$, and repeat the same steps. This iterative approach gives better results if the initial values are well chosen, i.e. if they are not too far from the "true" values of the parameters. If this procedure leads to $\alpha=0$ and/or to $\beta=0$, then we can fit the data with a STSP distribution model.

This estimation procedure is only a proposal, the problem of estimating α and β deserves a further and deeper investigation.

5. POSSIBLE EXTENSIONS PF STSP AND MTSP

The procedure that leaded us to the MTSP distribution may be extended in many ways. First of all, it is possible to change the domain from *[0,1]* to any closed interval *[a,b]* by applying the linear transformation:

$$W = a + (b-a)Y = a(1-Y) + bY \tag{21}$$

The expected value becomes then $E(W) = a + (b-a) E(Y)$, where $E(Y)$ is reported in (6), while the variance of W will be then equal to $(b-a)^2 \cdot V(Y)$.

Another possible extension may be given by moving the centre of the window (i.e. the mean value of the uniform component), from θ to another point of the domain. But this should split the domain in more parts, and the p.d.f. would become much more complicated. It is also possible to "open" more than one window, e.g. adding two uniform components in two distinct points of the domain.

Finally, it is even possible to define a multivariate STSP (as well as a multivariate MTSP); the simplest way should be to define a bivariate density function $f_{T,U}(t,u)$, whose marginal pdf's are STSP (or MTSP), with possibly different critical points (θ_T, θ_U) and the same value of κ. If the components are supposed independent, the joint p.d.f., referred to the domain $D_{T,U} \equiv [0,1] \times [0,1]$ is the following:

$$f_{T,U}(t,u) = \begin{cases} \kappa^2 \left(\dfrac{t \cdot u}{\theta_T \cdot \theta_U} \right)^{\kappa-1} & , \text{ if } 0 < t < \theta_T \, , \, 0 < u < \theta_U \\[3mm] \kappa^2 \left(\dfrac{t \cdot (1-u)}{\theta_T \cdot (1-\theta_U)} \right)^{\kappa-1} & , \text{ if } 0 < t < \theta_T \, , \, \theta_U < u < 1 \\[3mm] \kappa^2 \left(\dfrac{(1-t) \cdot u}{(1-\theta_T) \cdot \theta_U} \right)^{\kappa-1} & , \text{ if } \theta_T < t < 1 \, , \, 0 < u < \theta_U \\[3mm] \kappa^2 \left(\dfrac{(1-t) \cdot (1-u)}{(1-\theta_T) \cdot (1-\theta_U)} \right)^{\kappa-1} & , \text{ if } \theta_T < t < 1 \, , \, \theta_U < u < 1 \end{cases} \tag{22}$$

6. TWO FACED DISTRIBUTION

The recently proposed "Two-Faced distribution" (Brizzi, 2006) is a skewed distribution with a triangular left side and an exponential right side, divided by a hinge point θ which is the modal point. Let now Y be a r.v. following such a distribution.

Original parameters: θ (hinge point) , λ (exponential coefficient)

Probability density function:

$$f_Y(y) = \begin{cases} 0 \, , & y < 0 \\ \alpha \, y \, , & 0 < y < \theta \\ \beta e^{-\lambda y}, & y \geq \theta \end{cases} \tag{23}$$

where:

$$\alpha = \frac{2\lambda}{\theta(\lambda\theta + 2)}, \ \beta = \frac{2\lambda e^{\lambda\theta}}{\lambda\theta + 2}$$

It is useful to redefine the second parameter, denoting $\kappa = \lambda \cdot \theta$. Therefore, we can write the p.d.f. in a simpler way:

$$f_Y(y) = \begin{cases} 0 & \text{if } \ y < 0 \\ \dfrac{2\kappa}{\theta^2(\kappa + 2)} \ y & \text{if } \ 0 < y < \theta \\ \dfrac{2\kappa}{\theta(\kappa + 2)} e^{-\frac{\kappa}{\theta}(y - \theta)} & \text{if } \ y \geq \theta \end{cases} \tag{24}$$

