2

The Polygonal Distribution

Dimitris Karlis[1] and Evdokia Xekalaki[2]

[1] Department of Statistics, Athens University of Economics and Business
[2] Department of Statistics, Athens University of Economics and Business
Department of Statistics, University of California, Berkeley

Abstract: The triangular distribution, although simpler than the beta distribution both for mathematical treatment and for natural interpretation, has not been widely used in the literature as a modelling tool. Applications of this distribution as an alternative to the beta distribution appear to be limited in financial contexts and specifically in the assessment of risk and uncertainty and in modelling prices associated with trading single securities. One of the basic reasons is that it can have only a few shapes. In this paper, a new class of distributions stemming from finite mixtures of the triangular distribution is introduced. Their polygonal shape makes them appealing for modelling purposes since they can be used as simple approximations to several distribution functions. Properties of these distributions are studied and parameter estimation is discussed. Further, the distributions arising when using the triangular distribution instead of the beta distribution as the mixing distribution in the case of two well-known beta mixtures, the beta-binomial and the beta-negative binomial distribution, are examined.

Keywords and phrases: Triangular distribution, binomial mixtures, negative binomial mixtures, triangular-binomial distribution

2.1 Introduction

The probability density function (pdf) of the triangular distribution is given by

$$f(x \mid \theta) = \begin{cases} \frac{2x}{\theta}, & 0 \leq x \leq \theta \\ \frac{2(1-x)}{1-\theta}, & \theta \leq x \leq 1 \\ 0, & \text{elsewhere} \end{cases} \tag{2.1}$$

with "0/0" interpreted as 1. The above definition restricts the random variable X in the interval $[0,1]$. One can define in a similar manner triangular distributions in a finite

B.C. Arnold et al. (eds.), *Advances in Mathematical and Statistical Modeling*,
DOI: 10.1007/978-0-8176-4626-4_2,

interval $[\alpha, \beta]$ by considering the transformation $Y = \frac{X-\alpha}{\beta-\alpha}$. From (2.1), one can see that the density is linearly increasing in the interval $[0, \theta)$ and linearly decreasing in the interval $[\theta, 1]$ (θ is the mode of the distribution). The distribution is not symmetric except for the case $\theta = 1/2$. The parameter θ is allowed to take the values 0 and 1, using the appropriate part of the definition given in (2.1). More details about the triangular distribution can be found in van Dorp and Kotz (2004) and the references therein. Johnson (1997) and Johnson and Kotz (1999) refocused interest in the triangular distribution, which appeared to have been ignored as a modeling tool over the last decades, one of the most probable basic reasons being that it can have only a few shapes.

In this paper, a new class of distributions is introduced stemming from finite mixtures of the triangular distribution. Contrary to the triangular distribution, the members of this class have a shape flexibility that makes them appealing for modeling purposes. Because of their shape, which is polygonal, these distributions are termed in the sequel polygonal distributions.

The paper is organized as follows. Following a brief presentation of the triangular distribution in Section 2.2, the polygonal distribution is defined as a finite mixture of triangular component distributions in Section 2.3. Properties of it and estimation are discussed. In Section 2.4, mixture distributions arising when using the triangular as an approximation to a beta mixing distribution are examined. In particular, the cases of beta mixtures of binomial and negative binomial distributions are considered. The paper concludes with some remarks in Section 2.5.

2.2 The Triangular Distribution

We briefly review some properties of the triangular distribution that can have potential use in the context of polygonal distributions.

The triangular distribution consists of two parts that are truncated forms of the $Beta(\alpha, \beta)$ distribution with density

$$f(x) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 \le x \le 1, \quad \alpha, \beta > 0, B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}. \quad (2.2)$$

In particular, the first part is a tail truncated $Beta(2, 1)$ distribution, while the second part a head truncated $Beta(1, 2)$ distribution, both truncated at θ. The triangular distribution also arises as the distribution of the mean of two uniform random variables.

The $s-th$ simple moment of the distribution is given by

$$\mu_s = \frac{2(1-\theta^{s+1})}{(s+1)(s+2)(1-\theta)}. \quad (2.3)$$

The above expression holds for not necessarily integer values of s, which enables computation of non-integral moments of a triangular variate.

