International Journal of Uncertainty,
Fuzziness and Knowledge-Based Systems
Vol. 17, No. 3 (2009) 377–396

World Scientific
www.worldscientific.com

THE GENERALIZED BIPARABOLIC DISTRIBUTION

CATALINA BEATRIZ GARCÍA GARCÍA
Departamento de Métodos Cuantitativos para la Economía y la Empresa,
University of Granada, Campus Universitario de La Cartuja
Granada, 18071, Spain
cbgarcia@ugr.es

JOSÉ GARCÍA PÉREZ
Departamento de Economía Aplicada, University of Almería,
La Cañada de San Urbano, s/n
Almería, 04071, Spain
jgarcia@ual.es

SALVADOR CRUZ RAMBAUD
Departamento de Dirección y Gestión de Empresas,
University of Almería, La Cañada de San Urbano, s/n
Almería, 04071, Spain
scruz@ual.es

Received 26 February 2008
Revised 15 January 2009

Beta distributions have been applied in a variety of fields in part due to its similarity to the normal distribution while allowing for a larger flexibility of skewness and kurtosis coverage when compared to the normal distribution. In spite of these advantages, the two-sided power (TSP) distribution was presented as an alternative to the beta distribution to address some of its short-comings, such as not possessing a cumulative density function (cdf) in a closed form and a difficulty with the interpretation of its parameters. The introduction of the biparabolic distribution and its generalization in this paper may be thought of in the same vein. Similar to the TSP distribution, the generalized biparabolic (GBP) distribution also possesses a closed form cdf, but contrary to the TSP distribution its density function is smooth at the mode. We shall demonstrate, using a moment ratio diagram comparison, that the GBP distribution provides for a larger flexibility in skewness and kurtosis coverage than the beta distribution when restricted to the unimodal domain. A detailed mean-variance comparison of GBP, beta and TSP distributions is presented in a Project Evaluation and Review Technique (PERT) context. Finally, we shall fit a GBP distribution to an example of financial European stock data and demonstrate a favorable fit of the GBP distribution compared to other distributions that have traditionally been used in that field, including the beta distribution.

Keywords: Biparabolic; beta; TSP distribution; generating density; uncertainty; risk.

1. Introduction

The treatment of uncertainty is a problem of common interest for scientists. In general, going from uncertainty to risk is performed by assigning a probability distribution to the phenomenon to be studied and later by estimating its parameters. In order to do this, researchers sometimes resort either to the information supplied by experts or to the well-known elicitation methods. Following Garthwaite *et al.*,[1] "Elicitation is the process of formulating a person's knowledge and beliefs about one or more uncertain quantities into a (joint) probability distribution for those quantities". Another problem is that of calibration, that is to say, "the methodology (...) to reduce the level of variation among multiple experts participating in the elicitation".[2] Due to its flexibility the beta distribution has found some applications in a wide variety of fields. For example, the beta distribution has found a particular application in the context of the Project Evaluation and Review Technique (PERT) (see, e.g., Malcolm *et al.*[3]) to model the uncertainty on an activity duration. When modeling the uncertainty on physical variables, some examples include relative humidity,[4] soil moisture index,[5] and daily sunshine duration data.[6] Applications of the beta distribution in other fields are provided by Kirkpatrick and Levin[7] and, more recently, Sherrick *et al.*,[8] Ricciardi *et al.*[9] and Palekar.[10] The probability density function (pdf) of the beta distribution is given by:

$$f(x|\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{(x-a)^{\alpha-1}(b-x)^{\beta-1}}{(b-a)^{\alpha+\beta-1}}, \tag{1}$$

where $a < x < b,\ \alpha > 0,\ \beta > 0$.

In spite of its popularity, it has been documented that the beta pdf (1) suffers from some problems perhaps impeding an even wider application.[11–19] These problems pertain to the evaluation of its cdf (which is not available in a closed form and requires numerical procedures for its evaluation) and the interpretation of its parameters. Kotz and van Dorp[20] introduced a non-smooth alternative to the beta distribution called the TSP distribution with pdf:

$$f(x|a,\theta',b,n) = \frac{n}{b-a} \times \begin{cases} \left(\dfrac{x-a}{\theta'-a}\right)^{n-1}, & \text{if } a \le x \le \theta' \\ \left(\dfrac{b-x}{b-\theta'}\right)^{n-1}, & \text{if } \theta' \le x \le b \end{cases} \tag{2}$$

which does have a closed form cdf. Kotz and van Dorp[20] summarized this contribution, together with others, enlarging the available family of distributions with bounded support.

To facilitate beta parameters estimation, Malcolm *et al.*[3] introduced a procedure to solve for the four parameters of the pdf (1) through the specification of three (a lower bound a, a most likely value θ', and an upper bound b). This methodology of using three parameters to solve for four has been somewhat controversial[11–13,16,21,22] and resulted in Herrerías[23] who suggested a modified procedure for estimating the parameters of the pdf (1) by adding a fourth confidence parameter n to the lower and upper bounds, a and b, and the most likely value, θ', as suggested by Malcolm *et al.*[3] As it turns out to be the case, the confidence parameter n in the method of Herrerías[3] plays the same role in the mean value calculation of the TSP distribution (2), which is given by:

$$E[X] = \frac{a + (n-1)\theta' + b}{n+1}. \tag{3}$$

In this paper, we shall propose another alternative to the beta distribution and we will refer to it as the generalized biparabolic (GBP) distribution. Its pdf will be constructed in Section 3 and is given by:

$$g\{x|a, \theta', b, m\} = C(m) \times \begin{cases} \left(\dfrac{x-a}{\theta'-a}\right)^{2m} - 2\left(\dfrac{x-a}{\theta'-a}\right)^{m}, & \text{if } a \le x \le \theta' \\ \left(\dfrac{b-x}{b-\theta'}\right)^{2m} - 2\left(\dfrac{b-x}{b-\theta'}\right)^{m}, & \text{if } \theta' \le x \le b \end{cases} \tag{4}$$

where $C(m) = \frac{(2m+1)(m+1)}{(-3m-1)(b-a)}$. The pdf (4) is smooth over its entire domain as opposed to the TSP distribution which is non-differentiable at the threshold parameter θ'. However, similar to the TSP pdf (2), the GBP distribution has a closed form expression. Figure 1 presents the density function of the GBP distribution compared with the density function of the TSP distribution, and the shape of the noteworthy beta distribution presented by Malcolm *et al.*[3]

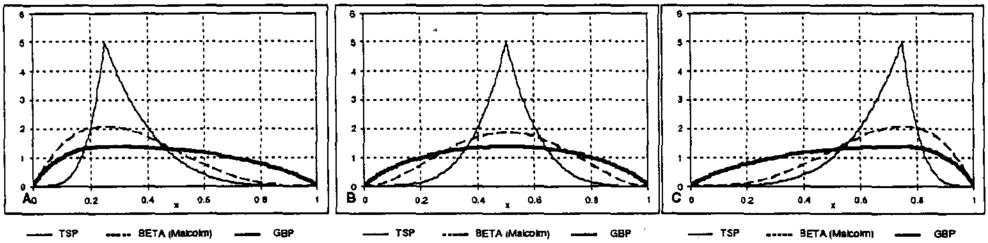

Fig. 1. Density function of the TSP distribution (thin solid), beta (Malcolm) distribution (dotted) and GBP distribution (thick solid) with a common weight for the most likely value in mean expressions. A. $\theta' = 0.25$, B. $\theta' = 0.5$, C. $\theta' = 0.75$.