where $\theta > 0$, $\kappa > 0$. In Figure 1 we have represented the p.d.f. as a function of κ; in Figure 2, we have represented it as a function of θ.
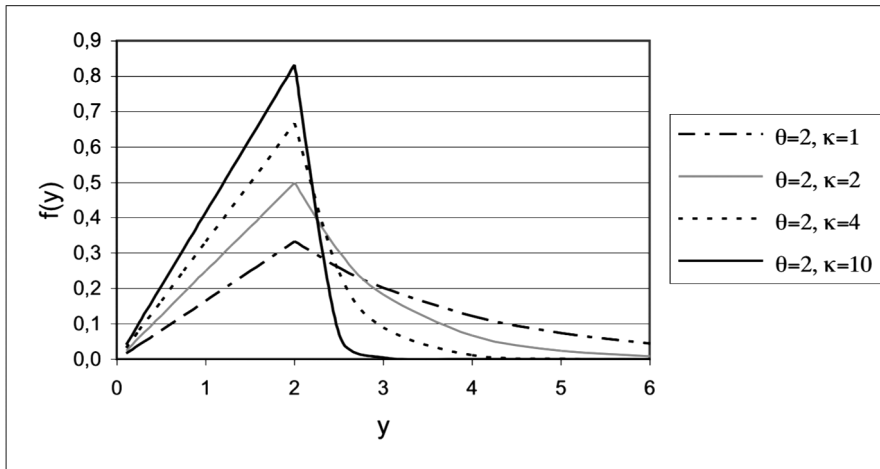

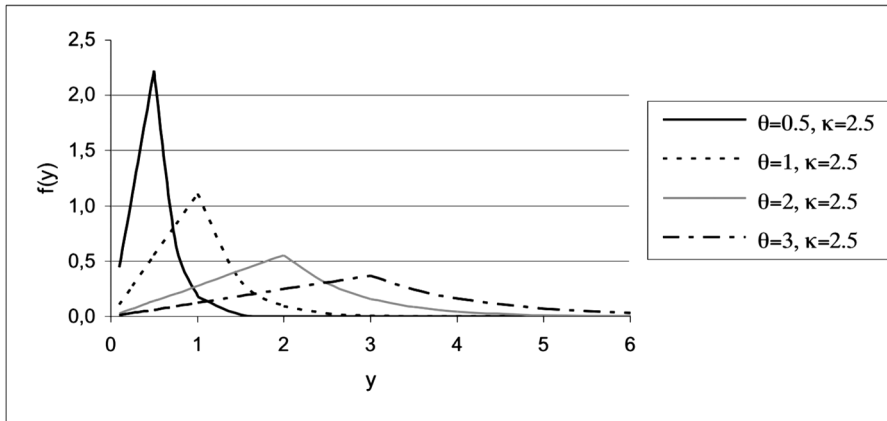*Figure 10* – Two-faced probability density function ($\theta=2$, κ variable).

Figure 11 – Two-faced probability density function (κ =2.5, θ variable)

The cumulative distribution function $F_Y(y)$ can be easily derived by integrating the density function given in (24):

$$F_Y(y) = \begin{cases} 0 & \text{if} \quad y < 0 \\[3mm] \dfrac{\kappa}{\theta^2(\kappa+2)}\, y^2 & \text{if} \quad 0 < y < \theta \\[3mm] 1 - \dfrac{2\mathrm{e}^{-\frac{\kappa}{\theta}(y-\theta)}}{\kappa+2} & \text{if} \quad y \geq \theta \end{cases} \tag{25}$$

The hinge point $y = \theta$ is the mode. The p.d.f. and c.d.f. of such point are, respectively:

$$f_Y(\theta) = \frac{2\lambda}{\lambda\theta+2} = \frac{2\kappa}{\theta(\kappa+2)}, F_Y(\theta) = \frac{\kappa}{\kappa+2} \tag{26}$$

If κ is less, equal or greater than *2*, the median is less, equal or greater than θ.

The Expected value of a two-faced distribution is to be calculated by solving two integrals, corresponding, respectively, to values less and greater than θ. It is useful to start from the expression of p.d.f. given in (23).

