Probability distribution theory, generalisations and applications of l-moments

Elsayed Ali Habib Elamir

Probability distribution theory, generalisations and applications of L-moments

A thesis presented for the degree
of Doctor of Philosophy at the
University of Durham

Elsayed Ali Habib Elamir

Department of Mathematical Sciences,

University of Durham,

Durham, DH1 3LE.

May 2001

Abstract

In this thesis, we have studied L-moments and trimmed L-moments (TL-moments) which are both linear functions of order statistics. We have derived expressions for exact variances and covariances of sample L-moments and of sample TL-moments for any sample size n in terms of first and second-order moments of order statistics from small conceptual sample sizes, which do not depend on the actual sample size n. Moreover, we have established a theorem which characterises the normal distribution in terms of these second-order moments and the characterisation suggests a new test of normality.

We have also derived a method of estimation based on TL-moments which gives zero weight to extreme observations. TL-moments have certain advantages over L-moments and method of moments. They exist whether or not the mean exists (for example the Cauchy distribution) and they are more robust to the presence of outliers.

Also, we have investigated four methods for estimating the parameters of a symmetric lambda distribution: maximum likelihood method in the case of one parameter and L-moments, LQ-moments and TL-moments in the case of three parameters. The L-moments and TL-moments estimators are in closed form and simple to use, while numerical methods are required for the other two methods, maximum likelihood and LQ-moments. Because of the flexibility and the simplicity of the lambda distribution, it is useful in fitting data when, as is often the case, the underlying distribution is unknown. Also, we have studied the symmetric plotting position for quantile plot assuming a symmetric lambda distribution and conclude that the choice of the plotting position parameter depends upon the shape of the distribution.

Finally, we propose exponentially weighted moving average (EWMA) control charts to monitor the process mean and dispersion using the sample L-mean and sample L-scale and charts based on trimmed versions of the same statistics. The proposed control charts limits are less influenced by extreme observations than classical EWMA control charts, and lead to tighter limits in the presence of out-of-control observations.

Dedication

I dedicate this to my parents, and all my brothers and sisters, for the support they have given me, especially over the last three years as a student in Durham.

Finally, to my family for always supporting me in everything that I do, and in particular, my wife, my children, Omnia and Ali, who could not possibly have known the extent to which the work for my Ph.D. would take over my life.

Acknowledgements

For a great many reasons, the three years I have spent working on my Ph.D have been the most interesting and enjoyable of my life. There are many people who have supported and encouraged me over the last few years. Unfortunately I have space to thank just a few of them here.

Thanks, first and foremost must go to my supervisor, Allan Seheult, for his patient help and wonderful insights over my thesis, especially, Chapter 2 and Chapter 4, and also for his understanding, both on a personal and academic level. I'd like to thank the whole Maths department, and the Statistics group, for their friendliness and inspiration during my Ph.D.

I would also like to thank the Zagazig University for the grant which has paid my way for the last three years and all my friends who have made my life easy in Durham.

Declaration

Some of the ideas in Chapters 2 and 4 were the results of joint work with Allan Seheult, particularly, the characterisation theorem of normality and the method of maximum likelihood. The rest of the work for this thesis is entirely my own.

I declare that the material presented in this thesis has not been submitted previously for any degree in either this or any other university.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The subject of order statistics deals with properties and applications of ordered random variables and of functions of these variables. If the random variables $\{X_i\}$, $i = 1, 2, ..., n$ are arranged in ascending order of magnitude and then written as

$$X_{1:n} \leq X_{2:n} \leq ... \leq X_{n:n} \tag{1.1}$$

then $X_{i:n}$ is said to be the ith-order statistics in a sample of size n. In the usual random sampling theory, the unordered X_i are assumed to be statistically independent and identically distributed. Because of the inequality relations among them, the order statistics $X_{i:n}$ are necessarily dependent. Some frequently encountered functions of order statistics are the extremes $X_{1:n}$ and $X_{n:n}$, the range $R = X_{n:n} - X_{1:n}$, the extreme deviate from the sample mean, $X_{n:n} - \overline{X}$, and for a random sample from a normal distribution $\mathrm{N}(\mu, \sigma^2)$, the studentized range, R/S_v, where S_v is a root mean square estimator of σ based on v degrees of freedom; see, for example, David (1981). All of these statistics have important applications. The extremes arise in the statistical study of floods and droughts, as well as in breaking strength and fatigue failure studies, the range is widely used in the field of quality control as a quick estimator of process standard deviation σ, the extreme deviate is a basic tool in procedures for detecting outliers and large values of $\left(X_{n:n} - \overline{X}\right)/\sigma$ suggest the presence of outliers, and when outliers are not confined to one direction, the studentized range is also useful in the detection process; see, for example, Barnett and Lewis (1994).