The particular beta distribution presented in Figure 1 (also known as the original PERT beta distribution) is characterized by the weight for the most likely value in the mean value calculation being equal to 4, i.e., making $n = 4$ in (3). For comparison purposes, in Figure 1 we present the TSP and the GBP density functions with this same parameter. From Figure 1, observe that, while the TSP distribution is an alternative to the PERT beta distribution with a lesser variance, the GBP distribution has a higher variance. In Section 4, we shall insist more on the comparison between the distributions in Figure 1.

The rest of the paper is organized as follows. In Section 2 we present the biparabolic (BP) distribution and its geometric derivation with support $[a, b]$, which is reminiscent of the geometric construction of the unimodal triangular distribution, but forcing smoothness at the mode. Section 3 shows that the BP distribution can be generalized through the generating density in the same manner proposed by van Dorp and Kotz,[24] giving rise to the generalized biparabolic (GBP) distribution in a manner similar to the generalization of the triangular to a TSP distribution. The moments of the standard generalized biparabolic (SGBP) distribution, the closed forms of the mean, variance, as well as an analysis of its moment ratio diagram are also provided in this Section. Section 4 introduces the GBP distribution within the PERT methodology and provides a mean-variance comparison with other distributions traditionally applied in this field. In an example involving European stock data, in Section 5 we provide a favorable fit to the GBP distribution compared to the beta, the TSP and the asymmetric Laplace distributions. Finally, Section 6 summarizes our findings and concludes.

2. The BP Distribution

The biparabolic (BP) distribution arises, hence its name, from the conjunction of two parabolas which share the same vertex (see Figure 2). It can be constructed as follows: the values a, θ' and b determine the parabolas from $(a, 0)$ and the vertex (θ', h) and the parabola using the point $(b, 0)$ and the vertex (θ', h). With the result in bold (Figure 2.B), the BP density function is presented:

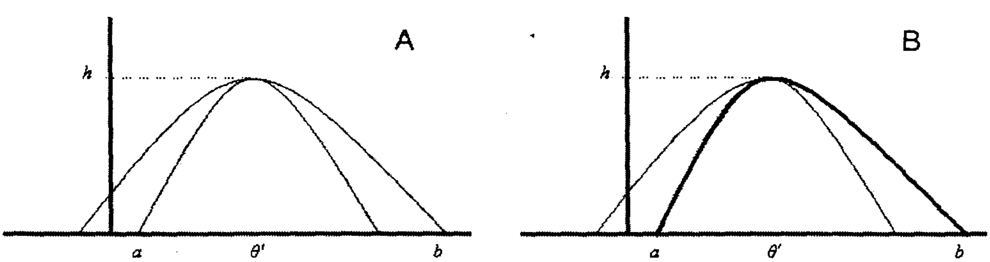

Fig. 2. Construction of the BP density function from the left and the right branches of two parabolas.

The density function of the BP distribution is defined as:

$$f(x) = \begin{cases} f_1(x), & \text{if } a \le x \le \theta' \\ f_2(x), & \text{if } \theta' \le x \le b \end{cases} \tag{5}$$

It can be deduced that:

$$f_1(x) = A(x-a)[x-(2\theta'-a)] \text{ and } f_2(x) = B(x-b)[x-(2\theta'-b)]. \tag{6}$$

The expressions of parameters A and B can be derived by requiring that $f(x)$ is a density function, $\int_a^b f(x)dx = 1$, and the continuity condition, $f_1(\theta') = f_2(\theta')$. Thus, it can be demonstrated that:

$$A = -\frac{3}{2}\frac{1}{(\theta'-a)^2(b-a)} \text{ and } B = -\frac{3}{2}\frac{1}{(b-\theta')^2(b-a)}, \tag{7}$$

obtaining the BP density function as:

$$f(x) = C(a,\theta',b) \times \begin{cases} (\theta'-b)^2[x^2 - 2\theta' x + (2\theta'-a)a], & \text{if } a \le x \le \theta' \\ (\theta'-a)^2[x^2 - 2\theta' x + (2\theta'-b)b], & \text{if } \theta' \le x \le b \end{cases} \tag{8}$$

where $C(a,\theta',b) = -\frac{3}{2}\frac{1}{(\theta'-a)^2(\theta'-b)^2(b-a)}$ and the distribution function as:

$$F(x) = C(a,\theta',b) \times \begin{cases} \dfrac{1}{2}\dfrac{(x-a)^2(-x+3\theta'-2a)}{(\theta'-a)^2(b-a)}, & \text{if } a \le x \le \theta' \\ 1 - \dfrac{1}{2}\dfrac{(x-b)^2(x-3\theta'+2b)}{(\theta'-b)^2(b-a)}, & \text{if } \theta' \le x \le b \end{cases} \tag{9}$$

By the change of variable $X = a + T(b-a)$, we can standardize the variable X and so the expression of the pdf (8) and the cdf (9), using the new variable $t \in [0,1]$, would be:

$$f(t) = -\frac{3}{2}\frac{1}{\theta^2(\theta-1)^2} \times \begin{cases} (\theta-1)^2[t^2 - 2\theta t], & \text{if } 0 \le t \le \theta \\ \theta^2[t^2 - 2\theta t + 2\theta - 1], & \text{if } \theta \le t \le 1 \end{cases} \tag{10}$$

and

$$F(t) = \begin{cases} \dfrac{1}{2}\dfrac{t^2(3\theta - t)}{\theta^2}, & \text{if } 0 \le t \le \theta \\ 1 - \dfrac{1}{2}\dfrac{(t-1)^2(t-3\theta+2)}{(\theta-1)^2}, & \text{if } \theta \le t \le 1 \end{cases} \tag{11}$$

where θ is the standardized parameter corresponding to θ'.

3. The GBP Distribution

Van Dorp and Kotz[24] demonstrate that, if $p(y|\Psi)$ is an appropriate density function defined on the interval $[0,1]$, with parameter or vector of parameters Ψ, the following unimodal density function can be constructed as a function of θ:

$$g\{t|\theta, p(y|\Psi)\} = \begin{cases} p\left(\dfrac{t}{\theta}|\Psi\right), & \text{if } 0 < t \le \theta \\ p\left(\dfrac{1-t}{1-\theta}|\Psi\right), & \text{if } \theta < t < 1 \end{cases} \tag{12}$$

where $p(y|\Psi)$ is the generating density of $g\{t|\theta, p(y|\Psi)\}$. The cdf associated with (11) is:

$$G\{t|\theta, p(y|\Psi)\} = \begin{cases} \theta P\left(\frac{t}{\theta}|\Psi\right), & \text{if } 0 < t \leq \theta \\ 1-(1-\theta)P\left(\frac{1-t}{1-\theta}|\Psi\right), & \text{if } \theta < t < 1 \end{cases} \tag{13}$$

where $P(y|\Psi)$ is the cdf of the generating density $p(y|\Psi)$. From (13), the following quantile function is obtained:

$$G^{-1}\{z|\theta, P^{-1}(y|\Psi)\} = \begin{cases} \theta P^{-1}\left(\frac{z}{\theta}|\Psi\right), & \text{if } 0 < z \leq \theta \\ 1-(1-\theta)P^{-1}\left(\frac{1-z}{1-\theta}|\Psi\right), & \text{if } \theta < z < 1 \end{cases} \tag{14}$$

Expression (15), presented by van Dorp and Kotz,[25] relates the moments of the generating variable Y with those of the variable T:

$$E[T^k] = \theta^{k+1}E[Y^k|\Psi] + \sum_{i=0}^{k}\binom{n}{i}(-1)^i(1-\theta)^{i+1}E[Y^i|\Psi]. \tag{15}$$

