Chapter 1

The Triangular Distribution

One of our goals in this book is to "dig out" suitable substitutes of the beta distribution. Only recently (less than 10 years ago) has the triangular distribution *specifically* been investigated by D. Johnson (1997) as a *proxy for the beta distribution*, even though its origins can be traced back to Thomas Simpson (1755) (about one century after the discovery of the beta distribution in a letter from Sir Isaac Newton to Henry Oldenberg). Very recently a "Handbook of Beta Distributions" edited by Gupta and Nadarajah (2004) has appeared (providing and emphasizing in a single monograph the attention that the beta distribution has attracted by both statistical theoreticians and practitioners over the last century, or so). On the other hand it appears that, in our opinion, the triangular distribution has been somewhat neglected in the statistical literature (perhaps even due to its simplicity which may discourage research efforts). In this chapter, we shall attempt to provide some chronology regarding the history of this distribution, state some of its properties and describe methods for estimating its parameters. Although the exposition is certainly not complete, we hope that it becomes apparent that the triangular distributions' "simplicity" is to a certain extent wrongly perceived and these distributions and their extensions are certainly worthy of further investigations.

1.1 An Historical Overview

Written records on the triangular distribution seem to originate in the middle of the 18-th century when problems of combinatorial probability were at their peak. A historically inclined reader may wish to consult the classical book by F.N. David (1962). One of the earliest mentions of the triangular distributions seems to be in Simpson[1] (1755, 1757). Thomas Simpson was a colorful personality in Georgian England. His life and adventures are described – in somewhat unflattering terms – in Pearson (1978) and Hald (1990). (Stigler (1986) gives a more sympathetic assessment of Simpson's work and character.) Stigler (1984) and more recently Farebrother (1990) provide some additional details on Thomas Simpson in particular on the correspondence with Roger Boscovich (1711-1787) a famous Italian astronomer and statistician of Serbian origin. The correspondence deals with the method of least absolute deviations regression problem which indirectly relates to triangular distributions (see, Farebrother (1990) and Stigler (1984)).

According to Seal (1949), Simpsons' object was to consider mathematically the method 'practised by Astronomers' of taking the mean of several observational readings "in order to diminish the errors arising from the imperfection of instruments and of the organs of sense". He supposes that any one reading errors in excess or defects are symmetrically disposed and have assignable upper and lower limits. He gives the probability that the mean of n observations falls between the boundaries $\pm z$ for the following discrete asymmetric triangular probability law:

$$p(x) = \begin{cases} \frac{h+1-|x|}{(h+1)^2} & x = \ -h,\ -h+1, \ldots -1, 0, 1, \ldots, h. \\ 0 & x = \text{ any other value.} \end{cases} \tag{1.1}$$

The solution for the case of a uniform discrete distribution, expressed as a gaming problem via a generalized die with k faces, was known by 1710, and Simpson's treatment by means of generating functions is the same as Abraham de Moivre[2] (1667-1754) (Todhunter (1865), p.85, Hald (1990)) which caused accusations of plagiarism. What is novel in Simpson's work appears in the four pages of additional material published in 1757. Here he extends the solution for the triangular case (1.1) to the limiting case $h \to \infty$ in such a way that the range of variation of an individual error remains

[1]Thomas Simpson (1710-1761) a prolific writer of mathematical textbooks and able teacher at the Royal Military Acadamy in Wolwich England has made original and important contributions to actuarial sciences.

[2]Abraham De Moivre (from a Huguenot family) left France in 1685 to seek asylum in England. He was a promininent probabilist who was the first to provide the normal approximation to the binomial distribution.

within ± 1. Seal (1949) points out that this is the *first* time a continuous (symmetric triangular) probability law is introduced. Hence, the continuous triangular distribution is certainly amongst the first *continuous* distributions to have been noticed by investigators during the 18-th century (when these types of problems were popular). For example, one of the first records that mentions the *continuous* uniform distribution is the famous paper by the reverend Thomas Bayes (1763) (only a few years after Simpons' written records in 1757).

The symmetric triangular distribution with probability density function (pdf)

$$f(x) = \begin{cases} 4x, & \text{for } 0 \leq x \leq \frac{1}{2}, \\ 4(1-x), & \text{for } \frac{1}{2} \leq x \leq 0, \\ 0, & \text{elsewhere} \end{cases} \tag{1.2}$$

and support $[0, 1]$ is depicted in Fig. 1.1A. R. Schmidt (1934) possibly was the first to notice that the pdf (1.2) follows as the distribution of the arithmetic average of two uniform random variables U_1 and U_2 on $[0, 1]$, i.e.

$$X = \frac{U_1 + U_2}{2}. \qquad (1.3).$$

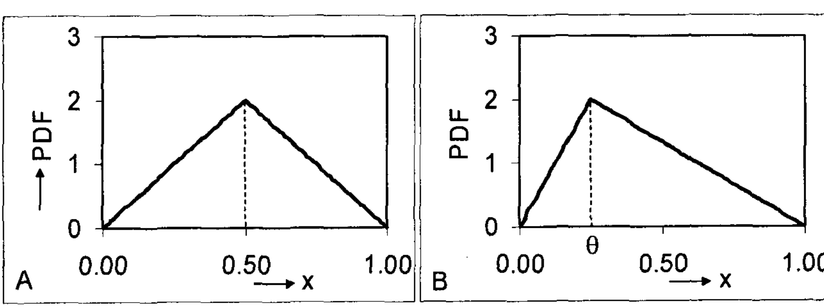

Fig. 1.1 A: Standard symmetric triangular distribution
B: Standard asymmetric triangular distribution with $\theta = 1/4$.

He referred to it as a *tine distribution* ("tine" is a slender projecting point). We were not able to find other Western sources dealing with triangular distributions between Simpson (1757) and Schmidt (1934) in the mainstream statistical literature. Asymmetric standard triangular distributions support $[0, 1]$ were studied by Ayyangar (1941). The pdf is given by

$$f(x|\theta) = \begin{cases} 2\frac{x}{\theta}, & \text{for } 0 \leq x \leq \theta, \\ 2\frac{1-x}{1-\theta}, & \text{for } \theta \leq x \leq 1, \\ 0, & \text{elsewhere.} \end{cases} \tag{1.4}$$

Substituting $\theta = 1/2$ yields the pdf (1.2). A standard asymmetric triangular distribution with $\theta = 1/4$ is depicted in Fig. 1.1B. The left ($\theta = 0$) and right ($\theta = 1$) triangular distribution (discussed in Rider (1963)) are depicted in Figs. 1.2A and 1.2B, respectively.

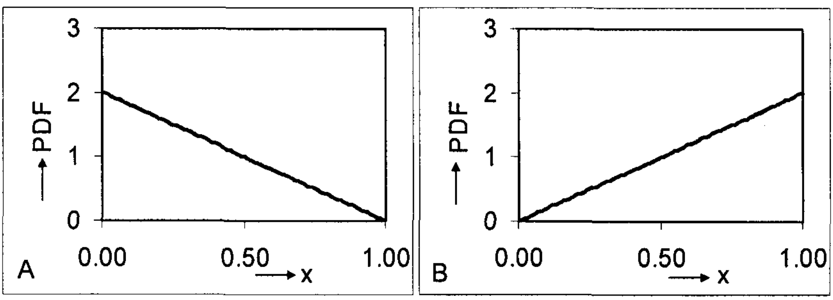

Fig. 1.2 A: Left triangular distribution ($\theta = 0$ in (1.4));
B: Right triangular distribution ($\theta = 1$ in (1.4)).

The left and right triangular distributions with support $[0, 1]$ are the only two members that the beta and triangular families have in common. Recall that the two parameter beta density is given by

$$f(x|\alpha, \beta) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, & \text{for } 0 \leq x \leq 1 \\ 0, & \text{elsewhere,} \end{cases} \tag{1.5}$$

where $\alpha > 0, \beta > 0$ and $\Gamma(\cdot)$ is the gamma function. Substituting $\alpha = 1$ and $\beta = 2$ in the beta pdf (1.5) yields the left triangular pdf ($\theta = 0$ in (1.4)). Substituting $\alpha = 2$ and $\beta = 1$ yields the right triangular one ($\theta = 1$ in (1.4)). Since 1941 up to the mid-sixties very few publications were devoted to the triangular distribution (Fullman (1953), Ostle *et al.* (1961) and Rider (1963)). The product of two identically independent distributed (i.i.d.) triangular random variables has been investigated by Donahue (1964). The sum of two independent triangular random variables (i.e., their convolution) sharing the same support (but not necessarily with the same mode) has – to the best of our knowledge – only been investigated very recently by Van Dorp and Kotz (2003b).