Noting that $1 - \theta^{s+1} = (1-\theta) \sum_{i=0}^{s} \theta^i$, for integer $s \ge 0$, simple moments can be rewritten as

$$\mu_s = \frac{2\sum_{i=0}^{s} \theta^i}{(s+1)(s+2)}, \qquad s = 1, 2, \ldots \tag{2.4}$$

and can be computed recursively using

$$\mu_{s+1} = \frac{2\theta^{s+1}}{(s+2)(s+3)} + \left(\frac{s+1}{s+3}\right)\mu_s \tag{2.5}$$

for $s = 0, 1, \ldots$ with $\mu_0 = 1$. It can easily be verified that

$$E(X) = \frac{(1+\theta)}{3} \qquad \text{and} \qquad Var(X) = \frac{(1-\theta+\theta^2)}{18}$$

implying that the mean ranges from 1/3 to 2/3 and the variance becomes minimum at $\theta = 0.5$. It holds also that the first inverse moment has the form

$$\mu_{-1} = E(X^{-1}) = \frac{-2\log(\theta)}{(1-\theta)},$$

and that the second inverse moment does not exist since the corresponding integral diverges.

Random variate generation from the triangular distribution is simple via the inversion method. Finally, ML estimation is described in Johnson and Kotz (1999) and van Dorp and Kotz (2004). Note that since the ML estimate is necessarily one of the observations, it suffices to evaluate the likelihood at all the observations to locate the maximum. Some references about the triangular distribution can be found in Johnson and Kotz (1999) and van Dorp and Kotz (2004).

2.3 The Polygonal Distribution

Let $f_j(x \mid \theta_j),\ j = 1, 2, \ldots, k$ be the probability densities of k independent triangular variables on $[0, 1]$ with parameters $\theta_j,\ j = 1, \ldots, k$. A broad family of distributions stemming from these densities with interesting properties arises from a mixture of them defined by the probability density function

$$f_k(x) = \sum_{j=1}^{k} p_j f_j(x \mid \theta_j) \tag{2.6}$$

with mixing proportions $\{p_j\}_{j=1}^k$ satisfying $p_j > 0$, $1 \le j \le k$ and $\sum_{j=1}^k p_j = 1$.

This distribution has a polygonal form with at most k points of inflection. Some members of the family of distributions are depicted in figure 2.1. Observe that the densities are piecewise linear, a feature that offers great flexibility as far as shape is concerned. The probability density function defined by (2.6) and (2.1) can take shapes which are not common in other distributions. For example, a 2-polygonal distribution with $p = 0.5$ and $\theta_1 = 0.25, \theta_2 = 0.75$ is flat over the interval $(0.25, 0.75)$, having a