This framework shall allow us to construct the GBP distribution obtaining its pdf and cdf expressions as well as its moments expressions. Taking into account Eq. (8), we can choose the following generating density and its cdf:

$$p(y|\Psi) = -\frac{3}{2}(y^2 - 2y), \tag{16}$$

$$P(y|\Psi) = \frac{1}{2}(3y^2 - y^3) \tag{17}$$

and then $g\{t|M, p(y|\Psi)\}$ will be the density function (10) of the standardized bi-parabolic (SBP) distribution, denoted by SBP(0, θ, 1). By the change of variable, we obtain the BP distribution, denoted by BP(a, θ', b). Finally, by introducing a fourth exponential parameter m and imposing the usual conditions of a density function, we obtain the following generating density and its cdf:

$$p(y|m) = \frac{(2m+1)(m+1)}{3m+1}(2y^m - y^{2m}) \tag{18}$$

and

$$P(y|m) = \frac{2y^{m+1}(2m+1) - y^{2m+1}(m+1)}{3m+1}. \tag{19}$$

The generating density (17) allows us to construct densities that are smoother at the threshold parameter θ. In effect, the so-called standard generalized biparabolic (SGBP) distribution is obtained, whose density function will be:

$$g\{t|\theta, p(y|m)\} = C(m) \times \begin{cases} \left(\dfrac{t}{\theta}\right)^{2m} - 2\left(\dfrac{t}{\theta}\right)^{m}, & \text{if } 0 < t \leq \theta \\ \left(\dfrac{1-t}{1-\theta}\right)^{2m} - 2\left(\dfrac{1-t}{1-\theta}\right)^{m}, & \text{if } \theta < t < 1 \end{cases} \tag{20}$$

where $C(m) = \frac{(2m+1)(m+1)}{-3m-1}$. Finally, the cdf associated to (20) will be:

$$G\{t|\theta, p(y|m)\} = \begin{cases} C(m)\theta \dfrac{\left(\frac{t}{\theta}\right)^{2m+1}}{2m+1} - \dfrac{2\left(\frac{t}{\theta}\right)^{m+1}}{m+1}, & \text{if } 0 < t \leq \theta \\ 1 + C(m)(\theta - 1)\dfrac{\left(\frac{1-t}{1-\theta}\right)^{2m+1}}{2m+1} - \dfrac{2\left(\frac{1-t}{1-\theta}\right)^{m+1}}{m+1}, & \text{if } \theta < t < 1 \end{cases} \tag{21}$$

Unstandardizing expression (20), we can obtain expression (4) which was provided in the introduction. As shown in Figure 1, the GBP distribution adopts some shapes similar to the normal distribution with the difference that the GBP can be asymmetric. Letting $m \to \infty$, the GBP distribution converges to a degenerate distribution with all the probability mass located at a point, and, when $m \to 0$, it adopts the shape of the uniform distribution.

3.1. *The inverse cumulative distribution function*

In order to obtain the inverse of the BP cdf, the inverse of the generating density (17) is calculated as:

$$P_y^{-1}(z) = 1 + \frac{1}{C(z)} + C(z), \tag{22}$$

where $C(z) = \left(\frac{1}{-1+z+\sqrt{-2z+z^2}}\right)^{\frac{1}{3}}$. By substituting (22) in (14), the inverse of the BP cdf can be constructed. Unfortunately, the quantile functions $G^{-1}\{z\}$ of the GBP distribution are not available in a closed form. However, a numerical algorithm has been developed whose Microsoft Excel spreadsheet is available from the authors upon request. The aim of this algorithm is to obtain the value of y that makes expression (19) equal to z. Next, we present the proposed algorithm step by step:

- STEP 1: $E_1 = z^{\frac{1}{2m+1}}$, $E_2 = z^{\frac{1}{m+1}}$.
- STEP 2: If $E_1 < E_2$ then $\text{LB} = E_1, \text{UB} = E_2$. Else $\text{LB} = E_2, \text{UB} = E_1$.
- STEP 3: $\bar{y} = (LB + UB)/2$.
- STEP 4: If $|P(\bar{y}|m) - z| < \epsilon$ then STOP.
- STEP 5: If $|P(\bar{y}|m) - z| < z$ then $UB = \bar{y}$.
- STEP 6: If $|P(\bar{y}|m) - z| > z$ then $LB = \bar{y}$. (23)

The starting interval for the bisection method is determined in Steps 1 and 2. Note that E_1 and E_2 in Step 1 are the quantile functions of the mixture components of $P(y|m)$ and the interval $[LB, UB]$, constructed in Step 2, by design needs to contain the solution of equation $z = P(y|m)$. From Step 3 to 6, the structure of a standard bisection algorithm is constructed.

As an example, we calculate the median of the GBP distribution displayed in Figure 1.A, whose parameters are $\theta = 0.25$ and $m = 0.8165$. By using the algorithm above presented and implemented in Excel by the authors, we can obtain that $G^{-1}(0.5) = 0.4331$ as showed in Figures 3.A and 3.B.

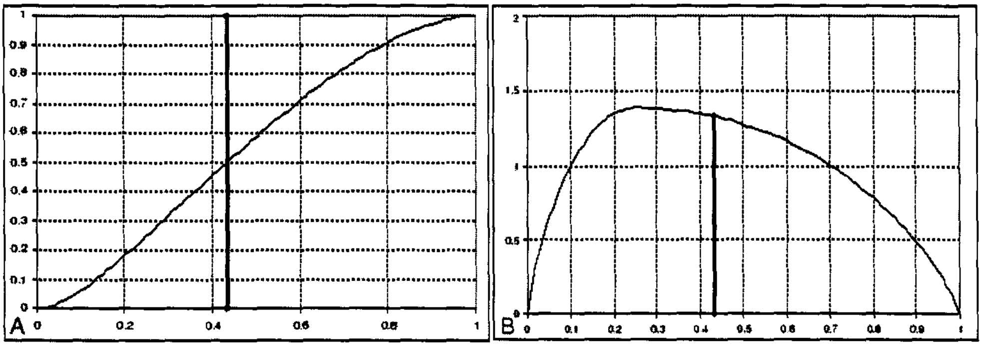

Fig. 3. Representation of the median value of the GBP distribution presented in Figure 1.A on the cumulative distribution function (A) and the probability distribution function (B).

Through this example we demonstrate that, although the GBP distribution does not have a closed expression for the inverse cumulative distribution function, a straightforward and computationally efficient algorithm (23) can be applied to evaluate the quantile function of this distribution.

3.2. *Moments expressions of the GBP distribution*

We can obtain the relation between the means and variances of variables T and Y:

$$E[T] = (2\theta - 1)E(Y|m) + (1 - \theta) \tag{24}$$

and

$$\mathrm{var}[T] = [\theta^3 + (1-\theta)^3]\mathrm{var}(Y|m) + \theta(1-\theta)[E(Y|m) - 1]^2. \tag{25}$$

At the same time, the moment of order k of the generating variable Y, defined on the interval $[0, 1]$, is:

$$E[Y^k] = \frac{(2m+1)(m+1)}{3m+1} \frac{3m+k+1}{(2m+k+1)(m+k+1)}. \tag{26}$$

Finally, the expressions of the expected value and the variance of the variable T that follows a GBP distribution are:

$$E[T] = \frac{6m^2\theta + 7m + 2}{6m^2 + 14m + 4} \tag{27}$$

and

$$\mathrm{var}[T] = \frac{1}{4(3m+1)^2(m+2)^2(2m+3)(m+3)} \times$$
$$[(148m^4 + 244m^3 + 40m^2)(\theta^2 - \theta) + (82m^4 + 247m^3 + 247m^2 + 96m + 12)]. \tag{28}$$

The expressions of the asymmetry and kurtosis coefficients could be obtained from (24), (26) and the well-known relation between the central and ordinary moments.[26] Closed form expressions for skewness and kurtosis of the GBP distribution are cumbersome and have been omitted. They are available from the authors upon request.