Since 1962 up to 1999, the distribution emerges in numerous papers dealing with the Project Evaluation and Review Technique – PERT (see, e.g., Clark (1962), Grubbs (1962), MacCrimmon and Ryaveck (1964), Moder and Rodgers (1968), Vãduva (1971), Williams (1992), Keefer and Verdini (1993), and D. Johnson (1997) amongst others). These papers deal with the asymmetric three-parameter triangular density

$$f(z|a,m,b)=\begin{cases}\frac{2}{b-a}\frac{z-a}{m-a}, & \text{for } a\le z\le m,\\ \frac{2}{b-a}\frac{b-z}{b-m}, & \text{for } m\le z\le b,\\ 0, & \text{elsewhere}\end{cases} \tag{1.6}$$

(with support $[a,b]$ and the mode m) which by means of the transformation $z=(x-a)/(b-a)$ reduces to its standard form (1.4) with the support $[0,1]$, where $\theta=(m-a)/(b-a)$. The parameters of the triangular distribution (1.6) are in one-to-one correspondence with a lower estimate $\widehat{a}$, a most likely estimate $\widehat{m}$, and an upper estimate $\widehat{b}$ of a characteristic under consideration. This leads to an intuitive appeal of the triangular distribution (see, e.g., Williams (1992)). In PERT these characteristics are the completion times of activities in a project network (see, Winston (1993)) whose uncertainties may be modeled by the distribution (1.6). N.L. Johnson and Kotz (1999) discuss the asymmetric triangular distribution in the context of YAWL distributions which have *inter alia* applications in modeling prices associated with orders placed by investors for single securities traded on the New York and American Stock Exchanges.

Recent popularity of the triangular distribution can be attributed to its use in Monte Carlo simulation modeling (see, e.g., Vose (1996) and Garvey (2002)), discrete system simulation (see, e.g., Banks *et al.* (2000), Altiok and Melamed (2001), Kelton *et al.* (2002)) and its use in standard uncertainty analysis software – such as @Risk (developed by the Palisade Corporation) or Crystal Ball (developed by Decision Engineering). These books and packages recommend the use of the triangular distribution when the underlying distribution is unknown, but a minimal value $\widehat{a}$, some maximal value $\widehat{b}$ and a most likely value $\widehat{m}$ are available. In Chapter 4, we shall discuss in some detail the appropriateness of this modeling approach given *only* these estimates.

1.2 Deriving the CDF utilizing a Geometric Argument

Instead of deriving the three-parameter cumulative distribution function (cdf) of the triangular distribution in the usual fashion from its pdf (1.6), we shall derive it using a geometric argument involving triangles (from which the triangular distribution derives its name). Figure 1.3A depicts the density function of a triangular distribution with parameters a, m and b, splitting the area underneath it into two triangles with area A_1 and A_2, respectively.

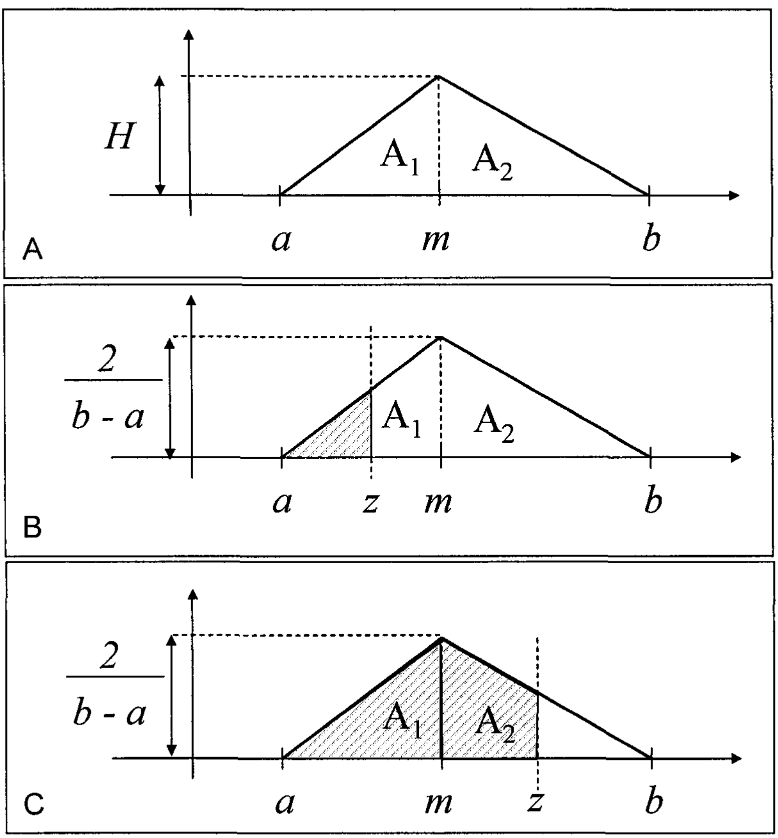

Fig. 1.3 Deriving of a triangular cdf utilizing areas of conforming triangles.

Since, from basic properties of a pdf it follows that $A_1 + A_2 = 1$ we have (see Fig. 1.3A)

$$(m-a)\frac{H}{2} + (b-m)\frac{H}{2} = 1 \Leftrightarrow H = \frac{2}{b-a}. \tag{1.7}$$

Hence, the density value at the mode m is not a function of the location of m, relative to the boundaries a and b (which is not obvious). Note that in Figs. 1.1 and 1.2 the density value at the mode equals 2 in all cases (since $a = 0$ and $b = 1$). In addition, from (1.7) we have

$$A_1 = \frac{m-a}{b-a} \text{ and } A_2 = \frac{b-m}{b-a}. \tag{1.8}$$

In other words, the probability mass to the left (the right) of the mode m, equals the *relative* distance of the mode m to the lower bound a (the upper bound b) compared to the whole range from a to b.

From Fig. 1.3B, Eq. (1.8) and utilizing conformity of the triangles, it immediately follows that for $a \le z \le m$:

$$Pr(Z \le z) = \left(\frac{z-a}{m-a}\right)^2 A_1 = \frac{m-a}{b-a}\left(\frac{z-a}{m-a}\right)^2 \tag{1.9}$$

and for $m \le z \le b$ using Fig. 1.3C and the complement rule $Pr(Z \le z) = 1 - Pr(Z > z)$:

$$Pr(Z \le z) = 1 - \left(\frac{b-z}{b-m}\right)^2 A_2 = 1 - \frac{b-m}{b-a}\left(\frac{b-z}{b-m}\right)^2 \tag{1.10}$$

Hence, the cdf is given by:

$$F(z) = Pr(Z \le z) = \begin{cases} \frac{m-a}{b-a}\left(\frac{z-a}{z-a}\right)^2, & \text{for } a \le z \le m, \\ 1 - \frac{b-m}{b-a}\left(\frac{b-z}{b-m}\right)^2 & \text{for } m \le z \le b. \end{cases} \tag{1.11}$$

Taking the derivative with respect to z in (1.11) we arrive at the pdf (1.6). The reader may wish to graph the cdf (1.11) for reasonable choices of a, m and b ($a \le m \le b$).

The inverse cdf of Z follows from (1.11) as

$$F^{-1}(y|a,m,b,n) = \tag{1.12}$$
$$\begin{cases} a+\sqrt{y(m-a)(b-a)}, & \text{for } 0 \le y \le \frac{m-a}{b-a} \\ b-\sqrt{(1-y)(b-m)(b-a)}, & \text{for } \frac{m-a}{b-a} \le y \le 1\,. \end{cases}$$

Equation (1.12) allows for straightforward sampling from a triangular distribution with support $[a,b]$ utilizing the inverse cdf transformation technique and a pseudo-random number generator of a uniform random variable on $[0,1]$ (see, e.g., Vose (1996)). Pseudo random number generators have become standard in spreadsheet software and are also utilized in uncertainty analysis packages such as @Risk (developed by the Palisade Corporation) and Crystal Ball (developed by Decision Engineering), and discrete event simulation software such as Arena (developed by Rockwell Software). The quality of the sample from a triangular distribution utilizing the inverse cdf transformation technique is identical to that obtained using the pseudo-random number generator. Banks *et al.* (2000) provide an excellent overview of desirable properties of and statistical tests for uniformity and independence of pseudo random number generators.

1.3 Moments of Triangular Distributions

The k-th moment about zero (which we shall denote by μ'_k) of a standard triangular distribution with support $[0,1]$ follows from the pdf (1.4) as

$$\mu'_k = E[X^k] = \int_0^1 x^k f(x|\theta)dx = \frac{2(1-\theta^{k+1})}{(k+1)(k+2)(1-\theta)}. \tag{1.13}$$

Here calculations are a bit lengthy but straightforward. The corresponding moments of a triangular variable Z with support $[a,b]$ and pdf (1.6) follow from (1.13) and the linear transformation $Z=(b-a)X+a$. Specifically,

$$E[Z^k] = E[\{(b-a)X+a\}^k] = \sum_{i=0}^{k}\binom{k}{i}(b-a)^i a^{k-1}E[X^i]. \tag{1.14}$$

Substituting $k=1$ and $k=2$ in (1.13) we arrive at the first and the second moments about zero of X:

$$\mu_1' = E[X] = \frac{\theta+1}{3}\ ; \mu_2' = E[X^2] = \frac{(1-\theta^3)}{6(1-\theta)} \tag{1.15}$$

and from the relation $\mu_2 = Var(X) = E[X^2] - E^2[X]$ we have

$$Var(X) = \mu_2 = \frac{1-\theta(1-\theta)}{18}. \tag{1.16}$$

Hence, the variance attains its minimum $3/72$ at $\theta = 1/2$ and its maximum $1/18$ at $\theta = 0$ or $\theta = 1$. Recall that the variance of a standard uniform distribution is much larger and equal to $1/12$.