Left Integral (from *0* to θ):

$$LI = \int_0^\theta \alpha\, y^2 dy = \alpha \int_0^\theta y^2 dy = \frac{\alpha\theta^3}{3} = \frac{2\kappa\theta}{3(\kappa+2)} \tag{27}$$

Right Integral (from θ to $+\infty$):

$$RI = \int_{\theta}^{+\infty} \beta \ y e^{-\lambda y} \ dy = \beta \int_{\theta}^{+\infty} y e^{-\lambda y} \ dy = \frac{2\lambda e^{\lambda\theta}}{\lambda\theta+2} \frac{\lambda\theta+1}{\lambda^2} e^{-\lambda\theta} = \frac{2 \ (\lambda\theta+1)}{\lambda \ (\lambda\theta+2)} \tag{28}$$

Applying the usual simplifying notation $\kappa = \lambda\theta$, we have:

$$RI = \frac{2\theta \ (\kappa+1)}{\kappa \ (\kappa+2)} \tag{29}$$

Adding these two components we derive the expected value $E(Y)$:

$$E(Y) = LI + RI = \frac{2\kappa\theta}{3(\kappa+2)} + \frac{2\theta \ (\kappa+1)}{\kappa \ (\kappa+2)} = \frac{2\kappa^2\theta + 6\theta \ (\kappa+1)}{3\kappa \ (\kappa+2)} = \frac{2(\kappa^2 + 3\kappa + 3)}{3\kappa \ (\kappa+2)}\theta \tag{30}$$

In Table 5 we have specified the expected value of a two-faced variable Y for some parameter values:

TABLE 5

Expected value of TF *distribution for some parameter values*

$\kappa =$	0.5	1	1.5	2	$\sqrt{6}$	3	5	10
$\theta =$								
0.5	1.267	0.778	0.619	0.542	0.500	0.467	0.410	0.369
1	2.533	1.556	1.238	1.083	1.000	0.933	0.819	0.739
2	5.067	3.111	2.476	2.167	2.000	1.867	1.638	1.478
3	7.600	4.667	3.714	3.250	3.000	2.800	2.457	2.217
4	10.133	6.222	4.952	4.333	4.000	3.733	3.276	2.956
6	15.200	9.333	7.429	6.500	6.000	5.600	4.914	4.433
8	20.267	12.444	9.905	8.667	8.000	7.467	6.552	5.911
10	25.333	15.556	12.381	10.833	10.000	9.333	8.190	7.389

In particular, when κ is equal to $\sqrt{6}$, then we have $E(Y) = \theta$. Indeed, if we put $\kappa = \sqrt{6}$ in (30), we have:

$$E(Y) = \frac{2\theta(6 + 3\sqrt{6} + 3)}{3\sqrt{6} \ (\sqrt{6} + 2)} = \frac{2(9 + 3\sqrt{6})}{18 + 6\sqrt{6}} \theta = \theta. \tag{31}$$

Coming back to the general TF distribution, and using an analogous method, we derive the second moment:

$$E(Y^2) = \frac{\kappa^3 + 4\kappa^2 + 8\kappa + 8}{2\kappa^2(\kappa + 2)} \theta^2 \tag{32}$$

Consequently, we derive the variance of a TF distribution:

$$V(Y) = E(Y^2) - [E(Y)]^2 = \frac{\kappa^4 + 6\kappa^3 + 24\kappa^2 + 72\kappa + 72}{18\kappa^2(\kappa+2)^2}\theta^2. \tag{33}$$

In Table 6 we have computed the variance of a two-faced variable Y for some parameter values:

TABLE 6

Variance of TF distribution for some parameter values

$\kappa =$	0.5	1	1.5	2	3	4	5	10
$\theta =$								
0.5	1.021	0.270	0.131	0.082	0.046	0.033	0.027	0.019
1	4.082	1.080	0.523	0.326	0.184	0.133	0.109	0.074
2	16.33	4.321	2.091	1.306	0.738	0.534	0.437	0.296
3	36.74	9.722	4.704	2.938	1.660	1.201	0.982	0.666
4	65.32	17.28	8.363	5.222	2.951	2.136	1.747	1.185
6	146.96	38.89	18.82	11.75	6.640	4.806	3.930	2.666
8	261.26	69.14	33.45	20.89	11.80	8.543	6.986	4.739
10	408.22	108.02	52.27	32.64	18.44	13.35	10.92	7.404