3.3. *The moment ratio diagram*

The moment ratio diagram[27] shall allow us to show the flexibility of the GBP distribution in terms of skewness and kurtosis coverage. In order to compare, Figures 4.A and 4.B present the moment ratio diagram of the GBP and beta densities, respectively. The parameter range of the GBP distribution is $0.01 < m < 1,000$ and $0 < \theta < 1$, and a comparable range of parameters α and β in (1) becomes $0.01 < \alpha < 1,000$ and $0.01 < \beta < 1,000$. Note that, in order to conserve the sign of μ_3, abscissa in Figures 4.A and 4.B is $\sqrt{\beta_1}$ instead of β_1.[28] On the other hand, theses figures present an infeasible (shaded) region, since for all distributions, $\beta_2 \geq \left(\sqrt{\beta_1}\right)^2 + 1$ holds.

The horizontally hatched area indicates the coverage of the unimodal GBP and beta distributions. It is observed that the beta family is richer than the GBP family when restricted to the U-shaped and the J-shaped regions since the GBP family does not exhibit these shapes. However, the coverage area of the beta family restricted to unimodal shapes in Figure 4.B is completely contained within the coverage area of the GBP family. Finally, possibly most important, the values of the kurtosis (β_2) for symmetric unimodal beta distributions with parameters restricted to $1 \leq \alpha \leq 1,000$ and $\beta = \alpha$ are strictly less than 2.88679, while, for symmetric unimodal GBP distributions, the values of the kurtosis can reach the value 5.0199. Therefore, it could be said that the GBP distribution has a larger flexibility in modeling skewed and peaked unimodal uncertainty phenomena.

4. The GBP Distribution in PERT Methodology

The beta distribution was proposed to be the underlying distribution of activity-time uncertainty in PERT methodology.[11] In the paper of Ben-Yair,[29] the use of

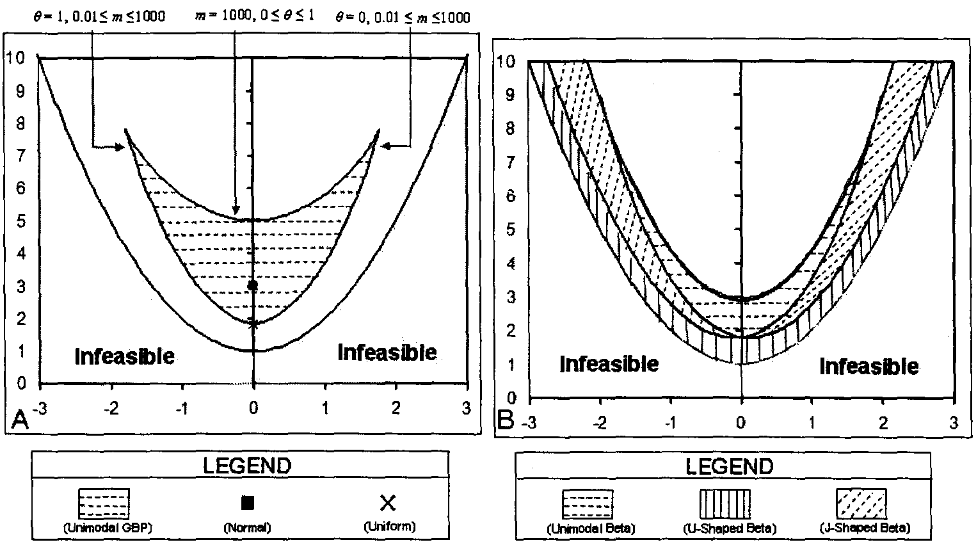

Fig. 4. $(\sqrt{\beta_1}, \beta_2)$ moment ratio diagram A. GBP distributions with parameter range $0.01 \leq m \leq 1{,}000$ and $0 \leq \theta \leq 1$. B. Beta distributions with parameter range $0.01 \leq \alpha \leq 1{,}000$, $0.01 \leq \beta \leq 1{,}000$.

a beta distribution is theoretically justified. As we are presenting the GBP distribution as an alternative to the beta and the TSP distributions, we shall compare their behavior within the PERT context.

PERT methodology originally assumed the beta distribution (1) as the underlying distribution and Malcolm *et al.*[3] suggested the following estimates of the mean and the variance values:

$$E[X] = \frac{a + 4\theta' + b}{6} \tag{29}$$

and

$$\mathrm{var}[X] = \frac{(b-a)^2}{36}. \tag{30}$$

Expressions (29) and (30) were suggested to overcome the difficulties with the interpretations and assessment of parameters α and β in (1). While the use of (29) and (30) results in the same kurtosis value as that of the normal distribution (i.e., 3), for the beta distribution, expressions (29) and (30) do not naturally follow from (1) and hence they have been subject to criticism since their introduction.[16,17,21,22,30]

Several researchers have proposed some modifications in the original PERT expressions. Most of them assume that the beta distribution is used to model the uncertainty in activity times. Herrerías[23] suggested a modification equivalent to (the original suggestion can be obtained substituting $n - 1 = s$):

$$\alpha = 1 + (n-1)\frac{\theta - a}{b - a} \text{ and } \beta = 1 + (n-1)\frac{b - \theta}{b - a}, \tag{31}$$

where $n > 0$ and $a < \theta < b$ in the beta pdf (1). The expressions of the mean and variance are given by:

$$E[X] = \frac{a + (n-1)\theta + b}{n+2} \tag{32}$$

and

$$\mathrm{var}[X] = \frac{n(b-a)^2 + (n-1)^2(b-\theta)(\theta-a)}{(n+2)(n+1)^2}. \tag{33}$$

On the other hand, various alternatives to the beta distribution have been suggested over time, such as triangular and, more recently, the two-sided power distribution and their generalizations[25] with the following expressions for the mean and the variance:

$$E[X] = \frac{a + (n-1)\theta + b}{n+1} \tag{34}$$

and

$$\mathrm{var}[X] = \frac{n(b-a)^2 - 2(n-1)(b-\theta)(\theta-a)}{(n+2)(n+1)^2}. \tag{35}$$

One of the main advantages of these alternatives is primarily that the cdf and quantile functions can be expressed in a closed form. García *et al.*[31] compared the use of the TSP and the beta distributions in a PERT context and concluded that the TSP distribution improves the behavior of the beta distribution in PERT methodology. However, the variance of the TSP distribution (34) in a unimodal domain is less than (33), as proposed by Herrerías.[23]

By taking the linear change $\theta = \frac{\theta'-a}{b-a}$ in expression (27), we obtain the following expression for the mean of a GBP distribution:

$$E[X|a, \theta', b, m] = \frac{(7m+2)a + 6m^2\theta' + (7m+2)b}{6m^2 + 14m + 4}. \tag{36}$$

To facilitate the comparison, we reparameterize (36) such that $6m^2 = n - 1$. The effect of the reparameterization above leads to the following expression for the mean value of a GBP distribution:

$$E[X|a, \theta', b, n] = \frac{f(n)a + (n-1)\theta' + f(n)b}{n - 1 + 2f(n)}, \tag{37}$$

where $f(n) = 7\sqrt{\frac{n-1}{6}} + 2$ and $n > 1$.