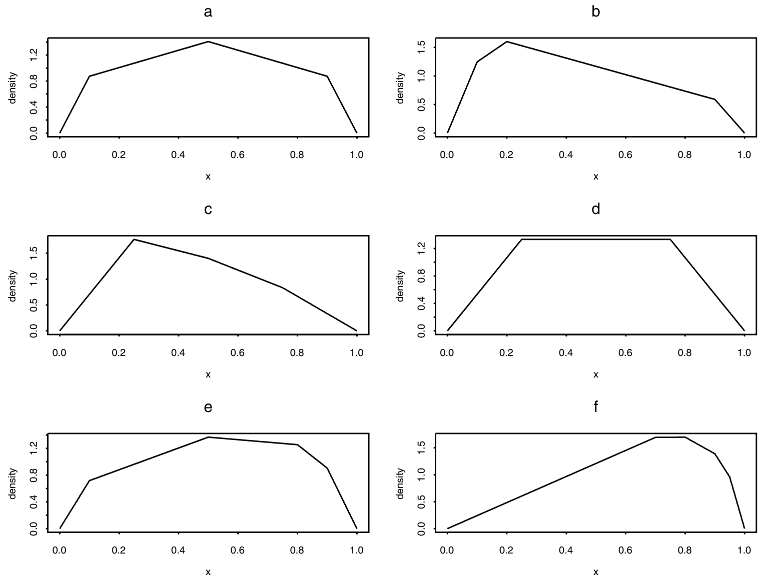

Figure 2.1. Examples of polygonal distributions. The depicted densities correspond to the following choices of parameter values: (a) $p_1 = p_2 = p_3 = 1/3$ and $\theta_1 = 0.1, \theta_2 = 0.5, \theta_3 = 0.9$, (b) $p_1 = 0.4, p_2 = 0.4, p_3 = 0.2$ and $\theta_1 = 0.1, \theta_2 = 0.2, \theta_3 = 0.9$, (c) $p_1 = 0.8, p_2 = 0.1, p_3 = 0.1$ and $\theta_1 = 0.25, \theta_2 = 0.5, \theta_3 = 0.25$, (d) $p_1 = p_2 = 0.5$ and $\theta_1 = 0.25, \theta_2 = 0.75$, (e) $p_1 = p_2 = p_3 = p_4 = 1/4$ and $\theta_1 = 0.1, \theta_2 = 0.5, p_3 = p_4 = 1/4$ and $\theta_1 = 0.7, \theta_2 = 0.8, \theta_3 = 0.9, \theta_4 = 0.95$

modal interval, rather than a mode. When components are close together and their number becomes larger and larger, the density approaches a smooth curve.

In the sequel, the distribution with k triangular components defined above is interchangeably referred to as the k-polygonal distribution or as the polygonal distribution. We also assume for simplicity that its components are ordered with respect to the values of their parameters θ_i.

The mean and variance of the polygonal distribution are given by

$$E(X) = \frac{1}{3} + \frac{1}{3}\sum_{j=1}^{k} p_j\theta_j \qquad \text{and}$$

$$Var(X) = \frac{1}{6} - \frac{1}{18}\sum_{j=1}^{k} p_j\theta_j + \frac{1}{6}\sum_{j=1}^{k} p_j{\theta_j}^2 - \frac{1}{9}\left(\sum_{j=1}^{k} p_j\theta_j\right)^2,$$

respectively. It can be seen that the mean lies in the interval $(1/3, 2/3)$.

Monotonicity It can be easily verified that the polygonal distribution has always a unique mode. This is supported by the fact that the segments of its density function over the intervals (θ_j, θ_{j+1}) given by

$$f_k(x) = \sum_{i=j+1}^{k} p_i \frac{2x}{\theta_i} + \sum_{i=1}^{j} p_i \frac{2(1-x)}{(1-\theta_i)}, \quad j = 0, \ldots, k, \quad \theta_0 = 0; \quad \theta_{k+1} = 1 \qquad (2.7)$$

have derivatives given by

$$f'_k(x) = 2\left(\sum_{i=j+1}^{k} \frac{p_i}{\theta_i} - \sum_{i=1}^{j} \frac{p_i}{(1-\theta_i)}\right). \tag{2.8}$$

Hence, the derivative of $f_k(x)$ is constant in each interval (θ_j, θ_{j+1}). It starts from a positive value, implying that the density is increasing in the first interval and, since in (2.8) at each interval a negative quantity replaces a positive one, continues to be increasing, but with a smaller and smaller slope over the following subintervals. As a result, the derivative corresponding to any of the sub-intervals is smaller than the derivative of the preceding sub-interval. The mode of the distribution occurs at the interval where the derivative becomes negative for the first time.

The position of the mode depends on the mixing proportions and it is not easy to be determined for general k. However, the mode is necessarily one of the $\theta's$ or a modal interval from θ_j to θ_{j+1}. For the mode of the 2-polygonal distribution, in particular, the following result holds

Proposition 1. *For a 2-polygonal distribution if* $p_1 > \frac{1-\theta_1}{\theta_2-\theta_1+1}$, *the mode is located at the point* θ_1, *otherwise, the mode is located at the point* θ_2 . *When* $p_1 = \frac{1-\theta_1}{\theta_2-\theta_1+1}$ *the distribution has a modal interval (the interval* (θ_1, θ_2)*) instead of a mode.*

For a proof, note that in accordance with the above, the density of the distribution is increasing in the first interval and decreasing in the third interval. Hence, the mode will be located at θ_1 or at θ_2 according as the derivative of the density over the second interval (θ_1, θ_2) given by $f'_k(x) = 2\left(\frac{p_2}{\theta_2} + \frac{p_1}{(1-\theta_1)}\right)$ is negative or positive; equivalently, according as p_1 exceeds or is exceeded by $\frac{1-\theta_1}{\theta_2-\theta_1+1}$. Note that if $p_1 = \frac{1-\theta_1}{1-\theta_1+\theta_2}$, the derivative $f'_k(x)$ equals 0, and thus the distribution has a modal interval from θ_1 to θ_2, where the density takes a constant value. Such a distribution will have a trapezoidal shape. Note that letting $\theta_1 \to 0$, while $\theta_2 \to 1$, the distribution tends to a uniform distribution.