Following the same pattern, we can easily derive the third and fourth moment of a TF distribution:

$$E(Y^3) = \frac{2(\kappa^4 + 5\kappa^3 + 15\kappa^2 + 30\kappa + 30)}{5\kappa^3(\kappa+2)}\theta^3 \tag{34}$$

$$E(Y^4) = \frac{\kappa^5 + 6\kappa^4 + 24\kappa^3 + 72\kappa^2 + 144\kappa + 144}{3\kappa^4(\kappa+2)}\theta^4 \tag{35}$$

and give a general formula for the *m*-th moment:

$$E(Y^m) = \frac{2}{m+2}\frac{\kappa+(m+2)\sum_{j=0}^{m}\frac{m!}{j!}\kappa^{j-m}}{\kappa+2}\theta^m \tag{36}$$

From the moments of third and fourth order we can derive the usual indices of skewness (γ_1) and kurtosis (γ_2). It is worth to observe that indices γ_1 and γ_2 do not depend on θ, due to the fact that moments are directly proportional to the corresponding powers of the parameter θ itself. We have reported, in Table 7, the first four moments, as well as the standard deviation *S(Y)*, and the indices of skewness γ_1 and kurtosis γ_2.

Looking at Table 3, we realise that moments are inversely correlated with κ; the level of skewness depends only on κ, being γ_1 positive for small values of this parameter and negative for larger values. The value of κ corresponding to $\gamma_1=0$ is, approximately, $\kappa = 6.15$. Nevertheless, ın this context, the p.d.f. will not be perfectly symmetric, having limited left tail and unlimited right tail, so this specific result seems to deserve some attention. Kurtosis, as well, decreases as κ becomes

larger, and the TF may be leptokurtic (small values of κ) or platikurtic (large values of κ). The threshold value is approximately $\kappa = 6.47$, corresponding to a normokurtic distribution (with $\gamma_2 = 3$, as in Gaussian model). It is interesting that threshold values, respectively corresponding to $\gamma_1 = 0$ and $\gamma_2 = 0$, are not coinciding but quite near to each other.

TABLE 7

Moments of TF distribution and related statistics

κ	E (Y)	E (Y²)	E(Y³)	E(Y⁴)	S(Y)	γ_1	γ_2
0.5	$2.533\,\theta$	$10.500\,\theta^2$	$63.280\,\theta^3$	$506.467\,\theta^4$	$2.020\,\theta$	1.939	8.760
1	$1.556\,\theta$	$3.500\,\theta^2$	$10.800\,\theta^3$	$43.444\,\theta^4$	$1.039\,\theta$	1.777	8.136
2	$1.083\,\theta$	$1.500\,\theta^2$	$2.575\,\theta^3$	$5.417\,\theta^4$	$0.571\,\theta$	1.302	6.465
$\sqrt{6}$	θ	$1.242\,\theta^2$	$1.853\,\theta^3$	$3.300\,\theta^4$	$0.492\,\theta$	1.070	5.763
3	$0.933\,\theta$	$1.056\,\theta^2$	$1.396\,\theta^3$	$2.141\,\theta^4$	$0.429\,\theta$	0.834	5.032
4	$0.861\,\theta$	$0.875\,\theta^2$	$1.006\,\theta^3$	$1.295\,\theta^4$	$0.365\,\theta$	0.469	4.074
5	$0.819\,\theta$	$0.780\,\theta^2$	$0.825\,\theta^3$	$0.955\,\theta^4$	$0.330\,\theta$	0.207	3.481
6	$0.792\,\theta$	$0.722\,\theta^2$	$0.724\,\theta^3$	$0.782\,\theta^4$	$0.309\,\theta$	0.023	3.119
8	$0.758\,\theta$	$0.656\,\theta^2$	$0.616\,\theta^3$	$0.615\,\theta^4$	$0.285\,\theta$	-0.203	2.749
10	$0.739\,\theta$	$0.620\,\theta^2$	$0.561\,\theta^3$	$0.536\,\theta^4$	$0.272\,\theta$	-0.324	2.590
15	$0.714\,\theta$	$0.576\,\theta^2$	$0.497\,\theta^3$	$0.450\,\theta^4$	$0.257\,\theta$	-0.456	2.458