Notice that expressions (32), (34) and (37) have in common that each mean expression is a weighted average of parameters a, θ' and b, where the weights sum up to 1. Moreover, each expression weighs the most likely value with the value $(n-1)$ in its numerator. Only the GBP distribution weighs the bounds with the value $f(n)$ in the numerator, whereas both, the TSP and beta distributions, assign

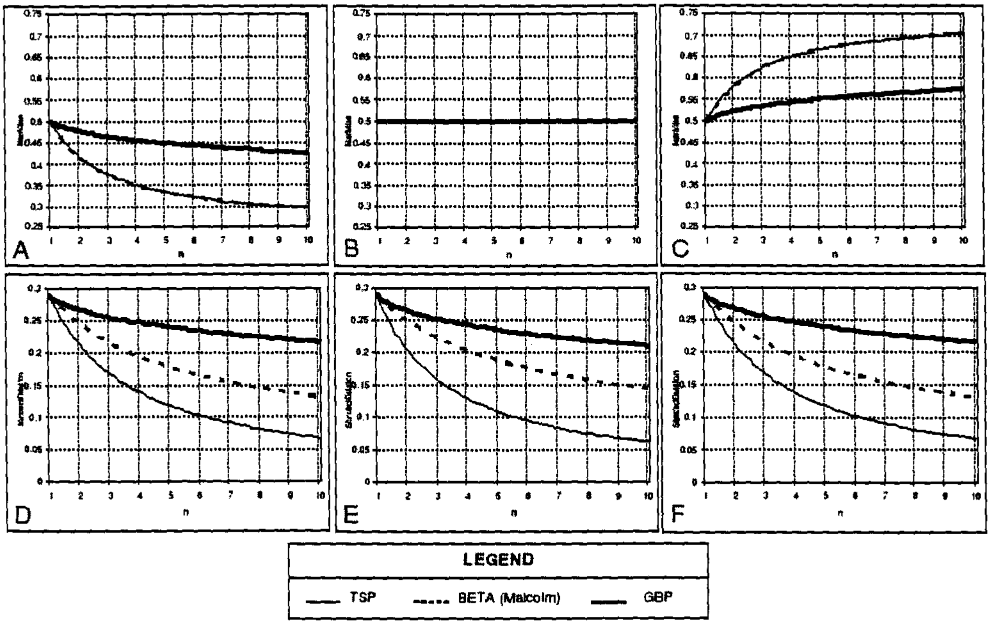

Fig. 5. Mean and variance of the GBP, TSP and beta distributions with a common weight of $(n-1)$ for the mode in mean value calculations. A. $\theta = 0.25$, B. $\theta = 0.5$, C. $\theta = 0.75$, D. $\theta = 0.25$, E. $\theta = 0.5$, F. $\theta = 0.75$.

here the value 1. Since $f(n) > 1$, as $n > 1$, it is deduced that the GBP distribution assigns a higher weight to the most likely value in the mean value calculation.

The value of n that equals the weights of both endpoints and the mode is calculated letting $f(n) = n-1$. This value is $n = 1.40407148$. Hence, for values of n greater than 1.40407148, the GBP distribution will weight the mode more than the endpoints, while, for lesser values, the GBP distribution will weight the endpoints more than the mode. The BP distribution, with pdf (8), is included in this last case.

Figure 5 plots on the y-axis the standardized mean (A, B and C) and the standard deviation (D, E and F) as a function of n for $n > 1$ (this coincides with the unimodal domain of the TSP and beta distributions), for different values of parameter θ. This is important to understand the effect of the choice of a certain distribution describing the uncertainty of completion time distribution in a network. From Figure 5, we can immediately conclude that, given a constant n, there is a strict order between the variance of these three families of distributions. Thus, the TSP variance is the smallest and the GBP is the largest one. On the other hand, the mean values of the beta and the TSP distributions agree for all values of n, whereas the mean value of the GBP distribution is larger (smaller), when $\theta < 0.5$ ($\theta > 0.5$), than that of the beta and the TSP distributions. The main conclusion from Figure 4 is that the GBP distribution exhibits a larger variance but a more moderate mean than TSP and beta distributions, while keeping the parameter n constant. Hence, regardless of the parameter value n, one can think of the GBP distributions being more conservative from an uncertainty analysis perspective than beta and TSP distributions, given the same lower bound, upper bound and most likely estimates.

5. An Example

We shall provide a data example within the financial field and compare the results using both the TSP and the beta distributions, as well as the asymmetric Laplace given by (38) which was recently proposed to be applied in this field by Kotz *et al.*[32]:

$$f(x|\theta,\kappa,\sigma) = \begin{cases} \dfrac{\sqrt{2}}{\sigma}\dfrac{\kappa}{1+\kappa^2}\left[-\sqrt{2}\dfrac{1}{\sigma\kappa}(\theta - x)\right], & \text{if } x < \theta \\ \\ \dfrac{\sqrt{2}}{\sigma}\dfrac{\kappa}{1+\kappa^2}\left[-\sqrt{2}\dfrac{\kappa}{\sigma}(x - \theta)\right], & \text{if } x \geq \theta \end{cases} \tag{38}$$

where κ, $\sigma > 0$ and $\theta \in \mathbb{R}$. More concretely, our database contains information about 1,000 daily price returns of BAY stock (DJ ES 50 market: for more details, see http://www.stoxx.com) from June 2001 to the end of May 2005. Fama[33] noted that the variable to be studied is $R_t = \ln\left(\frac{x_t}{x_{t-1}}\right)$, where x_t is the daily stock price in t $(t = 2, 3, 4, \ldots)$. The values x_t and one-step log differences (R_t) are displayed in Figures 6.A and 6.B, respectively.

With respect to the main characteristics of the data set, the mean is 0.00047, with a lower bound of -0.1843572 and an upper bound of 0.3228808. The standard deviation is 0.0006771. The coefficients of skewness and kurtosis are, respectively, 1.747 and 30.65. Because of the high coefficient of kurtosis, the normal distribution is not considered in this empirical application. Figure 7.A provides the empirical kernel density[34] and the empirical cdf is displayed in Figure 7.B.

The use of time series data requires to test for the existence of serial correlation which has been examined by using the auto-correlation function together with the Ljung-Box Q statistic. Table 1 contains the values of the auto-correlation function with lags equal to $1, 2, \ldots, 7$[35] together with the Ljung-Box Q statistic (LBQ). From the corresponding p-values, it immediately follows that the null hypothesis

Table 1. Auto-correlation function, Ljung-Box Q statistic and p-value for one step log differences.

Lag	1	2	3	4	5	6	7
ACF	0.0184	−0.0286	0.0312	0.0170	−0.0351	−0.0424	−0.0003
LBQ	0.3387	1.1589	2.1371	2.4294	3.6678	5.4840	5.4841
p-value	0.5606	0.5602	0.5444	0.6573	0.5982	0.4834	0.6011