In a similar manner, utilizing (1.15), substituting $k = 3$ and $k = 4$ in (1.13) and applying the definitions of the central moments

$$\begin{cases} \mu_3 = E[(X - E[X])]^3 = \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3 \\ \mu_4 = E[(X - E[X])]^4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - 3\mu_1'^4 \end{cases} \tag{1.17}$$

one obtains (see, Johnson and Kotz (1999)):

$$\begin{cases} \mu_3 = \frac{1}{270}(1-2\theta)(2-\theta)(1-\theta) \\ \mu_4 = \dfrac{1}{135}\{1-\theta(1-\theta)\}^2 \end{cases}$$

From the definitions of skewness $\sqrt{\beta_1}$ and kurtosis β_2 (see, e.g., Stuart and Ord (1994)) :

$$\sqrt{\beta_1} = Sign\{\mu_3\}\sqrt{\frac{\mu_3^2}{\mu_2^3}};\ \beta_2 = \frac{\mu_4}{\mu_2^2}, \tag{1.18}$$

(the skewness $\sqrt{\beta_1}$ retains the sign of the third central moment μ_3) we have

$$\sqrt{\beta_1} = \frac{(1-2\theta)(2-\theta)(1-\theta)}{\{1-\theta(1-\theta)\}^{\frac{3}{2}}}\frac{1}{5}\sqrt{2} \tag{1.19}$$

and

$$\beta_2 = 2.4. \tag{1.20}$$

(Compare with the kurtosis of a normal or Gaussian distribution, which equals 3.) The kurtosis β_2 (which is a combined measure of peakedness and heaviness of the tails of a distribution) here does not depend on θ.

Figure 1.4 plots skewness $\sqrt{\beta_1}$ as a function of θ. Observe that minimum skewness $-\frac{2}{5}\sqrt{2} \approx -0.566$ (which is a negative value) is attained for the right triangular distribution in Fig. 1.2B. It is important to note here that the left skewed distribution (with a heavier tail towards the left) has negative skewness and thus the designation *right* triangular distribution in Fig. 1.2B arises from the location of the mode θ being at the right boundary of the support. Similarly, the right skewed, *left* triangular distribution in Fig. 1.2A has the maximum positive skewness $\frac{2}{5}\sqrt{2} \approx 0.566$. The skewness of a symmetric triangular distributions $\sqrt{\beta_1} = 0$ is obtained from (1.19) by substituting $\theta = \frac{1}{2}$ (see Fig. 1.4).

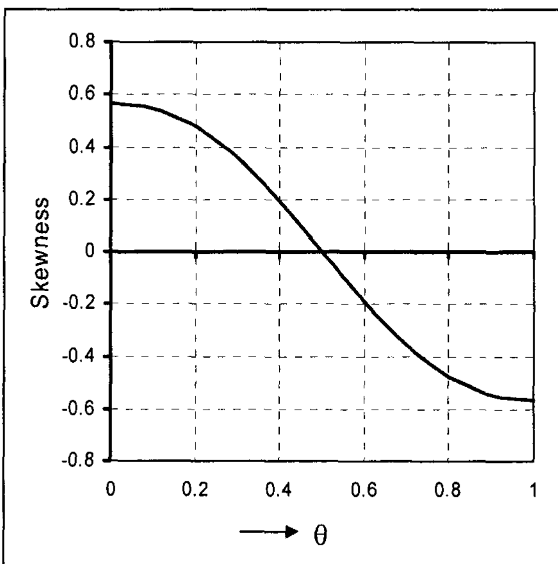

Fig. 1.4 Skewness $\sqrt{\beta_1}$ (Eq. (1.19)) as a function of θ.

Since the measures skewness and kurtosis are invariant under linear scale transformation it follows that (1.19) and (1.20), respectively, may be used for a triangular random variable Z with support $[a, b]$, pdf (1.6) and parameters a, m and b, utilizing $\theta = (m-a)/(b-a)$. From the linear transformation $Z = (b-a)X + a$, (1.15) and (1.16) we derive

$$E[Z] = \frac{a+m+b}{3} \tag{1.21}$$

and

$$Var[Z] = \frac{(b-a)^2}{18} \cdot \left\{1 - \frac{m-a}{b-a}\frac{b-m}{b-a}\right\}. \tag{1.22}$$

Note that from (1.21) it follows that the mean value of Z is the arithmetic average of the lower bound a, the mode m and the upper bound b. In our opinion, the popularity of the triangular distribution arises from the straightforward relationship (1.21) between the parameters and the mean of Z, a meaningful interpretation of the parameters a, m and b as well as from the property that the probability mass to the left of the mode m equals the relative distance of the mode m to the lower bound a over the whole support $[a, b]$ (i.e. $(m-a)/(b-a)$, see Eq. (1.8)).

1.4 Maximum Likelihood Method for the Threshold Parameter θ.

The structure of the standard triangular distribution (1.4) with support $[0, 1]$ leads to an illuminating procedure for estimating the threshold parameter θ. This parameter can be viewed as "dividing" (in the sense that it is related to two different analytical expressions appearing in the definition of the pdf (1.4)). The derivation of the ML estimator for θ in (1.4) seems to be quite instructive (and is similar, but simplified compared to the one presented in Johnson and Kotz (1999)).

Let for a random i.i.d. sample of size s, $\underline{X} = (X_1, \ldots, X_s)$, the order statistics be $X_{(1)} < X_{(2)} \ldots < X_{(s)}$. By definition, the likelihood for X with distribution (1.4) is

$$L(\underline{X}\,; \theta) = 2^s\{H(\underline{X}; \theta)\} \tag{1.23}$$

where

$$H(\underline{X};\theta) = \frac{\prod_{i=1}^{r} X_{(i)} \prod_{i=r+1}^{s} (1 - X_{(i)})}{\theta^r(1-\theta)^{s-r}} \tag{1.24}$$

and r is implicitly defined by $X_{(r)} \leq \theta < X_{(r+1)}$, $X_{(0)} \equiv 0$ and $X_{(s+1)} \equiv 1$.

Theorem 1.1: *Let* $\underline{X} = (X_1, \ldots, X_s)$ *be an i.i.d. sample from a triangular distribution with the pdf (1.4) and support* $[0, 1]$. *The ML estimator of* θ *maximizing the likelihood (1.23) over the parameter domain* $0 \leq \theta \leq 1$ *is*

$$\widehat{\theta} = X_{(\widehat{r})}, \tag{1.25}$$

where

$$\widehat{r} = \underset{r \in \{1, \ldots, s\}}{arg\,max} M(r) \tag{1.26}$$

and

$$M(r) = \prod_{i=1}^{r-1} \frac{X_{(i)}}{X_{(r)}} \prod_{i=r+1}^{s} \frac{1 - X_{(i)}}{1 - X_{(r)}}. \tag{1.27}$$

Proof: We shall provide a detailed proof of this basic theorem. (Another version of this theorem will be encountered in Chapter 5). To maximize the likelihood (1.23), we represent it as

$$\underset{0 \leq \theta \leq 1}{max} L(\underline{X};\theta) = 2^s \widehat{M}, \tag{1.28}$$

where

$$\widehat{M} = \underset{0 \leq \theta \leq 1}{max} H(\underline{X};\theta), \tag{1.29}$$

$H(\underline{X};\theta)$ is defined by (1.24) and $X_{(r)} \leq \theta \leq X_{(r+1)}$, with $X_{(0)} \equiv 0$, $X_{(s+1)} \equiv 1$. Utilizing (1.29) one can therefore write

$$\widehat{M} = \underset{r \in \{0, \ldots, s\}}{max} H(r), \tag{1.30}$$

where

$$H(r) = \max_{X_{(r)} \leq \theta \leq X_{(r+1)}} H(\underline{X}\,;\theta), \tag{1.31}$$

$r = 0, \ldots, s$, $X_{(0)} \equiv 0$ and $X_{(s+1)} \equiv 1$. The three non-overlapping cases: $r \in \{1, \ldots, s-1\}$, $r = 0$ and $r = s$ will be discussed separately.