Sarhan and Greenberg (1962) used linear functions of order statistics in conjunction with the Gauss-Markov theorem to systematically estimate location and scale parameters in both complete and censored samples. They provided tables of the coefficients necessary for the calculations of these estimates from samples varying in size from 2 to 20. Other applications of order statistics arise in the study of reliability systems. A system of n components is called a k-out-of-n system if it remains operational only if at least k components continue to function. For components with independent lifetime distributions, the time to failure of the system is thus the $(n-k+1)$ th-order statistic. The special cases $k=1$ and $k=n$ correspond respectively to parallel and series systems.

A major impetus for the study of order statistics has been provided by the development of modern computers. Through their use it is feasible to make repeated examinations of the same data in many different ways. Tukey (1970) and Mosteller and Tukey (1977) have employed various informal techniques in the analysis of data. It is possible to determine quickly if the data are in accord with an assumed distribution and with an assumed model. A plot of the ordered observations against some simple functions of their ranks, preferably on probability paper appropriate for the assumed distribution, will often prove helpful in making such determinations.

The term robust statistics has many meanings, we use it in a relatively narrow sense: *...robustness signifies insensitivity to small deviations from the assumption of normality....* see Huber (1981).

Tukey (1960) points out that for a sample from $\mathrm{N}(\mu, \sigma^2)$ the mean deviation has asymptotic efficiency 0.88 relative to the standard deviation in estimating σ. The situation is changed drastically if some contamination by a wider normal, for example $\mathrm{N}(\mu, 9\sigma^2)$ is present: as little as 0.008 of the wider population will render the mean deviation asymptotically superior. Nevertheless there are flaws: the efficiency of the mean is very small for a uniform parent, and for any parent a single wild observation may render $\overline{X}$ useless. It has long been known that the midpoint, $(X_{1:n} + X_{n:n})/2$, is optimal in the former case but much worse than $\overline{X}$ in the latter, and that the median is preferable in the latter case but worse in the former. Obviously, we must not expect an estimator to be good under too wide a set of circumstances.

1.1 Motivation and outline of the thesis

It is standard statistical practice to summarise a probability distribution or an observed data set by some of its moments. It is also common, when fitting a parametric distribution to a data set, to estimate the parameters by equating the sample moments to those of the fitted distribution. The method of moments is not always satisfactory: sometimes it is difficult to assess exactly what information about the shape of a distribution is conveyed by its moments of third and higher order, the numerical values of sample moments, particularly when the sample is small, can be very different from those of the probability distribution from which the sample was drawn, and the estimated parameters of distributions fitted by the method of moments are often markedly less accurate than those obtainable by other estimation procedures such as the method of maximum likelihood; see, for example, Vogel and Fennessey (1993) and Kirby (1974).

Many statistical techniques are based on the use of linear combinations of order statistics but there has not been developed a unified theory covering the characterisation of probability distributions, the summarisation of observed data samples, the fitting of probability distributions to data and the testing of hypotheses about fitted distributions, until Hosking introduced L-moments in 1990, the L in L-moments emphasises the construction of L-moments from linear combinations of order statistics.

Greenwood et al. (1979) have introduced probability weighted moments, and they used them as a basis for estimating the parameters of some known distributions, for example, the Gumbel distribution. Hosking (1990) has studied an alternative approach based on quantities, which he called L-moments, which are analogous to the conventional moments but can be estimated by linear combinations of order statistics (L-statistics). L-moments have the theoretical advantages over conventional moments of being able to characterise a wider range of distributions and of being more robust to the presence of outliers of the data.