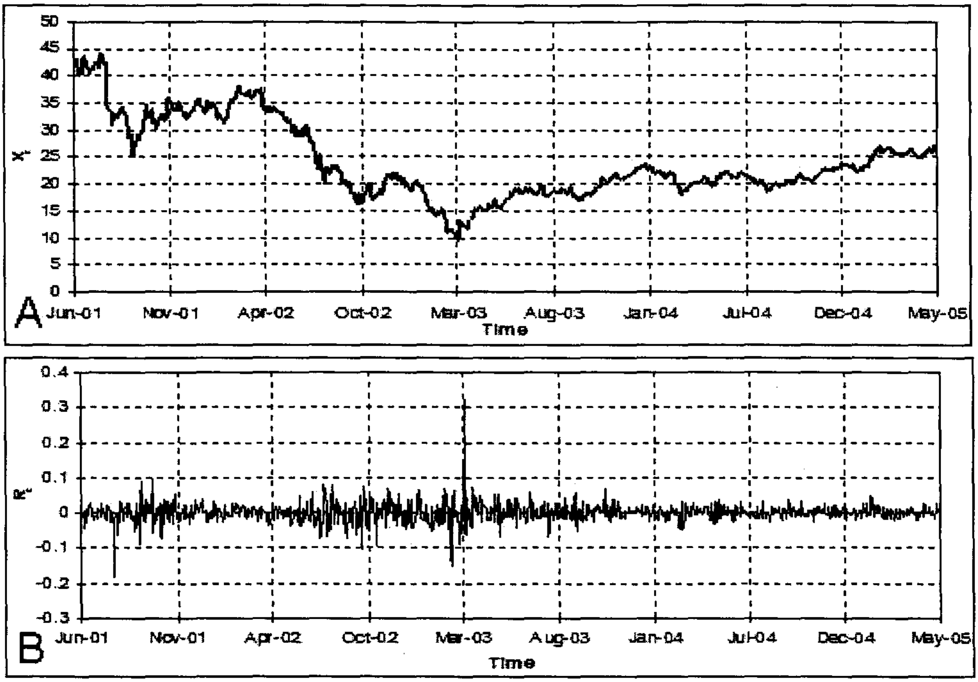

Fig. 6. A. BAY Daily stock prices (x_t), B. BAY stock prices one-step log-differences (R_t).

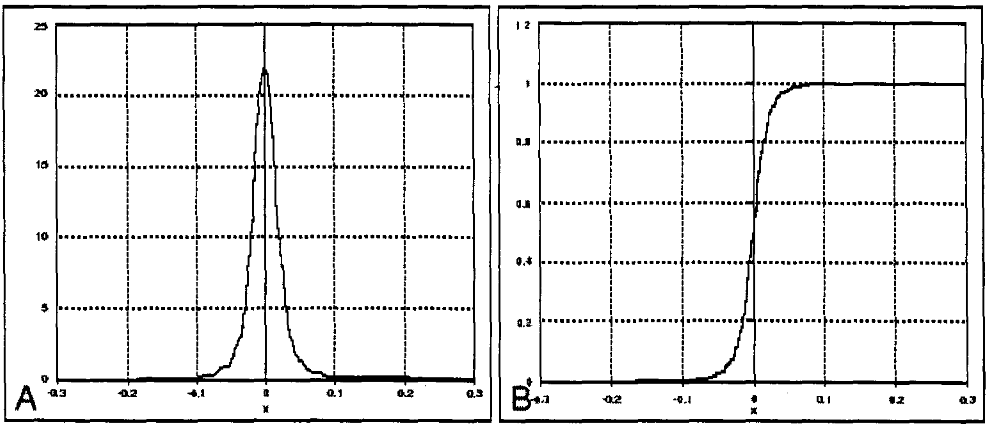

Fig. 7. A. Empirical kernel density estimate of R_t, B. Empirical cdf of R_t.

(i.e., the autocorrelations for all lags are equal to zero) is accepted. Therefore, we may reasonably conclude that the time series presented in Figure 6.B is serially uncorrelated.

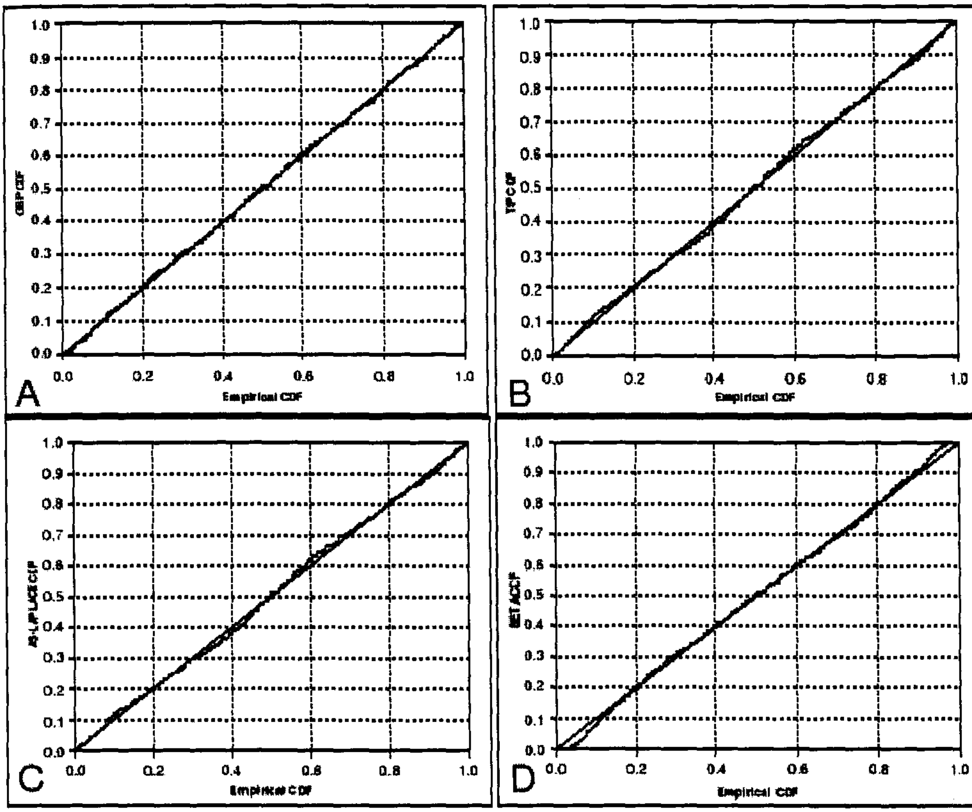

Fig. 8. QQ-Plot optimal pdf A. GBP distribution, B. TSP distribution, C. Asymmetric Laplace distribution, D. Beta distribution.

The method of least squares is used to determine the shape parameters (a, b, α, β) for the beta distribution, (a, θ', b, n) for the TSP distribution, (a, θ', b, m) for the GBP distribution, and (θ, σ, κ) for the asymmetric Laplace, that best fit to the data. The aim is to minimize the difference between the theoretical and the empirical cdf. We consider the lowest and highest value of the data set as the starting points for the lower bound (a) and the upper bound (b), respectively. The rest of the parameters have been changed until getting a relatively close fit to the empirical kernel density in Figure 7.A. Table 2 provides the starting solutions and resulting least squares estimates for the applied distributions.

Figure 8 depicts the QQ Plot of the empirical cdf and the optimal cdf of each distribution fitted to data. Figure 9 displays the pdf optimal solution. The GBP, the TSP and the asymmetric Laplace distributions appear to represent the sample data reasonably well, but the beta distribution appears to fit the data poorly.

While the plots of the empirical and the optimal fitted pdf provide useful visual evidence, the formal goodness of fit test is used to assess the adequacy of each distribution in representing the sample data with additional statistical evidence.