2.3.1 Estimation

ML estimation can be carried out using the finite mixture representation via an EM algorithm. This comprises the following steps.

Step 1 (E-step): Given the current values for the parameters, say ${\theta_j}^{old}$ and ${p_j}^{old}$, $j = 1, \ldots, k$, calculate

$$w_{ij} = \frac{{p_j}^{old} f(x_i \mid {\theta_j}^{old})}{f_k(x_i)},$$

where $f(x \mid \theta)$ and $f_k(x)$ are given in (2.1) and (2.6) respectively.

Step 2 (M-step): For each component j, $j = 1, \ldots, k$ update p_j by

$$p_j^{new} = \sum_{i=1}^{n} w_{ij}/n,$$

then update θ_j by solving $L_j(\theta) = \sum_{i=1}^{n} w_{ij} \log f(x_i \mid \theta)$.

The maximization can be easily carried out, since the solution is one of the observations and thus evaluating L_j at all the observations suffices to locate the maximum. Note that w_{ij}'s do not depend on the estimate and hence the monotonicity of the likelihood holds as in Johnson and Kotz (1999).

2.4 The Polygonal Distribution as a Mixing Density

This section looks at the polygonal distribution as a mixing distribution in mixtures of discrete distributions $\{p_\theta(x); x = 0, 1, \ldots, m\}, m \in Z^+ \cup \{0\}$ with parameter $\theta \in (0, 1)$. These can obviously be seen as finite mixtures of triangular mixtures on θ of $p_\theta(\cdot)$ due to the associative property of finite mixtures. This is particularly appealing in the context of applications since then one can focus on triangular mixtures that are of a simpler structure. The application potential of the triangular and hence the polygonal distribution, is particularly enhanced in the area of mixtures by Johnson's (1997) result, which shows that any beta distribution can always be closely approximated by a triangular distribution.

In the remainder of this section we discuss the distributions to which two well known beta mixtures, the beta-binomial and the beta-negative binomial, are converted when the beta form of their mixing density is transitioned to a triangular form.

2.4.1 The binomial–triangular distribution

The binomial distribution is a prominent member of the family of discrete distributions. Mixtures of the binomial distribution with respect to the parameter p have been considered in the literature. Such mixtures have probability functions of the form

$$P(X = x) = \binom{n}{x} \int_0^1 p^x (1-p)^{n-x} dG(p), \qquad x = 0, 1, \ldots, n. \tag{2.9}$$

Note that $G(p)$ denotes a generic mixing distribution that can be either a finite step distribution giving positive probabilities at only a finite number of points or a continuous distribution. Some identifiability problems arise for small values of n (see, for example, Follmann and Lambert (1991)). The distribution is identifiable only up to the first n moments of the mixing distribution.

The beta-binomial (B-B) is the best known member of the family of binomial mixture distributions. It arises when the parameter p follows a beta distribution (see, for example, Tripathi and Gurland (1994) and the references therein). Only a few other binomial mixtures have been developed, mainly due to numerical difficulties (see Alanko and Duffy (1996), Horsnell (1957), Brooks et al. (1997)).

Assume that the parameter p has a triangular distribution given in (2.1). Then the resulting probability function is given by

$$P(X = x) = 2\binom{n}{x}\left(\frac{1}{\theta}\int_0^\theta p^{x+1}(1-p)^{n-x}dp + \frac{1}{1-\theta}\int_\theta^1 p^x(1-p)^{n-x+1}dp\right). \tag{2.10}$$

Both integrals are in fact incomplete beta integrals (see Abramowitz and Stegun (1974)) defined as $B_x(\alpha, \beta) = \int_0^x t^{\alpha-1}(1-t)^{\beta-1}dt$. Using the representation

$$I_x(\alpha, \beta) = \frac{B_x(\alpha, \beta)}{B(\alpha, \beta)},$$

with $I_x(\alpha, \beta) = 1 - I_{1-x}(\beta, \alpha)$ and $I_x(\alpha, \beta) = xI_x(\alpha-1, \beta) + (1-x)I_x(\alpha, \beta-1)$, and after tedious algebraical manipulations one can write the probability function of the binomial-triangular (B-T) distribution as

$$P(X = x) = 2\binom{n}{x}(\theta^{-1}B_\theta(x+2, n-x+1) + \\ + (1-\theta)^{-1}B(x+1, n-x+2) + (1-\theta)^{-1}B_\theta(x+1, n-x+2)).$$

This probability function is quite awkward for calculations as it involves incomplete beta functions. One can improve by considering recurrence relationships for beta integrals and incomplete beta integrals. A simpler method can be used for calculating the probabilities, based on a finite series representation of the probability mass function.