In Chapter 2, we review L-moments and probability weighted moments. We derive the exact variances and covariances of sample L-moments in terms of first and second-moments of the order statistics from small samples. We also characterise the normal distribution in terms of the covariances between certain sample L-moments. Also, we have derived distribution-free unbiased estimators of variances and covariances of sample L-moments. We also discuss probability weighted moments and their relation to L-moments and obtain exact variances and covariances of the sample probability weighted moments.

In Chapter 3, we review LQ-moments and extend the idea of L-moments to trimmed L-moments (TL-moments) and show that population TL-moments are able to characterise a wider range of distributions, for example a location measure of the Cauchy distribution. Also, we show TL-mean is a robust measure of location, protects against outliers and gives different weights for the observations. We finally define and study trimmed probability weighted moments (TPWM).

In Chapter 4, we investigate four methods of estimating the parameters of the symmetric lambda distribution: maximum likelihood in the case of a single parameter and L-moments, LQ-moments and TL-moments in the case of three parameters. We have also shown that the estimators based on L-moments and TL-moments are in closed form and simple to compute. Also, we have studied symmetric plotting position for quantile plot when sampling from a symmetric lambda distribution.

In Chapter 5, we review control charts and develop exponentially weighted moving average control charts for a process mean and standard deviation which incorporates an L-scale estimate of the process standard deviation, and we also describe trimmed versions of these charts. We have investigated out of control points of these charts by simulation.

Chapter 2

L-moments

2.1 Introduction

Hosking (1990) introduced population L-moments $\lambda_1, \lambda_2, ...$ as robust alternatives to classical measures of location, dispersion, skewness and kurtosis based on central moments and has studied properties of their corresponding sample L-moments $l_1, \ldots, l_n$ for samples of size n from any continuous distribution. Sample L-moments which can be expressed as linear combinations of the sample order statistics, are unbiased for the corresponding population quantities $\lambda_1, \ldots, \lambda_n$, and Hosking (1990) has given expressions for their asymptotic variances and covariances. An example of a sample L-moment is Gini's mean difference scale estimate g which is twice the sample L-moment l_2 and therefore has expectation $2\lambda_2$. Nair (1936) derived the standard error of g for any continuous distributions and Lomoniki (1951) obtained in a different way a general expression for the standard error of g when sampling is from any continuous distribution.

In this chapter, see also Elamir and Seheult (2001b), we derive expressions for the exact variances and covariances of sample L-moments in terms of first and second-moments of order statistics from small samples. For example, the variance of Gini's mean difference g depends only on the mean and covariance structure of the order statistics for conceptual samples of sizes 1, 2 and 3. We give examples of the use of these formulae for various distributions.

In section 2.2 we review classical moments. In sections 2.3 and 2.4 definitions and equivalent expressions for population and sample L-moments are given. In section 2.5 we derive exact results for the mean and variance-covariance structure of sample L-moments for any univariate continuous distribution. In section 2.6 we show how to derive distribution-free unbiased estimators of the variances and covariances of sample

L-moments and give two examples. In section 2.7 we establish a theorem which characterises the normal distribution in terms of sample L-moments. In section 2.8 we apply these results to obtain exact variances and covariances for sample probability weighted moments.

2.2 Moments

Let $X_1, X_2, ..., X_n$ be a sample from a distribution function $F_X(.)$. The shape of a unimodal probability distribution has traditionally been described by the moments of the distribution. The moments are the mean

$$\mu = \mathrm{E}(X) \tag{2.1}$$

and the higher central moments

$$\mu_r = \mathrm{E}(X - \mu)^r, \qquad r = 2, 3, ... \tag{2.2}$$

The mean is the centre of location of the distribution. The dispersion of the distribution about its centre is measured by the standard deviation

$$\sigma = \mu_2^{1/2} \tag{2.3}$$

or the variance σ^2.

Dimensionless higher moments $\mu_r/\mu_2^{r/2}$ are also used, in particular the skewness

$$\beta_1 = \mu_3/\mu_2^{3/2} \tag{2.4}$$

and the kurtosis

$$\beta_2 = \mu_4/\mu_2^2 \tag{2.5}$$

Analogous quantities can be computed from a data sample $X_1, X_2, ..., X_n$. The sample mean

$$\overline{X} = \sum_{i=1}^{n} X_i/n \tag{2.6}$$

is the natural estimator of μ.