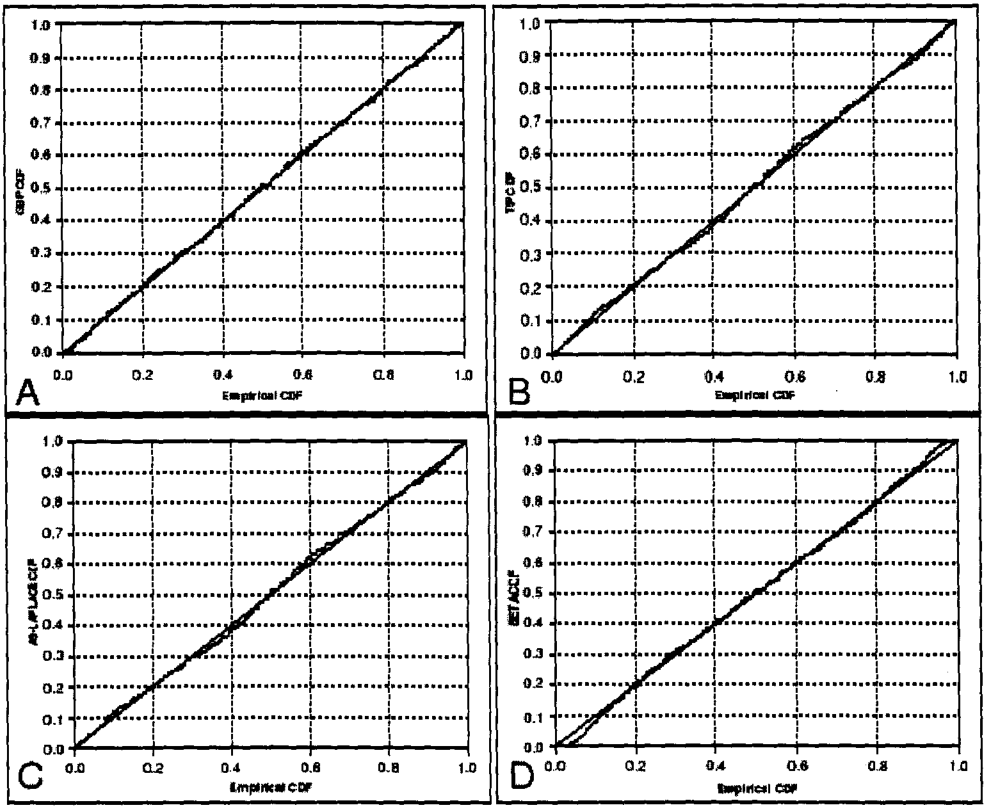

Fig. 9. Optimal pdf. A. GBP distribution, B. TSP distribution, C. Asymmetric Laplace distribution, D. Beta distribution.

Table 2. Starting points and optimal solutions obtained by the least squares method.

GBP		TSP		AL		BETA	
Start	Optimal	Start	Optimal	Start	Optimal	Start	Optimal
$a = -0.18436$	$a = 0.32012$	$a = -0.18436$	$a = -0.32665$	$\theta = 0$	$\theta = -0.00027$	$a = -0.1844$	$a = -0.18836$
$\theta' = 0$	$\theta' = -0.0006$	$\theta' = 0$	$\theta' = -0.00032$	$\sigma = 0.03$	$\sigma = 0.02418$	$b = 0.3228$	$b = 0.33054$
$b = 0.32288$	$b = 0.32288$	$b = 0.32288$	$b = 0.32288$	$\kappa = 1$	$\kappa = 1.07132$	$\alpha = 65.25468$	$\alpha = 65.25469$
$n = 15$	$m = 21.5496$	$n = 11$	$n = 18.50599$			$\beta = 114.74532$	$\beta = 114.7453$

The Chi-square statistic is calculated utilizing 32 bins;[36] therefore, its expression is given by:

$$\sum_{i=1}^{32} \frac{(O_i - E_i)^2}{E_i}, \tag{39}$$

where O_i is the number of observations of each bin and E_i the expected number of observations calculated as:

$$E_i = 1000 \times \{F(UB_i|\Theta) - F(LB_i|\Theta)\}. \tag{40}$$

Table 3. Bin boundaries and construction of the Chi-Square statistic (39) for the different applied distributions.

				GBP	TSP	AL	BETA
	LB_i	UB_i	O_i	$\frac{(O_i-E_i)^2}{E_i}$	$\frac{(O_i-E_i)^2}{E_i}$	$\frac{(O_i-E_i)^2}{E_i}$	$\frac{(O_i-E_i)^2}{E_i}$
1	< −0.18	−4.61E−02	31	5.05	0.00	0.51	128.11
2	−4.61E−02	−3.44E−02	31	0.02	0.40	0.32	1.42
3	−3.44E−02	−2.67E−02	31	2.36	1.97	1.52	3.97
4	−2.67E−02	−2.24E−02	32	0.22	0.00	0.02	1.96
5	−2.24E−02	−1.95E−02	32	0.22	1.06	1.46	0.25
6	−1.95E−02	−1.67E−02	31	0.05	0.14	0.28	1.24
7	−1.67E−02	−1.43E−02	30	0.29	0.01	0.04	1.59
8	−1.43E−02	−1.24E−02	32	0.04	0.62	0.82	0.15
9	−1.24E−02	−1.05E−02	31	0.02	0.14	0.21	0.29
10	−1.05E−02	−9.07E−03	31	0.87	1.90	2.08	0.50
11	−9.07E−03	−7.86E−03	31	2.32	3.51	3.67	2.05
12	−7.86E−03	−6.06E−03	32	0.72	0.41	0.39	0.60
13	−6.06E−03	−4.63E−03	31	0.00	0.00	0.00	0.05
14	−4.63E−03	−3.35E−03	31	0.14	0.03	0.02	0.52
15	−3.35E−03	−1.77E−03	31	0.81	1.91	2.13	0.20
16	−1.77E−03	0.00E+00	52	2.54	0.20	0.09	5.15
17	0.00E+00	1.07E−03	11	7.96	11.35	11.99	6.21
18	1.07E−03	2.43E−03	31	0.00	0.36	0.50	0.18
19	2.43E−03	3.84E−03	31	0.01	0.16	0.23	0.08
20	3.84E−03	5.45E−03	32	0.18	0.24	0.29	0.02
21	5.45E−03	6.48E−03	31	4.35	4.77	4.70	5.03
22	6.48E−03	8.21E−03	31	0.22	0.06	0.05	0.19
23	8.21E−03	9.77E−03	31	0.22	0.72	0.79	0.14
24	9.77E−03	1.15E−02	32	0.29	0.98	1.12	0.07
25	1.15E−02	1.31E−02	31	1.44	2.90	3.24	0.63
26	1.31E−02	1.61E−02	31	2.57	1.11	0.87	4.71
27	1.61E−02	1.87E−02	31	0.03	0.52	0.76	0.29
28	1.87E−02	2.18E−02	32	0.04	0.51	0.80	0.37
29	2.18E−02	2.51E−02	31	0.80	1.76	2.29	0.01
30	2.51E−02	3.25E−02	31	3.12	2.40	1.87	5.53
31	3.25E−02	4.39E−02	31	0.27	0.67	0.51	0.00
32	4.39E−02	> 0.32	32	2.80	0.02	0.65	48.01
		Total	1,000	40.00	40.80	44.20	219.55

The bin boundaries are presented in Table 3 and are selected to be the same for every analyzed distribution in a similar way that the "equal probability method of constructing classes".[37]

We have found that the beta distribution produces the worst fit with a joint contribution of about 176 to the tail bins 1 and 32. This fact indicates that the beta distribution is not able to fit distributions with high kurtosis and, at the same time, heavy tails. While the TSP distribution offers the lowest contribution to the tail bins, the GBP distribution is the best capturing the "peak" in this data set with the lowest contribution in bins 16 and 17. Taking into account that the degrees of freedom is $32 - k - 1$ (k = number of fitted parameters), we can calculate the p-value for the Chi Square Statistic, reported in Table 4, and conclude that the GBP

Table 4. Goodness of fit analysis of least squares fitted distributions.

	GBP	TSP	AL	BETA
p-value Chi-Square Test	0.0513	0.0430	0.0266	4.46E-32
LSQ	0.0269	0.0944	0.1246	0.1298
KS-Statistic	0.0132	0.0225	0.0243	0.0342

distribution outperforms the other distributions for the data example in question. Table 4 also includes the least square value and the Kolmogorov-Smirnov Statistic which lead to the same conclusion.