Case $r \in \{1, \ldots, s-1\}$: Here, $X_{(r)} \leq \theta \leq X_{(r+1)}$. The function

$$g(\theta) = \theta^r (1-\theta)^{s-r} \tag{1.32}$$

in the denominator of the definition of $H(\underline{X}\,;\theta)$ (1.24) is proportional to an unimodal beta density since $r \in \{1, \ldots, s-1\}$. Thus,

$$\min_{X_{(r)} \leq \theta \leq X_{(r+1)}} g(\theta) = \min_{\theta \in \{X_{(r)}, X_{(r+1)}\}} g(\theta) \tag{1.33}$$

and, from (1.24), (1.31) and (1.33),

$$H(r) = \max_{r' \in \{r, r+1\}} \prod_{i=1}^{r'-1} \frac{X_{(i)}}{X_{(r')}} \prod_{i=r'+1}^{s} \frac{1-X_{(i)}}{1-X_{(r')}}. \tag{1.34}$$

Case $r = 0$: Here $0 \leq \theta \leq X_{(1)}$. From (1.24) and (1.31) it follows that now

$$H(0) = \max_{0 \leq \theta \leq X_{(1)}} \prod_{i=1}^{s} \frac{1-X_{(i)}}{1-\theta}.$$

Hence $H(0)$ becomes the product

$$H(0) = \prod_{i=1}^{s} \frac{1-X_{(i)}}{1-X_{(1)}} \equiv \prod_{i=2}^{s} \frac{1-X_{(i)}}{1-X_{(1)}}. \tag{1.35}$$

Case $r = s$: Here $X_{(s)} \leq \theta \leq 1$. From (1.24) and (1.31) it follows that in this case

$$H(s) = \max_{X_{(s)} \leq \theta \leq 1} \prod_{i=1}^{s} \frac{X_{(i)}}{\theta}.$$

Hence $H(s)$ becomes the product

$$H(s) = \prod_{i=1}^{s} \frac{X_{(i)}}{X_{(s)}} \equiv \prod_{i=1}^{s-1} \frac{X_{(i)}}{X_{(s)}}. \qquad (1.36)$$

(Compare with (1.35).) From (1.30), (1.34), (1.35) and (1.36) we obtain that

$$\widehat{M} = \max_{r \in \{1,\ldots,s\}} M(r), \qquad (1.37)$$

where $M(r)$ is defined by (1.27). Hence, the ML estimator of the threshold parameter θ equals the order statistic $X_{(\widehat{r})}$, where $\widehat{r}$ is given by (1.26). □

The ML estimator $\widehat{\theta} = X_{(\widehat{r})}$ given in (1.25) is quite intuitive (if one recalls the ML estimator $\widehat{\theta} = X_{(s)}$ of the parameter of a uniform distribution on $[0, \theta]$ for a sample of size s).

1.4.1 An illustrative example

We shall illustrate the ML estimation procedure for the parameter θ of a standard triangular distribution (1.4) by means of the following hypothetical order statistics

$$(X_{(1)}, \ldots, X_{(8)}) = (0.10, 0.25, 0.30, 0.40, 0.45, 0.60, 0.75, 0.80). \qquad (1.38)$$

This data was also used in Johnson and Kotz (1999)[3]. Consider the matrix $A = [a_{i,r}]$ with the entries :

$$a_{i,r} = \begin{cases} \frac{X_{(i)}}{X_{(r)}} & i < r \\ \frac{1-X_{(i)}}{1-X_{(r)}} & i \geq r. \end{cases} \qquad (1.39)$$

Table 1.1 summarizes calculations of the matrix A for the order statistics given in (1.38). The last row in the table contains the products of the matrix entries in the r-th column which are equal to the values of $M(r)$ given by

[3]Note: The values in Johnson and Kotz (1999) corresponding to the last three entries in the last row of Table 1.1 contain the following typos; 0.00547 should read 0.00543, 0.00137 should replace 0.00364 and 0.00029 should be 0.00290.

(1.27), $r = 1, \ldots, s$. Here $s = 8$. From the last row of Table 1.1 and utilizing (1.37), (1.26) and (1.25) we calculate

$$\widehat{M} = 0.011;\ \widehat{r} = 3; \\ \widehat{\theta} = X_{(\widehat{r})} = 0.30. \tag{1.40}$$

Table 1.1 ML estimation for a triangular distribution with pdf (1.4) using the data given by (1.38).

		r	1	2	3	4	5	6	7	8
			$X_{(1)}$	$X_{(2)}$	$X_{(3)}$	$X_{(4)}$	$X_{(5)}$	$X_{(6)}$	$X_{(7)}$	$X_{(8)}$
i			**0.10**	**0.25**	**0.30**	**0.40**	**0.45**	**0.60**	**0.75**	**0.80**
1	$X_{(1)}$	**0.10**	1	0.400	0.333	0.250	0.222	0.167	0.133	0.125
2	$X_{(2)}$	**0.25**	0.833	1	0.833	0.625	0.556	0.417	0.333	0.313
3	$X_{(3)}$	**0.30**	0.778	0.933	1	0.750	0.667	0.500	0.400	0.375
4	$X_{(4)}$	**0.40**	0.667	0.800	0.857	1	0.889	0.667	0.533	0.500
5	$X_{(5)}$	**0.45**	0.611	0.733	0.786	0.917	1	0.750	0.600	0.563
6	$X_{(6)}$	**0.60**	0.444	0.533	0.571	0.667	0.727	1	0.800	0.750
7	$X_{(7)}$	**0.75**	0.278	0.333	0.357	0.417	0.455	0.625	1	0.938
8	$X_{(8)}$	**0.80**	0.222	0.267	0.286	0.333	0.364	0.500	0.800	1
	M(r)		0.007	0.010	***0.011***	0.010	0.009	0.005	0.004	0.003

Figure 1.5 displays the function $H(\underline{X}\,;\theta)$ defined by Eq. (1.24) and shows that for the data in (1.38) the maximum value $\widehat{M} = 0.011$ of $H(\underline{X};\theta)$ over $\theta \in [0,1]$ is attained at $X_{(3)} = 0.30$. From (1.28), $\widehat{M} = 0.011$ and $s = 8$ we have $L(\underline{X};\widehat{\theta}) \approx 2.79$. Also observe that the maximum value $H(r)$ (see, Eq. (1.31)) of $H(\underline{X}\,;\theta)$ over $\theta \in [X_{(r)},\ X_{(r+1)}]$ is attained at either $X_{(r)}$ or $X_{(r+1)}$ for all $r = 0, \ldots, s$.

The ML estimation of the mode θ of the triangular pdf (1.4) with support $[0,1]$ can easily be modified to the ML estimation of the mode m of the triangular pdf (1.6) with support $[a,b]$, using the linear scale transformation $Z = (b-a)X + a$; recall that the parameters a and b are fixed and the parameter $m = (b-a)\theta + a$. The ML estimator of the parameter m of the distribution (1.6) utilizing the order statistics $(Z_{(1)}, \ldots, Z_{(s)})$ are

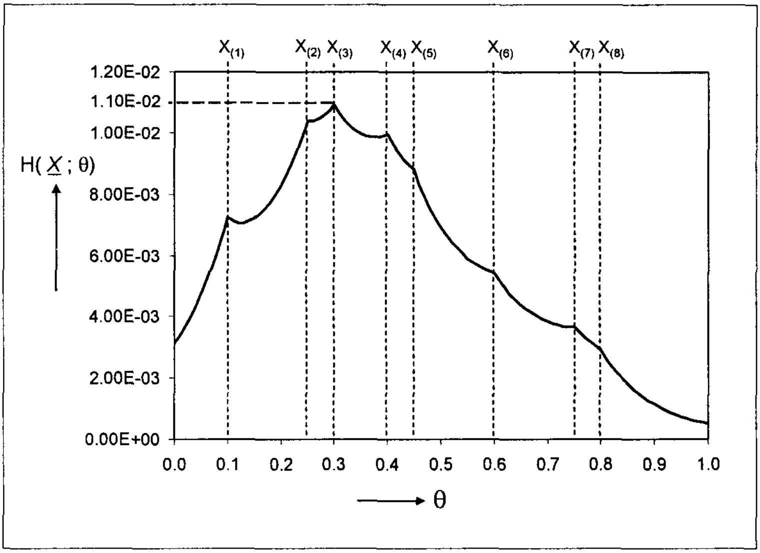

Fig. 1.5 Graph of $H(X\ ;\theta)$ (1.24) for the data in (1.38). (Observe that maxima over the sets $[X_{(r)}, X_{(r+1)}]$ are attained here solely at the order statistics, $r = 1, \ldots, s$).

$$\widehat{m}(a,b) = Z_{(\widehat{r}(a,b))} \tag{1.41}$$

where, as above,

$$\widehat{r}(a,b) = \underset{r \in \{1,\ldots,s\}}{arg\ max} M(a,b,r) \tag{1.42}$$

and

$$M(a,b,r) = \prod_{i=1}^{r-1} \frac{Z_{(i)} - a}{Z_{(r)} - a} \prod_{i=r+1}^{s} \frac{b - Z_{(i)}}{b - Z_{(r)}}. \tag{1.43}$$

Compare with equations (1.25), (1.26) and (1.27).

1.5 Three Parameter Maximum Likelihood Estimation

This lengthy section involves some non-standard interesting derivations of the ML procedure of the three-parameter triangular distributions which are closely related to a non-regular case of ML estimation for continuous distributions (see, e.g., Cheng and Amin (1983)).