The higher sample moments

$$m_r = n^{-1} \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^r \qquad r = 2, 3, ... \tag{2.7}$$

are reasonable estimators of the μ_r, but are not unbiased. For more details about moments; see, for example, Kendall and Stuart (1987) and Mood et al. (1974).

The method of moments estimates the parameters by finding expressions for them in terms of the lowest possible order moments and then substituting sample moments into the expressions.

For example, suppose we wish to estimate two parameters, θ_1 and θ_2. If θ_1 and θ_2 can be expressed in terms of the first two moments as

$$\theta_1 = f_1(\mu_1, \mu_2) \quad \text{and} \quad \theta_2 = f_2(\mu_1, \mu_2) \tag{2.8}$$

then the method of moments estimates are

$$\widehat{\theta}_1 = f_1(m_1, m_2) \quad \text{and} \quad \widehat{\theta}_2 = f_2(m_1, m_2) \tag{2.9}$$

2.3 Population L-moments

Let $X_1, X_2, ..., X_n$ be a sample from a continuous distribution function $F_X(.)$ with quantile function $Q_X(u) = F_X^{-1}(u)$ or $Q(u)$ for simplicity, where $0 < u < 1$. Denote by $X_{1:n} \leq X_{2:n} \leq ... \leq X_{n:n}$ the order statistics from a random sample $X_1, X_2, ..., X_n$ of size n drawn from $F(x)$.

Sillito (1969) and Hosking (1990) defined the population L-moments λ_r as follows

$$\lambda_r = r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} \mathrm{E}\left(X_{r-k:r}\right) \qquad r = 1, 2, ... \tag{2.10}$$

Hosking (1990) gives the expression

$$\lambda_r = \int_0^1 P_{r-1}(u) Q(u) du \qquad r = 1, 2, ... \tag{2.11}$$

where

$$P_r(u) = \sum_{k=0}^{r} c_{r,k} u^k \tag{2.12}$$

and

$$c_{r,k} = (-1)^{r-k} \binom{r}{k} \binom{r+k}{k} \tag{2.13}$$

It is straightforward to establish from (2.10) and (2.11) the following equivalent expressions for the first four L-moment

$$
\begin{aligned}
\lambda_1 &= \mathrm{E}\left(X_{1:1}\right) = \int_0^1 Q\left(u\right)\, du \\
\lambda_2 &= \frac{1}{2}\mathrm{E}\left(X_{2:2} - X_{1:2}\right) = \int_0^1 \left(2u-1\right) Q(u)\, du \\
\lambda_3 &= \frac{1}{3}\mathrm{E}\left(X_{3:3} - 2X_{2:3} + X_{1:3}\right) = \int_0^1 \left(6u^2 - 6u + 1\right) Q(u)\, du \\
\lambda_4 &= \frac{1}{4}\mathrm{E}\left(X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}\right) = \int_0^1 \left(20u^3 - 30u^2 + 12u - 1\right) Q(u)\, du
\end{aligned}
$$

Where λ_1 is a measure of location (mean) and λ_2 is a measure of scale. The scale-free quantities $\tau_3 = \lambda_3/\lambda_2$ and $\tau_4 = \lambda_4/\lambda_2$ are measure of skewness and kurtosis which are less sensitive to the extreme tails of the distribution than β_1 and β_2, the usual measures of skweness and kurtosis. For more details; see, for example, Kirby (1974), Oja (1981), Kaigh and Driscoll (1987) and Vogel and Fennessey (1993).