Sivaganesan and Berger (1993) showed that for a general $G(p)$ the resulting mixed binomial distribution can be written as

$$P(X = k) = \sum_{j=k}^{n} h(j, k)E(p^j), \qquad k = 0, 1, \ldots, n, \tag{2.11}$$

where $h(j, k) = (-1)^{j-k}\frac{n!}{k!(j-k)!(n-j)!}$ for $j \geq k$ and 0 if $j < k$, $E(p^r)$ denotes the $r-th$ simple moment of the mixing distribution. For the case of the triangular distribution, we obtain

$$P(X = k) = \sum_{j=k}^{n} h(j, k)\frac{2\sum_{i=0}^{j}\theta^i}{(j+1)(j+2)}, \qquad k = 0, 1, \ldots, n. \tag{2.12}$$

Computationally, this form is particularly convenient,since the coefficients $h(j, k)$ can be easily computed recursively using

$$h(0, 0) = 1, \qquad h(j+1, j+1) = \frac{n-j}{j+1}h(j, j), \quad j = 0, 1, \ldots, n \text{ and}$$

$$h(j+1, k) = -\frac{n-j}{j-k+1}h(j, k), \qquad j = k =, \ldots, n-1,$$

while the moments of the triangular distribution can be derived recursively. Evaluation of the probability function using the above form is easy and inexpensive. Even for large values of n near 200, no overflows were encountered for the entire range of values of θ.

A graphical comparison of the resulting distribution to the B-B distribution with the same mean and variance indicates a close agreement (see figure 2.2).

The mean and the variance of the B-T distribution are given by

$$E(X) = nE(p) = \frac{n(1+\theta)}{3} \quad \text{and} \quad Var(X) = \frac{n(n+3)}{18} - \frac{n(n-3)\theta(1-\theta)}{18},$$

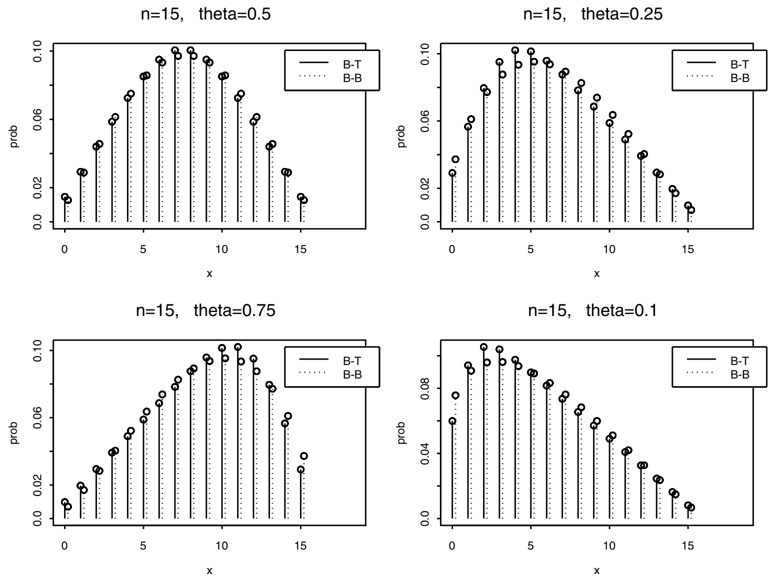

Figure 2.2. Plots of the probability function of the B-T distribution for $n = 15$ and $\theta = 0.5, 0.25, 0.75, 0.1$ superimposed by plots of the B-B with the same mean and variance

respectively. Note that there is a symmetry analogous to that existing in the case of the simple binomial distribution. So if X follows a B-T distribution with parameter θ, then $Y = 1 - X$ follows a B-T distribution with parameter $1 - \theta$.