Let Z be a random variable with pdf (1.6). For a random sample $\underline{Z} = (Z_1, \ldots, Z_s)$ with size s from a triangular distribution with support $[a, b]$ and mode m, let the order statistics be $Z_{(1)} < Z_{(2)} < \ldots < Z_{(s)}$. Utilizing (1.6), the likelihood for $\underline{Z}$ is by definition

$$L(\underline{Z}; a, m, b) = \left(\frac{2}{b-a}\right)^s \left\{ \prod_{i=1}^{r} \frac{Z_{(i)} - a}{m - a} \prod_{i=r+1}^{s} \frac{b - Z_{(i)}}{b - m} \right\} \quad (1.44)$$

where r is implicitly defined by $Z_{(r)} \leq m < Z_{(r+1)}$, $Z_{(0)} \equiv a$ and $Z_{(s+1)} \equiv b$. Thus, analogously to (1.28) it follows that for fixed values of a and b, satisfying

$$a < Z_{(1)} \text{ and } b > Z_{(s)},$$

we have

$$\max_{a \leq m \leq b} L(\underline{Z}; a, m, b) = \left(\frac{2}{b-a}\right)^s \left\{ M(a, b, \widehat{r}(a, b)) \right\} \quad (1.45)$$

where $\widehat{r}(a, b)$ and $M(a, b, r)$ are given by (1.42) and (1.43), respectively. The ML estimator for the mode m (as a function of a and b) is given by Eq. (1.41). Note that, the function $\widehat{r}(a, b)$ is an *index* function indicating at which order statistic the ML estimate of the parameter m is attained as a function of the lower bound a and upper bound b (we shall elaborate on the index function $\widehat{r}(a, b)$ below).

From (1.45) we have that

$$\max_{S(a,m,b)} \left[Log\{L(\underline{Z}; a, m, b)\} \right] = \quad (1.46)$$
$$\max_{a < X_{(1)}, b > X_{(s)}} \left[sLog2 + G(a, b) \right],$$

where the set

$$S(a,m,b) = \{(a,m,b)|\, a < Z_{(1)}, b > Z_{(s)}, a \le m \le b\} \tag{1.47}$$

and the function

$$G(a,b) = Log\{M(a,b,\widehat{r}(a,b)\} - sLog\{b-a\}. \tag{1.48}$$

This is an interesting function to be discussed below. Again recall the definitions of $M(a,b,r)$ and that of $\widehat{r}(a,b)$ in Eqs. (1.42) and (1.43). Note that $G(a,b)$ is only defined for values of $a < Z_{(1)}$ and $b > Z_{(s)}$ (see Eq. (1.47)). To summarize, the three-dimensional optimization problem of maximizing the likelihood (1.44) reduces to a two-dimensional case of maximizing $G(a,b)$ over the region $a < Z_{(1)}$ and $b > Z_{(s)}$. From the structure of (1.44), however, we can immediately conclude that for all values of m such that

$$Z_{(1)} < m < Z_{(s)} \tag{1.49}$$

the likelihood $L(\underline{Z};a,m,b) \to 0$ (and hence $Log\{L(\underline{Z};a,m,b)\} \to -\infty$) when $a \uparrow Z_{(1)}$ or $b \downarrow Z_{(s)}$. Thus, when a modal value can be observed in the data (via, for example, a histogram) indicating the validity of Eq. (1.49), it would seem that the ML estimators for a and b are *not* the order statistics $Z_{(1)}$ and $Z_{(s)}$, respectively. This is in contrast with the well-known fact that the ML estimators of a uniform distribution with support $[a,b]$ *are* given by smallest order statistic $X_{(1)}$ and the largest one $X_{(s)}$ (see, e.g., Devore (2004)).

We shall demonstrate the above fitting characteristic of a triangular distribution for civil engineering data consisting of a sample of 85 hauling times (Source, AbouRizk (1990)) rather than the hypothetical 8 point example given by (1.38) since we are now fitting a three parametric distribution instead of a distribution with one parameter θ given by (1.4). Figure 1.6 depicts the empirical pdf for the data in Table 1.2 which seems to have a mode in the vicinity of the center of the range $[Z_{(1)}, Z_{(85)}] = [3.20, 8.60]$. Hence, Eq. (1.49) is satisfied. In addition, Fig. 1.6 depicts the ML fitted triangular distribution with ML estimates of parameters

Table 1.2 Civil engineering data consisting of 85 hailing times (Source: AbouRizk (1990)).

4.79	4.75	5.40	4.70	6.50
5.30	6.00	5.90	4.80	6.70
6.00	4.95	7.90	5.40	3.50
4.54	6.90	5.80	5.40	5.70
8.00	5.40	5.60	7.50	7.00
4.60	3.20	3.90	5.90	3.40
5.20	5.90	4.40	5.20	7.40
5.70	6.00	3.60	6.20	5.70
5.80	5.90	6.00	5.15	6.00
4.82	5.90	6.00	7.30	7.10
4.73	5.90	3.60	6.30	7.00
5.10	6.00	6.60	4.40	6.80
5.60	5.90	5.90	8.60	6.00
5.80	5.40	6.50	4.80	6.40
4.15	4.90	6.50	8.20	7.00
8.50	5.90	4.40	5.80	4.30
5.10	5.90	4.70	3.50	6.80

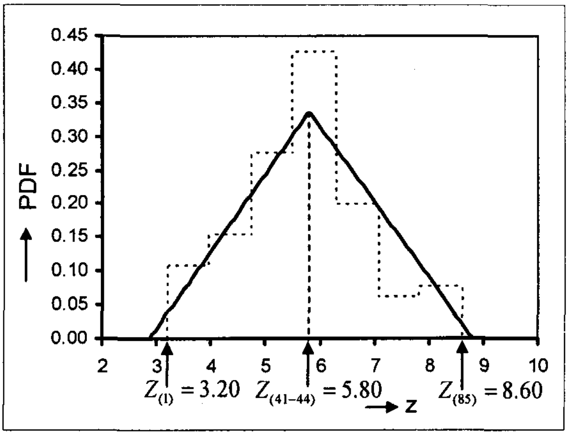

Fig. 1.6 Empirical pdf for the data in Table 1.2 together with a ML fitted three-parameter triangular distribution $\widehat{a} = 2.87$, $\widehat{m} = Z_{(41-44)} = 5.80, \widehat{b} = 8.80$.

$$\widehat{a} = 2.87 < Z_{(1)} = 3.20,\ \widehat{m} = Z_{(41-44)} = 5.80, \qquad (1.50)$$
$$\widehat{b} = 8.80 > Z_{(s)} = 8.60.$$

Observe that the example data in Table 1.2 actually contains ties, resulting in the ML estimator $\widehat{m}$ to be attained at either one of the order statistics $Z_{(41)}$ through $Z_{(44)}$. Also note that the triangular distribution in Fig. 1.6 does not quite capture the 'peak' of the empirical pdf in Fig. 1.6. In Chapter 4 we shall fit a four parameter generalization of the triangular distribution that does capture this 'peak' and present a more formal fit analysis using the chi-square test (see, e.g., Devore (2004)).

Figure 1.7 provides the form of the function $G(a, b)$ given by (1.48) that was maximized to arrive at the ML estimators for the lower and upper bounds a and b in (1.50) for the data in Table 1.2. Figure 1.8A (Figure 1.8B) depicts a likelihood profile of the function $G(a, b)$ displayed in Fig. 1.7 for the data in Table 1.2 as function of the parameter a (parameter b) for different fixed values of the parameter b (parameter a). Note the behavior of $G(a, b)$ for $b = 8.6$ (for $a = 3.2$) in Fig. 1.8A (Fig. 1.8B). The ML estimates $\widehat{a} = 2.87$ and $\widehat{b} = 8.80$ are indicated by means of a vertical solid line in Figs. 1.8A and 1.8B, respectively. Observe the apparent mirror symmetry of the graphs in Figs. 1.8A and 1.8B for the data in Table 1.2. A further investigation of the function would be appropriate (see Sec. 1.5.1). Moreover, note that the profile log-likelihood of the function $G(a, b)$ in Fig. 1.8A (Fig. 1.8B) for the value of the largest order statistic $Z_{(s)} = 8.6$ (smallest order statistic $Z_{(1)} = 3.2$) is located below the other two, which indicates that $Z_{(s)}$ (that $Z_{(1)}$) is *not* the ML estimator for the upper bound b (lower bound a).

Readers interested in more statistical aspects of the three-parameter triangular distribution may omit Secs. 1.5.1 and 1.5.2 (with its subsections) during an initial reading.