2.4 Sample L-moments

Hosking (1990) defined the sample L-moment l_r, corresponding to the population L-moment λ_r given in section 2.3 as follows

$$
l_r = \binom{n}{r}^{-1} \sum_{1 \le i_1} \sum_{< i_2 <} \cdots \sum_{< i_r \le n} \frac{1}{r} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} X_{i_{r-k}:n} \qquad r = 1, \ldots, n \tag{2.14}
$$

Thus, from (2.14) the four sample moments corresponding to λ_1, λ_2, λ_3 and λ_4 are

$$
\begin{aligned}
l_1 &= \frac{1}{n} \sum_{1 \le i \le n} X_{i:n} \\
l_2 &= \frac{1}{2\binom{n}{2}} \sum_{1 \le i} \sum_{< j \le n} \left(X_{j:n} - X_{i:n}\right) \\
l_3 &= \frac{1}{3\binom{n}{3}} \sum_{1 \le i} \sum_{< j <} \sum_{k \le n} \left(X_{k:n} - 2X_{j:n} + X_{i:n}\right) \\
l_4 &= \frac{1}{4\binom{n}{4}} \sum_{1 \le i} \sum_{< j} \sum_{< k <} \sum_{l \le n} \left(X_{l:n} - 3X_{k:n} + 3X_{j:n} - X_{i:n}\right)
\end{aligned}
$$

As Hosking (1990) has pointed out, it is not necessary to iterate over all subsamples of size r when calculating l_r, as it can be written as linear combination of order statistics; see also, Blom (1989). Hosking et al. (1985) show that l_r may be written

$$l_r = \sum_{k=0}^{r-1} c_{r-1,k} b_k \tag{2.15}$$

where $c_{r,k}$ is given in (2.13) and

$$b_k = \frac{1}{n} \sum_{i=1}^{n} \frac{(i-1)(i-2)\ldots(i-k)}{(n-1)(n-2)\cdots(n-k)} X_{i:n} \tag{2.16}$$

which can be written more compactly as

$$b_k = \frac{1}{n^{(k+1)}} \sum_{i=1}^{n} (i-1)^{(k)} X_{i:n} \tag{2.17}$$

where $n^{(r)} = n(n-1)\cdots(n-r+1)$.

Thus, for example, we can re-express the first four sample L-moments in the readily computable forms

$$\begin{aligned}
l_1 &= \frac{1}{n} \sum_{i=1}^{n} X_{i:n} \\
l_2 &= \frac{1}{n(n-1)} \sum_{i=1}^{n} (2i-1-n) X_{i:n} \\
l_3 &= \frac{1}{n(n-1)(n-2)} \sum_{i=1}^{n} [6\,(i-1)\,(i-2) - 6\,(n-2)\,(i-1) \\
&\quad + (n-1)\,(n-2)]\, X_{i:n} \\
l_4 &= \frac{1}{n(n-1)(n-2)(n-3)} \sum_{i=1}^{n} [20\,(i-1)\,(i-2)\,(i-3) - 30\,(n-3) \\
&\quad \times (i-1)\,(i-2) + 12\,(n-2)\,(n-3)\,(i-1) - (n-1)\,(n-2)\,(n-3)]\, X_{i:n}
\end{aligned}$$

We see that l_1 is the sample mean, l_2 is half Gini's mean difference g, l_3 is used by Sillito (1969) as a measure of symmetry and by Locke and Spurrier (1976) to test for symmetry, and l_4 is used by Hosking (1990) as a measure of kurtosis. Standardised, unit-free versions of the symmetry and kurtosis measures are $t_3 = l_3/l_2$ and $t_4 = l_4/l_2$ corresponding to the populations versions $\tau_3 = \lambda_3/\lambda_2$ and $\tau_4 = \lambda_4/\lambda_2$ described in section 2.3.

Hosking and Wallis (1995) gave the following biased estimator of population L-moments λ_r given in section 2.3 as

$$\widetilde{l}_{r+1} = (-1)^r \sum_{k=0}^{r} (-1)^{r-k} \binom{r}{k} \binom{r+k}{k} \widetilde{\alpha_k} \tag{2.18}$$

where

$$\widetilde{\alpha}_k = n^{-1} \sum_{i=1}^{n} (1 - p_{i:n})^k X_{i:n} \quad \text{and} \quad p_{i:n} = \frac{i+\gamma}{n+\delta} \quad \text{with} \quad \delta > \gamma > -1 \tag{2.19}$$

where $p_{i:n}$ is the plotting position, a distribution-free estimator of $F_X(x_{i:n})$, the non-exceedance probability of $X_{i:n}$. $\widetilde{l}_{r+1}$ gives good estimates of the tails of the quantile $Q(u)$; see, for example, Landweher et al. (1979), Hosking et al. (1985) and Hosking and Wallis (1987). We shall discuss the choice of γ and δ for $p_{i:n}$ in Chapter 4.