The simple moments can be derived easily from the simple moments of the binomial distribution. It holds, in particular, that the simple moments of the B-T distribution are given by

$$\mu_r = 2 \sum_{j=0}^{r} \frac{S(r,j)n!}{(n-r)!} \frac{\sum_{i=0}^{j} \theta^i}{(j+1)(j+2)},$$

where $S(r, j)$ denote the Stirling numbers of the second kind .

Moment estimates of the parameter θ can be obtained through equating the mean with the sample mean. This yields the unbiased estimator $\hat{\theta} = 3n^{-1}\bar{x} - 1$, which leads to parameter estimates whenever $\bar{x}$ is in the range $(n/3, 2n/3)$. The variance of the moment estimator is given by $Var(\hat{\theta}) = \frac{9}{n^2} \frac{Var(X)}{N}$, where N denotes the sample size.

From (2.12), we can see that the probability function is a polynomial with respect the parameter θ. The same is true for the likelihood. Direct maximization is not easy because of the sum involved in the probability function, but grid search is not prohibitive since we have only one parameter distribution in a limited range of values.

2.4.2 An application

As an application of the B-T distribution illustrating a notable closeness to the B-B distribution, consider the data in table 2.1. The data refer to the numbers of courses taken by a class of 65 students from the first year of the Department of Statistics of Athens University of Economics. The students enrolled in this class attended 8 courses during the first year of their study. The total numbers of successful examinations (including resits) were recorded. For this data set, $n = 8$ and $\bar{x} = 5.2$.

The binomial distribution with $\hat{p} = 0.65$ provided a very poor fit. This was expected since it would not be reasonable to consider the probability of success p to be constant for all the students. Considering the students as having different probability of success according to their ability would be more natural.

Assuming that the probability of success varies according to a triangular distribution, the B-T distribution was fitted to the data. The moment estimate of θ was found to be 0.95. The likelihood was maximized at $\hat{\theta} = 1$. The maximized loglikelihood was -134.85. Assuming a beta distribution for p and fitting the data by the B-B distribution with parameter estimates $\hat{\alpha} = 1.825$ and $\hat{\beta} = 0.968$, yielded a maximized loglikelihood of -134.76. It is evident that the improvement of the loglikelihood from the B-T model to the B-B model is very small, taking into account that one parameter is added. The fits as judged by the χ^2 goodness of fit test give some indication of the closeness of the B-T distribution to the B-B distribution

Sivaganesan and Berger (1993) showed that for a general $G(p)$ the resulting posterior expectation of θ can be obtained as

$$E(\theta \mid X = k) = \frac{\sum_{j=k}^{n} h(j,k) E(p^{j+1})}{P(X = k)}, \qquad k = 0, 1, \ldots, n,$$

where $P(X = k)$ is given in (2.11), and it can be useful for Bayesian approaches, beyond the well known case of a conjugate Beta prior distribution. The values of $E(\theta \mid x)$ given in table 2.1 are indicative of a linear behavior, which is not in general true, and it is due to the value of θ estimated to be equal to 1.

Table 2.1. Data concerning the number of passed courses for a class of 65 students at the Dept. of Statistics, Athens University of Economics ($n = 8$). (The asterisk indicates grouped cells)

x	observed	expected			$E(\theta \mid x)$
		BB	BT	Binomial	
0	1	1.80	1.45	0.01*	0.1818
1	4	3.28	2.89	0.22*	0.2727
2	4	4.65	4.35	1.41*	0.3636
3	8	5.97	5.78	5.25	0.4545
4	9	7.25	7.23	12.18	0.5455
5	6	8.51	8.67	18.10	0.6364
6	8	9.78	10.12	16.82	0.7273
7	12	11.11	11.56	8.92	0.8182
8	13	12.65	12.97	2.08	0.9091
χ^2		1.45	3.15	105.4	
df		6	7	5	
p-value		0.96	0.88	0.00	

2.4.3 The negative binomial–triangular distribution

The triangular distribution can be also used as the mixing distribution for some other discrete distributions, having a parameter defined in the interval $[0,1]$. Such examples are the geometric and the negative binomial distributions. The negative binomial distribution has probability function given by

$$P(X=x)=\frac{\Gamma(\alpha+x)}{\Gamma(\alpha)x!}p^{\alpha}(1-p)^{x}, \qquad x=0,1,\ldots,n,\alpha>0,0\leq p\leq 1. \tag{2.13}$$