1.5.1 Some details about the functions $G(a, b)$ and $\widehat{r}(a, b)$

While the function $G(a, b)$ given by (1.48) is continuous over its domain $a < Z_{(1)}$ and $b > Z_{(s)}$, the partial derivatives with respect to a or b may not be unique at a finite $(s - 1)$ number of points. The source of non-differentiability at these points is due to the behavior of the index function

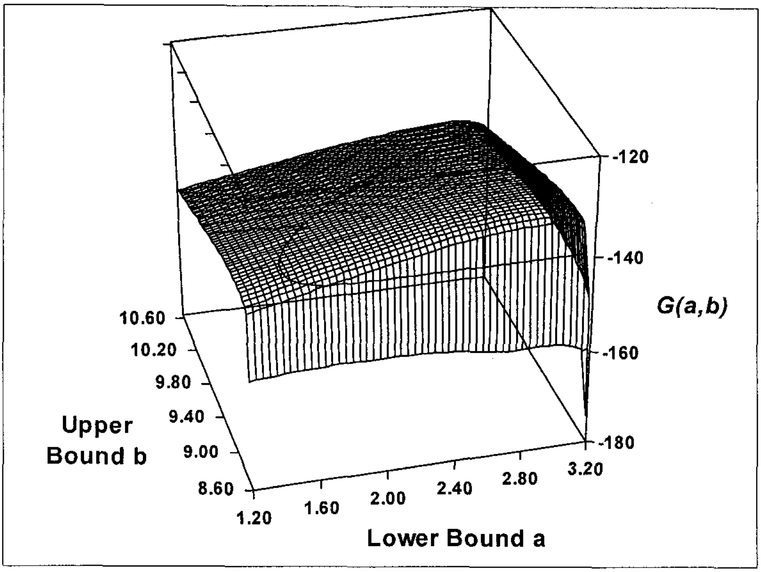

Fig. 1.7 The function $G(a,b)$ given by (1.48) for the data in Table 1.2.

$\widehat{r}(a,b)$ given by (1.42) as a function of the parameters a and b. In fact, the following properties can be derived for $\widehat{r}(a,b)$ as a function of b, keeping $a < X_{(1)}$ fixed (recall that $\widehat{r}(a,b)$ is an *index* function indicating at which order statistic the ML estimate of the parameter m is attained);

(1) The order statistic index $\widehat{r}(a,b)$ is decreasing in b;

(2) $\lim_{b \to \infty} \widehat{r}(a,b) = 1$;

(3) $\lim_{b \downarrow X_{(s)}} \widehat{r}(a,b) = s$;

(4) $\widehat{r}(a,b)$ as a function of b has $(s-1)$ discontinuities at the points

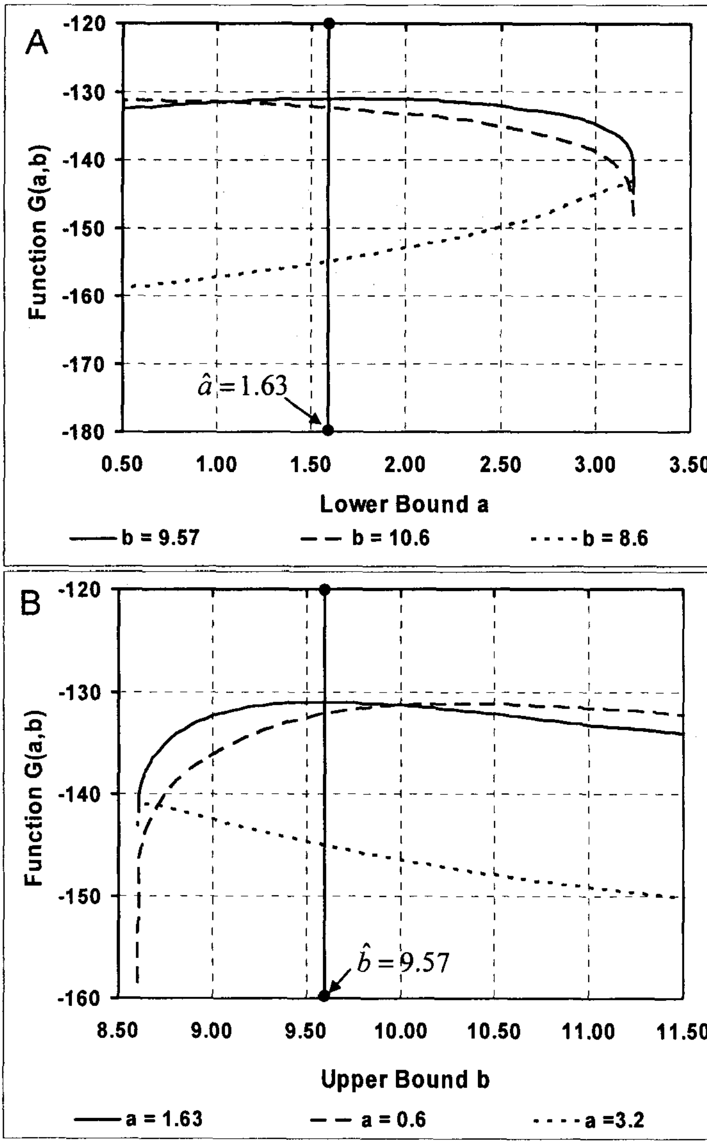

Fig. 1.8 Profiles of the function $G(a, b)$ given by (1.48) for the data in Table 1.2: Graph A: as a function of the lower bound a; Graph B: as a function of the upper bound b. The ML estimates $\widehat{a} = 2.87; \widehat{b} = 8.80$ in (1.50) are indicated by means of a vertical dotted line in Figs. 1.8A and 1.8B, respectively.

$$f_b(a,r) = \frac{X_{(r+1)} - X_{(r)} \sqrt[s-r]{\left(\frac{X_{(r)}-a}{X_{(r+1)}-a}\right)^r}}{1 - \sqrt[s-r]{\left(\frac{X_{(r)}-a}{X_{(r+1)}-a}\right)^r}}, \quad r \in \{1, \ldots, s-1\}. \tag{1.51}$$

(Note that the parameter a is fixed.) Similar properties can be derived for $\widehat{r}(a,b)$ as a function of a, while keeping $b > X_{(s)}$ fixed. Figure 1.9 gives the form of the function $\widehat{r}(a,b)$ (Eq. (1.42)) for the data in (1.38). The function $\widehat{r}(a,b)$ may be viewed as a bivariate step-function or a *winding staircase function*, which could serve as a useful tool for studying non-differentiable bivariate distributions. We are purposely using only a the small set of 8 data points in (1.38) in Fig. 1.9 to emphasize the stepwise behavior of the function $\widehat{r}(a,b)$, which would have been less apparent visually when using, for example, the whole data set in Table 1.2. The central axis of the "winding staircase" in Fig. 1.9 is located at $a = Z_{(1)} = 0.10$ and $b = Z_{(s)} = 0.80$. For a fixed a, the value of $f_b(a,r)$ (1.51) identifies the location of the r-th step (in terms of b) of the winding stare case. Note that at the central axis ($a = X_{(1)}$, $b = X_{(s)}$), the $(s-1)$ discontinuities $f_b(a,r)$ of the index function $\widehat{r}(a,b)$ converge.

Discarding the points of discontinuity of the function $\widehat{r}(a,b)$, the function $G(a,b)$ becomes differentiable with respect to a and b. From (1.48) we obtain:

$$\frac{\partial}{\partial a}G(a,b) = \frac{\frac{\partial}{\partial a}M(a,b,\widehat{r}(a,b))}{M(a,b,\widehat{r}(a,b))} + \frac{s}{b-a} \tag{1.52}$$

and

$$\frac{\partial}{\partial b}G(a,b) = \frac{\frac{\partial}{\partial b}M(a,b,\widehat{r}(a,b))}{M(a,b,\widehat{r}(a,b))} - \frac{s}{b-a}, \tag{1.53}$$

where the partial derivatives of $M(a,b,r)$ (1.43) are

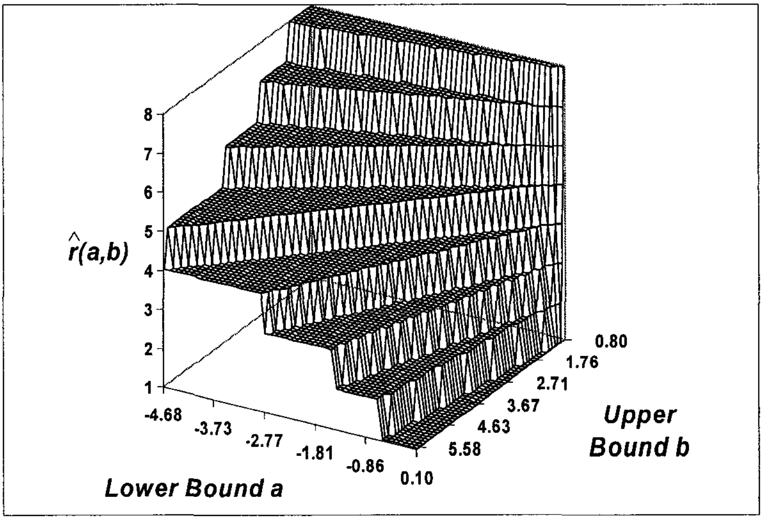

Fig. 1.9 The index function $\widehat{r}(a,b)$ given by Eq. (1.42) for the data in (1.38).