7. PARAMETER ESTIMATION FOR A TWO – FACED DISTRIBUTION

The two-faced distributions, as defined and described above has two parameters to estimate: the hinge point θ and the shape parameter κ. If the hinge point value, say θ_0, is known, say the ML estimator of the shape parameter is:

$$\hat{\kappa} = \sqrt{1 + \dfrac{2n\theta_0}{\displaystyle\sum_{y_j > \theta_0}(y_j - \theta_0)}} - 1 \tag{37}$$

and it depends, essentially, on the right tail values of the sample. The main problem is to estimate θ, when unknown, since the p.d.f. depends directly on its value. Being θ the modal point, the following step-by-step procedure seems suitable:

a) Find the sample subset where the points seem to be more dense, thus identifying a set of (say) *n'* sample values which are not too far from each other;

b) Let θ be equal to each of the *n'* points belonging to the subset and calculate the corresponding maximum likelihood estimate κ^* by applying (35). We find, with this method, *n'* couples of parameter estimates $\{\theta_j^*, \kappa_j^*; j = 1,2,..., n'\}$

c) Calculate the log-likelihood of each couple of parameter estimates, and consider the couple $(\theta^{**}, \kappa^{**})$, yielding the maximum log-likelihood;

d) Change slightly the value of θ^{**} in both directions; if the resulting likelihood does not decrease, we can keep the couple $(\theta^{**}, \kappa^{**})$ as a maximum likelihood estimation.

A specific socio-demographic application of the TF distribution has been proposed in Brizzi (2006), concerning the distribution of municipal population of Emilia – Romagna, a region of northern Italy having a little more than 300 municipalities. The TF model resulted to be very well-fitting for such set of data.

7. CONCLUDING REMARKS

There are many natural and social phenomena having a different behaviour, depending on we are considering values which are smaller or greater than a threshold point. The distribution models representing variables related with such kind of phenomena do need to have a different shape on the left and on the right of the above mentioned threshold value, which becomes the hinge point of the distribution. It is surely useful to dispose of general models, which may include a great variety of particular cases; families of distributions like STSP, MTSP and TF, developed in this paper, may be appropriate in such context, due to the wide range of models included. Evidently, this subject is worth of further investigation, and it would be surely possible and useful to derive new hinge-point models by modifying or generalising the distribution families considered here. For instance, the triangular-exponential TF distribution may be extended by changing the shape of right tail section, which may allow us to define a triangular-gamma distribution, a triangular-gaussian and so on. It is also possible to define models with more that one hinge point. As always, such "new" models should be fitted to real data, in order to check their capability to represent some specific natural or social variable. Even here, as in many other fields, theory can give useful tools to practice, and practice may give a strong validation to theory.

Dipartimento di Scienze statistiche "Paolo Fortunati"   MAURIZIO BRIZZI
Università di Bologna

ACKNOWLEDGEMENTS

REFERENCES

M. BRIZZI (2006), *A skewed model combining triangular and exponential features: the two – faced distribution and its statistical properties*, "Austrian Journal of Statistics", 35, pp. 455-462.

N.L. JOHNSON (1997), *The triangular distribution as a proxy for the Beta distribution in risk analysis*, "The Statistician", 46, pp. 387-398.

N.L. JOHNSON, S. KOTZ (1999), *Non–smooth sailing or triangular distributions revisited after some 50 years*, "The Statistician", 48, pp. 179-187.

J.R. VAN DORP, S. KOTZ (2002a), *The standard two-sided power distribution and its properties, with applications to financial engineering*, "The American Statistician", 56, pp. 90-99.

J.R. VAN DORP, S. KOTZ (2002b), *A novel extension of triangular distribution and its parameter estimation*, "The Statistician", 51, pp. 63-79.

J.R. VAN DORP, S. KOTZ (2003), *Generalized trapezoidal distributions*, "Metrika", 58, pp. 85-97.

SUMMARY

Some considerations and developments about recently proposed families of distributions with hinge points

This paper deals with probability density functions with hinge points, useful when representing phenomena which have different statistical features, depending on the fact that a threshold point has been overtaken or not. The STSP distribution, proposed by Van Dorp and Kotz (2002 a) is recalled and developed, and a modified model (MTSP), a mixture of STSP and Uniform, is defined and studied. Moreover, the Two Faced distribution, already proposed in Brizzi (2006), is recalled and thoroughly analysed, and some estimation problems, related to both models considered, are faced by a semi-empirical approach.