Mixtures of the negative binomial distribution with respect the parameter p can be developed by allowing the parameter p to vary according to some distribution $G(p)$. Such a mixture has probability function of the form

$$P(X=x)=\frac{\Gamma(\alpha+x)}{\Gamma(\alpha)x!}\int_0^1 p^{\alpha}(1-p)^{x}dG(p), \qquad x=0,1,\ldots,n. \tag{2.14}$$

The literature on mixtures of the negative binomial is rather sparse. Note that one can define mixtures with respect to either of the parameters α and p. Allowing $G(p)$ to have a $beta(\alpha,\beta)$ form, the generalized Waring distribution arises (see for example, Xekalaki (1983)).

Expanding $(1-p)^x$ in (2.14), one obtains that

$$\int_0^1 p^{\alpha}(1-p)^{x}dG(p)=\int_0^1 p^{\alpha}\sum_{k=0}^{x}\binom{x}{k}(-1)^{x-k}p^{x-k}dG(p)=$$
$$=\int_0^1\sum_{k=0}^{x}\binom{x}{k}(-1)^{x-k}p^{\alpha+x-k}dG(p)$$
$$=\sum_{k=0}^{x}\binom{x}{k}(-1)^{x-k}\int_0^1 p^{\alpha+x-k}dG(p)=\sum_{k=0}^{x}\binom{x}{k}(-1)^{x-k}E(p^{\alpha+x-k}),$$

thus leading to

$$P(X=x)=\frac{\Gamma(\alpha+x)}{\Gamma(\alpha)x!}\sum_{k=0}^{x}\binom{x}{k}(-1)^{x-k}E(p^{\alpha+x-k}).$$

In other words, the probability density function can be written as a finite series of non-integral moments of the mixing distribution.

Assuming a triangular distribution as a mixing distribution, one obtains the negative binomial-triangular distribution with probability function given by

$$P(X=x)=\frac{\Gamma(\alpha+x)}{\Gamma(\alpha)x!}\sum_{k=0}^{x}\binom{x}{k}\frac{(-1)^{x-k}2(1-\theta^{\alpha+x-k+1})}{(\alpha+x-k+1)(\alpha+x-k+2)(1-\theta)}.$$

The above formula can be used for calculating the probability function. A similar scheme as the one proposed for the binomial-triangular distribution is applicable. However, since now the values of x are not restricted in a finite range, minor anomalies may

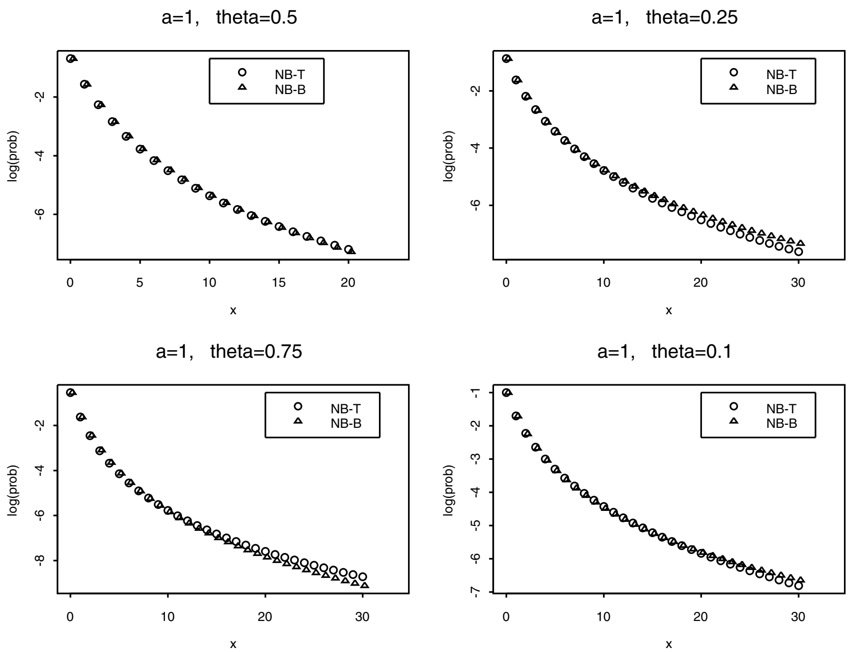

Figure 2.3. Plots of the probability function in log scale of the NB-T distribution for $\alpha = 1$ and $\theta = 0.5, 0.25, 0.75, 0.1$ superimposed by plots of the Generalized Waring (NB-B) distribution with the same mean

be found at the tail. Alternatively, the probability function can be written via incomplete beta functions in a similar manner as for the simple binomial case.