6. Concluding Remarks

For the treatment of risk and uncertainty several alternatives to the beta distribution have been suggested over time such as the rectangular, the triangular and, more recently, the two-sided power distributions and their generalizations.[25] All these distributions are unimodal and defined on a restricted domain which can be determined from the lower and upper bounds, and the most likely estimate supplied by an expert. This paper presents the GBP distribution which has these same characteristics and, moreover, is differentiable at the mode.

From the analysis of the moment ratio diagram, we can conclude that the family of beta distributions is richer than the family of GBP distributions when restricted to the U-shaped and the J-shaped areas. However, the coverage area of the beta family restricted to unimodal shapes is completely contained within the coverage area of the GBP family restricted to unimodal shapes. Therefore, as the beta distribution is applied to different fields, we propose the GBP distribution as an alternative to model unimodal phenomena and, concretely, when the smooth behavior of the density function at its mode is a requirement. We provide a favorable fit of the GBP distribution in the financial field when compared to the beta, the TSP and the asymmetric Laplace distributions.

Finally, we have compared the GBP distribution to the beta and the TSP distributions in a PERT context. Within such a context, the GBP distribution resulted in a higher variance than the beta and the TSP distributions given the same lower, upper and most likely estimates.

Acknowledgements

The authors would like to thank two anonymous referees whose suggestions greatly improved an earlier draft.

References

1. P. H. Garthwaite, J. B. Kadane and A. O'Hagan, Statistical methods for eliciting probability distributions, *J. Amer. Statistical Assoc.* **100**(470) (2005) 680-686.
2. R. Unal, C. B. Keating, T. M. Chytka and B. A. Conway, Calibration of expert judgments applied to uncertainty assessment, *Engineering Management Journal* **17**(2) (2005) 34.

3. D. Malcolm, J. Roseboom, C. Clark and W. Fazar, Application of a technique for research and development program evaluation, *Oper. Res.* **7** (1959) 646-669.
4. A. Y. M. Yao, A statistical model for the surface relative humidity, *J. Appl. Meteorol.* **15** (1974) 17-21.
5. A. C. Ravelo and W. L. Decker, The probability distribution of a soil moisture index, *Agr. Meteorol.* **20** (1979) 301-312.
6. F. Y. Ettoumi, A. Mefti, A. Adane and M. Y. Bouroubi, Statistical analysis of solar measurements in Algeria using beta distribution, *Renew. Energ.* **26** (2002) 47-67.
7. C. A. Kirkpatrick and R. I. Levin, *Planning and Control with PERT/CPM* (McGraw-Hill Inc., New York, 1966).
8. B. J. Sherrick, F. C. Zanini, G. D. Schnitkey and S. H. Irwin, Crop insurance valuation under alternative yield distributions, *Amer. J. Agr. Econ.* **86** (2004) 406-419.
9. K. L. Ricciardi, G. F. Pinder and K. Belitz, Comparison of the lognormal and beta distribution functions to describe the uncertainty in permeability, *J. Hydrol.* **313** (2005) 248-256.
10. U. S. Palekar, On estimating the distribution of optimal traveling salesman tour lengths using heuristics, *Eur. J. Oper. Res.* **186** (2008) 111-119.
11. C. E. Clark, The PERT model for the distribution of an activity time, *Oper. Res.* **10** (1962) 405-406.
12. F. E. Grubbs, Attempts to validate certain PERT statistics or a picking on PERT, *Oper. Res.* **10** (1962) 912-915.
13. J. J. Moder and E. G. Rogers, Judgment estimate of the moments of PERT type distributions, *Manage. Sci.* **18**(2) (1968) B76-B83.
14. S. D. Moitra, Skewness and the beta distribution, *J. Op. Res. Soc.* **41**(10) (1990) 953-961.
15. K. C. Chae and S. Kim, Estimating the mean and variante of PERT activity time likelihood of the mode and the midpoint, *I.I.E. Trans.* **22**(3) (1990) 198-203.
16. M. W. Sasieni, A note on PERT times, *Manage. Sci.* **32**(12) (1986) 1652-1653.
17. J. Kamburowski, New validations of PERT times, *Omega - Int. J. Manage. S.* **25**(3) (1997) 323-328.
18. D. Johnson, The triangular distribution as a proxy for the beta distribution risk analysis, *The Statistician* **46**(3) (1997) 387-398.
19. H. S. Lau, A. H. L. Lau and Y. Zhang, An improved approach for estimating the mean and standard deviation of a subjective probability distribution, *J. Forecasting* **16** (1997) 83-95.
20. S. Kotz and J. R. van Dorp, *Beyond Beta, Other Continuous Families of Distributions with Bounded Support and Applications* (World Scientific Press, Singapore, 2004).
21. C. Gallagher, A note on PERT assumptions, *Manage. Sci.* **33**(10) (1987) 1360.
22. T. K. Littlefield Jr. and P. H. Randolph, Another note on PERT times, *Manage. Sci.* **33** (1987) 1357-1359.
23. R. Herrerías, Utilización de modelos probabilísticos alternativos para el método PERT: aplicaciones al análisis de inversiones, *Estudios de Economía Aplicada*, Secretariado de Publicaciones de la Universidad de Valladolid (1989) 89-112.
24. J. R. Van Dorp and S. Kotz, Generalizations of two-sided power distributions and their convolution, *Commun. Stat. - Theor. M.* **32**(9) (2003) 1703-1723.
25. J. R. Van Dorp and S. Kotz, The standard two-sided power distribution and its properties: with applications in financial engineering, *Am. Stat.* **56**(2) (2002) 90-99.
26. A. Stuart and J. K. Ord, *Kendall's Advanced Theory of Statistics in Distribution Theory* (vol. 1) (Wiley, New York, 1994).

27. W. P. Elderton and N. L. Johnson, *Systems of Frequency Curves* (Cambridge University Press, London, 1969).
28. S. Kotz and N. L. Johnson, Moment ratio diagrams, in *Encyclopedia of Statistical Sciences* (vol. 5), eds. S. Kotz, C. B. Read, N. Balakrishnan and B. Vidakovic (Wiley, New York, 1985) pp. 602-604.
29. A. Ben-Yair, Upon implementing the beta distribution in project management, Working paper of the Department of Industrial Engineering and Management, Negev Academic College of Engineering, Israel. Available in:
    http://braude.ort.org.il/industrial/13thconf/html/.%5Cfiles%5C113_p.pdf
30. R. Herrerías, J. García and S. Cruz, A note on the reasonableness of PERT hypotheses, *Oper. Res. Lett.* **31**(1) (2003) 60-62.
31. J. García, S. Cruz and C. B. García, The two-sided power distribution for the treatment of uncertainty in PERT, *Statistical Methods and Applications* **14** (2005) 209-222.
32. S. Kotz, T. J. Kozubowski and K. Podgorrski, *The Laplace Distribution and Generalizations* (Birkhauser, Boston, 2001).
33. E. F. Fama, The behavior of stock market prices, *J. Bus.* **38** (1965) 34-105.
34. A. J. Izenman, Recent developments in nonparametric density estimation, *J. Am. Stat. Assoc.* **86**(413) (1991) 205-224.
35. R. S. Tsay, *Analysis of Financial Time Series* (Wiley, New York, 2002).
36. J. Banks, J. S. Carson, B. L. Nelson and D. M. Nicol, *Discrete-Event System Simulation* (4th ed.) (Prentice-Hall, Upper Saddle River, 2005).
37. A. Stuart, K. Ord and S. Arnold, *Kendall's Advanced Theory of Statistics (Vol. 2A): Classical Inference and the Linear Model* (6th ed.) (Edward Arnold, London, 1999).