$$\frac{\partial}{\partial a}M(a,b,\widehat{r}(a,b)) = M(a,b,\widehat{r}(a,b)) \times \left\{ \sum_{j=1}^{\widehat{r}-1} \frac{Z_{(j)} - Z_{(\widehat{r})}}{(Z_{(\widehat{r})} - a)(Z_{(j)} - a)} \right\} < 0 \tag{1.54}$$

and

$$\frac{\partial}{\partial b}M(a,b,\widehat{r}(a,b)) = M(a,b,\widehat{r}(a,b)) \times \left\{ \sum_{j=\widehat{r}+1}^{s} \frac{Z_{(j)} - Z_{(\widehat{r})}}{(b - Z_{(\widehat{r})})(b - Z_{(j)})} \right\} > 0. \tag{1.55}$$

A routine $BSearch$ has been developed utilizing (1.51), (1.53) and (1.55) to determine $\widetilde{b}\ (a)$ for fixed a, where

$$\widetilde{b}\ (a) = arg \max_{b > Z_{(s)}} \Big[G(a,b) \Big]. \tag{1.56}$$

This routine follows a bisection approach (see, e.g., Press *et al.* (1989)) and is described in the next subsection. Having the routine $BSearch$ to determine $\widetilde{b}\ (a)$ for fixed a, we next compile a routine $ABSearch$ which determines $\widehat{a}$ and $\widetilde{b}\ (\widehat{a})$ such that

$$\widehat{a} = \underset{a < Z_{(1)}}{argmax} \left[G(a, \widetilde{b}\ (a)) \right].$$

The latter routine utilizes (1.52), (1.54) and is also based on a bisection approach. (It is described in the next subsection.) The routine $ABSearch$ evaluates the maximum of the likelihood, namely the RHS of (1.46), by successively utilizing $BSearch$ and yields the following ML estimators :

$$\widehat{a},\ \widehat{b} = \widetilde{b}\ (\ \widehat{a}\ ),\ \widehat{m}(\widehat{a},\ \widehat{b}) = Z_{(\widehat{r}(\widehat{a},\ \widehat{b}))} \text{ and}$$
$$\widehat{n}(\widehat{a}, \widehat{b}) = -\frac{s}{Log\{M(\widehat{a}, \widehat{b}, \widehat{r}(\widehat{a}, \widehat{b})\}}$$

where $\widetilde{b}\ (\cdot)$ and $\widehat{r}(a, b)$ are defined in (1.56) and (1.42), respectively. For ease of implementation, the ML procedure above is summarized in Pseudo Pascal in the next subsection. We emphasize that the procedure – although straightforward – requires utilization of a number of variables and careful analysis of the consecutive steps and their interconnection.

1.5.2 ML estimation procedure in pseudo Pascal

The numerical routines below in Pseudo Pascal require separate algorithms to evaluate:

$M(a_k, b_k, r_k)$: Eq. (1.43), $G(a_k, b_k, r_k)$: Eq. (1.48),

$\frac{\partial}{\partial a} G(a_k, b_k, r_k)$: Eq. (1.52), $\frac{\partial}{\partial b} G(a_k, b_k, r_k)$: Eq. (1.53),

$\frac{\partial}{\partial a} M(a_k, b_k, r_k)$: Eq. (1.54) and $\frac{\partial}{\partial b} M(a_k, b_k, r_k)$: Eq. (1.55).

Output parameters of routines below are indicated in bold.

1.5.2.1 The search routine Bsearch

Let $G(a, b)$ be the function defined by (1.48). For a given value of the parameter a the set of discontinuities in the parameter b of the function $G(a, b)$ is a (finite) null-set and one could thus utilize the partial derivatives with respect to b (1.53) and (1.55) to determine an ascending search direction with respect to $G(a, b)$ for b. Define

$$B(a) = \underset{r \in \{1, \ldots, s-1\}}{Max} \Big[f_b(a, r) \Big], \tag{1.57}$$

where $f_b(a, r)$ are the discontinuity points given by (1.51). From the properties of $\widehat{r}(a, b)$ (1.42) mentioned at the beginning of Sec. 1.5.1, it follows that for $b > B(a)$ (*outside* the discontinuity locations) :

$$G(a, b) = Log\left\{ \prod_{i=2}^{s} \frac{b - Z_{(i)}}{b - Z_{(1)}} \right\} - sLog\{b - a\}$$

and

$$\frac{\partial}{\partial a} G(a, b) = \frac{s}{b - a} > 0. \tag{1.58}$$

Compare with the derivative (1.52). Hence, it follows from (1.58) that necessary conditions for a local maximum of $G(a, b)$ (i.e. $\frac{\partial}{\partial a} G(a, b) = 0$ and $\frac{\partial}{\partial b} G(a, b) = 0$) cannot be satisfied for $b > B(a)$. Thus, $BSearch$ maximizing $G(a, b)$ as a function of b with a fixed, can be confined to the interval $(Z_{(s)}, B(a))$ only. The routine $BSearch$ below evaluates $\widetilde{b}\ (a)$ (1.56), follows a bisection approach (see, e.g., Press *et al.* (1989)) and requires a separate algorithm to evaluate $B(a)$ (1.57).

$BSearch(a_k, \underline{Z}, \boldsymbol{b_k}, \boldsymbol{M_k}, \boldsymbol{r_k})$

$Step\,1:\quad l_k^b = Z_{(s)}$

$Step\,2:\quad u_k^b = B(a_k), b_k = \frac{l_k^b + u_k^b}{2},$

$\qquad M_k = M(a_k, b_k, r_k), G_k = \frac{\partial}{\partial b} G(a_k, b_k, M_k, r_k)$

$Step\,3:\quad If\ Abs(G_k) \geq \delta\ \ then$

$\qquad\qquad If\ G_k < 0\ then\ u_k^b :\ = b_k\ \ Else\ l_k^b :\ = b_k$

$\qquad Else\ Stop$

$$Step\,4: \qquad If\ (u_k^b - l_k^b) \geq \delta\ Goto\ Step\,2\ \ Else\ Stop$$

1.5.2.2 The search routine ABSearch

Let as above $G(a,b)$ be the function defined by (1.48). For a given value of the parameter b the set of discontinuities in the parameter a of the function $G(a,b)$ is a null-set and one could thus utilize the partial derivatives with respect to a (Eqs. (1.52) and (1.54)) to determine an ascending search direction with respect to $G(a,b)$ for a. The routine $ABSearch$ starts by establishing an interval $[A, X_{(1)}]$ such that

$$\frac{\partial}{\partial a} G(A, \widehat{b}(A)) > 0, \tag{1.59}$$

where $\widehat{b}(A)$ maximizes $G(A,b)$ as a function of b (and is calculated using the $BSearch$ routine in Sec. 1.5.1.1). To determine A in (1.59) one may utilize (1.52) and (1.54). From

$$\lim_{a \to -\infty} \left[\widehat{r}(a,b)\right] = s,$$

it follows that for any given b, there is a sufficiently small a such that

$$G(a,b) = Log\left\{\prod_{i=1}^{s-1} \frac{Z_{(i)} - a}{Z_{(s)} - a}\right\} - sLog\{b-a\}$$

and

$$\frac{\partial}{\partial a} G(a,b) = \sum_{i=1}^{s-1} \frac{Z_{(i)} - Z_{(s)}}{(Z_{(i)} - a)(Z_{(s)} - a)} + \frac{s}{b-a}. \tag{1.60}$$

It thus follows from (1.60) that for any given b there exists an a sufficiently small such that $\frac{\partial}{\partial a} G(a,b) > 0$. So far we can only conjecture that an A satisfying (1.59) does exist. Numerical analyses support this conjecture. Having established the search interval $[A, X_{(1)}]$, the routine $ABSearch$ follows (analogously to the routine $BSearch$) a bisection approach (see, e.g., Press *et al.* (1989)) and evaluates the RHS of (1.46) by successively utilizing the routine $BSearch$.

$$
\begin{array}{lll}
ABSearch(\underline{Z}\,,\, \boldsymbol{a_k},\, \boldsymbol{b_k}, \boldsymbol{m_k}) & & \\
Step\,1 & : & u_k^a = Z_{(1)},\, l_k^a = Z_{(1)} - (Z_{(s)} - Z_{(1)}) \\
Step\,2 & : & BSearch(l_k^a, \underline{Z}\,,\, \boldsymbol{b_k}, \boldsymbol{M_k}, \boldsymbol{r_k}), \\
 & & G_k = \frac{\partial}{\partial a} G(l_k^a,\, b_k, M_k, r_k) \\
Step\,3 & : & If\, G_k < 0\, then \\
 & & \qquad u_k^a = l_k^a, l_k^a = l_k^a - (Z_{(s)} - Z_{(1)}), Goto\, Step\,2. \\
Step\,4 & : & a_k = \frac{l_k^a + u_k^a}{2}, \\
 & & BSearch(\, a_k, \underline{Z}\,,\, \boldsymbol{b_k}, \boldsymbol{M_k}, \boldsymbol{r_k})\,, \\
 & & G_k = \frac{\partial}{\partial a} G(\, a_k, b_k, M_k, r_k) \\
Step\,5 & : & If\, Abs(G_k) \geq \delta\; then \\
 & & \qquad If\, G_k < 0\, then\, u_k^b := a_k\, Else\, l_k^b := a_k \\
 & & Else\, Goto\, Step\,7 \\
Step\,6 & : & If\, (u_k^a - l_k^a) \geq \delta\, then\, Goto\, Step\,4 \\
 & & Else\, Goto\, Step\,7. \\
Step\,7 & : & m_k = Z_{(r_k)}
\end{array}
$$