From figure 2.3, there appears a close agreement between the NB-T distribution and the NB-B (Generalized Waring) distribution with the same mean for $\alpha = 1$ and different values of θ.

Setting $\alpha = 1$, a geometric-triangular mixture is obtained. Now the moments used are of integral order and thus the recursive relationships for the moments of the triangular distribution can be used. Similar is the case when the Pascal distribution is considered.

The mean of the negative binomial-triangular distribution is

$$E(X) = \int_0^1 \frac{\alpha(1-p)}{p} g(p)dp = \alpha E(p^{-1}) - \alpha = \alpha \left(\frac{-2\log(\theta)}{1-\theta} - 1 \right).$$

Since $0 \leq \theta \leq 1$, it holds that $E(X) > \alpha$ for every value of θ. The variance does not exist, since it involves the second inverse moment of the triangular distribution which does not exist. The distribution exhibits a very long tail.

Note that mixtures of the negative binomial distribution with respect to the parameter p are in fact mixtures of the Poisson distribution, with mixing distribution a gamma mixture. For example, the negative binomial-triangular distribution defined above is a Poisson mixture with mixing distribution the mixture of a Gamma density with a triangular mixing density.

2.5 Discussion

A new class of distributions has been introduced stemming from the triangular distribution. Their polygonal shape offers them an appealing application potential and enhances their plausibility as modelling tools in areas ranging from risk analysis assessment, where its simplest member, the triangular distribution has been used, to developing envelope functions for rejection algorithms in simulation studies and as approximations to the beta distributions.

A notable feature is that the members of this family have always one mode (or modal interval), and the number of angles of the polygon depicting their density depends on the number of triangular components used in their finite mixture representation. In this sense, the polygonal distribution generalizes the trapezoidal distribution studied by van Dorp and Kotz (2003)

Finally, one may expand the definition of polygonal distributions beyond the interval $[0, 1]$ over a more general interval $[\alpha, \beta]$ using the transformation $y = \frac{x-\alpha}{\beta-\alpha}$. Alternatively, one may expand the polygonal distribution to the positive real line by the transformation $y = \frac{x}{1-x}$.

References

Abramowitz, M. and Stegun, I.A. (1974). *Handbook of Mathematical Functions.* Dover, New York.

Alanko, T. and Duffy, J. (1996). Compound binomial distributions for modelling consumption data. *The Statistician,* 45:269–286.

Brooks, S. P., Morgan, B. J. T., Ridout, M. S. and Pack, S. E. (1997). Finite mixture models for proportions. *Biometrics*, 53:1097–1115.

Follmann, D. and Lambert, D. (1991). Identifiability of finite mixtures of logistic regression models. *Journal of Statistical Planning and Inference*, 27:375–381.

Horsnell, G. (1957). Economic acceptance sampling schemes. *Journal of the Royal Statistical Society, Series A*, 120:148–191.

Johnson, D. (1997). The triangular distribution as a proxy for the beta distribution in risk analysis. *The Statistician*, 46:387–398.

Johnson, N. and Kotz, S. (1999). Non-smooth sailing or triangular distributions revisited after some 50 years. *The Statistician*, 48:179–187.

Sivaganesan, S. and Berger, J. O. (1993). Robust Bayesian analysis of the binomial empirical Bayes problem. *The Canadian Journal of Statistics*, 21:107–119.

Tripathi, G. R., R. and Gurland, J. (1994). Estimation of the parameters in the beta binomial model. *Annals of the Institute of Statistical Mathematics*, 46:317–331.

van Dorp, J. and Kotz, S. (2003). Generalized trapezoidal distributions. *Metrika,* 58:85–97.

van Dorp, J. and Kotz, S. (2004). *Beyond Beta: Other continuous distributions with bounded support and applications*. World Scientific, New York.

Xekalaki, E. (1983). The univariate generalized waring distribution in relation to accidents theory: Proneness, spells or contagion? *Biometrics*, 39:887–895.