1.6 Solving for a and b using a Lower and Upper Quantile Estimate

We shall conclude our discussion of the triangular distribution by providing an appealing and smooth method of using quantile estimates to solve for a and b. Let Z be a triangular pdf with support $[a, b]$ and mode m with the pdf (1.6) and the cdf (1.11). As mentioned above, the recent popularity of the triangular distribution could perhaps be attributed to its use in uncertainty analysis packages such as @Risk (developed by the Palisade corporation). The package @Risk allows definition of a triangular distribution (via the function TRIGEN) by specifying a lower quantile a_p, a most likely value m and an upper quantile b_r such that

$$a < a_p \leq m \leq b_r < b. \tag{1.61}$$

The latter avoids having to specify the lower and upper extremes a and b that by definition have a zero likelihood of occurrence (since, the triangular density equals zero at the bounds a and b). The software @Risk does not provide details, however, regarding how the bounds a and b are calculated given values for a_p, m and b_r. Keefer and Bodily (1983) formulated this problem in terms of two quadratic equations from which the unknowns a and b had to be solved numerically for the values $p = 0.05$ and $r = 0.95$. Although the numerical solution of their equations and their generalizations to other values of p and r do not pose any difficulties, we shall present here a slightly simplified version that only requires to solve numerically a single equation in the unknown quantity

$$q = \frac{m-a}{b-a}. \tag{1.62}$$

It follows from the cdf (1.11) that the quantity q equals the probability mass to the left of the mode m (and also equals the relative distance of the mode m to the lower bound a over the whole support $[a, b]$, which is unknown here).

From the definition of a_p, $(F(a_p|a, m, b) = p)$, we have from (1.11) and (1.62) that

$$a_p = a + (m-a)\sqrt{\frac{p}{q}}. \tag{1.63}$$

There is no direct relation between p and q here (contrary to the common notation when dealing with proportions and/or the binomial distribution), except that from (1.61) and (1.62) it follows that $0 < p < q < 1$. Solving for the parameter a from (1.63), yields using (1.62)

$$a \equiv a(q) = \frac{a_p - m\sqrt{\frac{p}{q}}}{1-\sqrt{\frac{p}{q}}} < \frac{a_p - a_p\sqrt{\frac{p}{q}}}{1-\sqrt{\frac{p}{q}}} = a_p. \tag{1.64}$$

(We use the notation $a(q)$ instead of a to emphasize that the lower bound a is a function of q, provided the p-th percentile a_p and the most likely value m are given.) Analogously to (1.64), we have for $m < b_r$ (using $b(q)$ in place of b):

$$b \equiv b(q) = \frac{b_r - m\sqrt{\frac{1-r}{1-q}}}{1-\sqrt{\frac{1-r}{1-q}}} > \frac{b_r - b_r\sqrt{\frac{1-r}{1-q}}}{1-\sqrt{\frac{1-r}{1-q}}} = b_r. \tag{1.65}$$

(Here we have from (1.61) and (1.62) that $1-q > 1-r > 0$).

Substituting $a(q)$ and $b(q)$ as given by (1.64) and (1.65) into (1.62), we arrive at the following *basic equation*

$$q = g(q) \tag{1.66}$$

where

$$g(q) = \frac{m - a(q)}{b(q) - a(q)} = \tag{1.67}$$

$$\frac{(m - a_p)\left(1 - \sqrt{\frac{1-r}{1-q}}\right)}{(b_r - m)\left(1 - \sqrt{\frac{p}{q}}\right) + (m - a_p)\left(1 - \sqrt{\frac{1-r}{1-q}}\right)}.$$

Observe the rather "structured" relation between $g(q)$ and q. Indeed, from its structure it immediately follows that $0 \le g(q) \le 1$ (as it should be since $g(q)$ represents the probability mass to the left of the mode m). In fact, setting $q = p$ $(q = r)$ in the RHS of (1.67) yields $g(p) = 1$ $(g(r) = 0)$. In addition, the denominator of the RHS (1.67) is "almost" a linear combination of the distances of the quantiles a_p and b_r from the mode m, with the weights that are determined by the quantile probability masses p and r and the probability mass q to the left of the mode m. In Chapter 4 (Sec. 4.3.3.3) we shall show that a *generalized* version of the Eq. (1.67) has a unique solution $q^* \in [p, r]$. One can solve numerically for q^* utilizing (1.66), the definition of $g(q)$ (1.67) and our favorite bisection method (see, e.g., Press *et al.* (1989)) with the starting interval $[p, r]$. After solving for the unique solution q^* of Eq. (1.66) one could calculate the associated lower and upper bounds $a(q^*)$ and $b(q^*)$ from Eqs. (1.64) and (1.65), respectively.

We shall illustrate the above procedure via the example:

$$a_p = 6.5, m = 7, \text{ and } b_r = 10.5, p = 0.10 \text{ and } r = 0.90. \tag{1.68}$$

Figure 1.10 depicts the function $g(q)$ (1.67) for the example above. Note that as stated above $g(p) = 1$, $g(r) = 0$ in this case and that the unique solution $q^* = 0.2198$ is the intersection of the function $g(q)$ with the positive diagonal of the unit square (indicated by a dotted line in Fig. 1.10). We calculated the value of q^* using the standard root finding algorithm GOALSEEK available in Microsoft Excel. Next, from (1.64) and (1.65) and utilizing $q^* = 0.2198$, we obtain for the lower and upper bounds

$$a(q^*) = 5.464 \text{ and } b(q^*) = 12.452,$$

respectively.

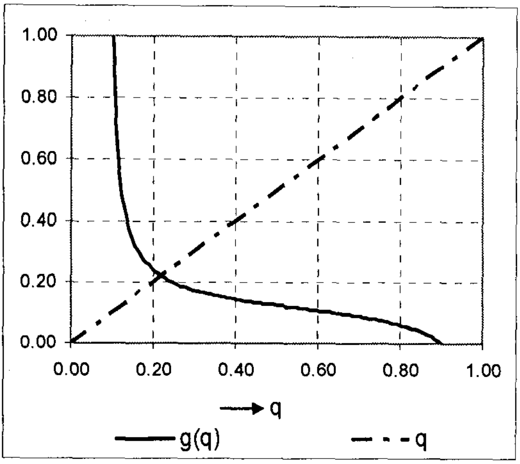

Fig. 1.10 The function $g(q)$ given by (1.67) with $a_p = 6.5, m = 7,$ and $b_r = 10.5, p = 0.10$ and $r = 0.90$.

1.7 Concluding Remarks

We have presented some details and properties of the triangular distribution which possibly have not been sufficiently addressed in the statistical literature. For example, to the best of our knowledge, the three-parameter ML method for the triangular distribution was first presented in Van Dorp and Kotz (2002b). The software BESTFIT developed by the Palisade corporation (which has been already available for a number of years now), however, does yield exactly the same estimates for the parameters a, m and b for the data in Table 1.2. Unfortunately, the authors do not provide specific details on their method for obtaining these estimates. On the other hand, another fitting software package called INPUT ANALYZER (developed by Rockwell Software) does *not* yield the same parameter estimates for the data in Table 1.2. (Again, no details are provided about the estimation procedure.)

A careful reader would have noticed that the method of moments for the standard triangular distribution (1.4) with support $[0, 1]$ has not been explicitly discussed in this chapter due to its obvious simplicity. One may in fact directly solve for its the threshold parameter θ from the expression for the mean in (1.15). A three parameter method of moments procedure for a triangular distribution with support $[a, b]$, would require to solve for the bounds a and b and the mode m, numerically. For example, for a fixed a and b one could standardize the data on $[0, 1]$ and next solve for $\widehat{\theta}$ using the simple expression for the mean (1.15). Next, one could evaluate the least squares error of the second and third central moment of the standardized data utilizing the straightforward expressions for the variance and third moment about the mean as given by (1.16) and (1.17), respectively, and minimize this least squares error over the domain $a < Z_{(1)}$, $b > Z_{(s)}$ (similar to the maximization of the likelihood function $G(a, b)$ (1.48) introduced in Sec. 1.5). We suggest here a minimization procedure since there is no guarantee that a solution will be obtained when equating the first three sample moments to the theoretical ones. Steps used in the outlined methods of moments procedure may be somewhat tedious, but do not pose any intrinsic difficulties.

There are of course many topics and applications of triangular distributions which we were not able to cover in this chapter mainly due to space limitations. For completeness we are including in the bibliography citations of a number of papers not mentioned in the text that could be of interest to our diligent readers. These are appended by a star.