2.4.1 Asymptotic variances and covariances of sample L-moments

Asymptotic theory for linear combinations of order statistics, developed by Chernoff et al. (1967), Moor (1968) and Stigler (1974), can be immediately applied to estimators of L-moments. The asymptotic theory usually provides a good approximation to the exact distribution for samples of size $n > 50$, and is often adequate even for $n \geq 20$, see Figures 2.1, 2.2, 2.3 and 2.4 and Table 2.3. Hosking (1986) proved that

I. $n^{1/2}(l_r - \lambda_r)$, $r = 1, 2, ..., m, m \leq n$, converges in distribution to the multivariate normal distribution $\mathrm{N}(0, \Lambda)$, where the elements Λ_{rs} $(r, s = 1, 2, ..., m)$ of Λ are given by

$$\begin{aligned} \Lambda_{rs} &= \int\int_{x<y} \{P_{r-1}[F(x)] P_{s-1}[F(y)] + P_{s-1}[F(x)] P_{r-1}[F(y)]\} \\ &\quad \times F(x)(1 - F(y))\, dxdy \end{aligned} \tag{2.20}$$

where $P_r(F)$ is as in equation (2.12).

II. Let $\tau_r = \lambda_r / \lambda_2$ and $t_r = l_r / l_2$, $r = 3, 4, ..., m$. Then as $n \to \infty$ the vector

$$n^{1/2}\left[(l_1 - \lambda_1)\ (l_2 - \lambda_2)\ (t_3 - \tau_3)\ (t_4 - \tau_4) \ldots (t_m - \tau_m)\right]^T$$

converges in distribution to the multivariate normal distribution $\mathrm{N}(0, T)$ where the elements T_{rs} $(r, s = 1, 2, ..., m)$ of T are given by

$$T_{rs} = \begin{cases} \Lambda_{rs} & \text{if } r \leq 2,\ s \leq 2 \\ \left(\Lambda_{rs} - \tau_r \Lambda_{2s}\right)/\lambda_2 & \text{if } r \geq 3,\ s \leq 2 \\ \left(\Lambda_{rs} - \tau_r \Lambda_{2s} - \tau_s \Lambda_{2r} + \tau_r \tau_s \Lambda_{22}\right)/\lambda_2^2 & \text{if } r \geq 3,\ s \geq 3 \end{cases} \tag{2.21}$$

Asymptotic theory has many practical applications. It makes possible the construction of confidence limits for population L-moments and it can give approximations to finite sample distributions. Also, Taylor series expansion enables the asymptotic theory to be applied to functions of sample L-moments.

Figures 2.1 and 2.2 show histograms and quantile plots of l_1, l_2, t_3 and t_4 for samples of sizes 15 and 50 drawn from standard normal distribution and Figures 2.3 and 2.4 show histograms and quantile plots of l_1, l_2, t_3 and t_4 for samples of sizes 15 and 50 drawn from standard exponential distribution. Figure 2.3 shows a slight curvature for a sample size 15 from the exponential distribution. Gail and Gastwirth (1978) prove that the normal approximation to l_2 is correct asymptotically, which is also illustrated by Figure 2.4 when the sample from exponential distribution is of size 50.

2.5 Exact mean and covariance structure for sample L-moments

It would appear from equation (2.15) that to compute the variance structure of l_1, l_2, l_3, l_4 we need the full covariance structure of $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$. The main theme in this section is that much less is required.

In what follows we will assume that we are working with certain convenient standardised versions

$$Y_{i:n} = \frac{X_{i:n} - \mu}{\sigma} \qquad i = 1, 2, ..., n \tag{2.22}$$

of the order statistics $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ usually from a scale-location family

$$F\left(\frac{x-\mu}{\sigma}\right)$$

with baseline distribution function $F(y)$; for example, $F(y)$ may be the unit Gaussian distribution $\mathrm{N}(0, 1)$. Thus, results for expectations, variances and covariances for $Y_{1:n} \leq Y_{2:n} \leq ... \leq Y_{n:n}$ readily convert to those for $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ in the usual way.

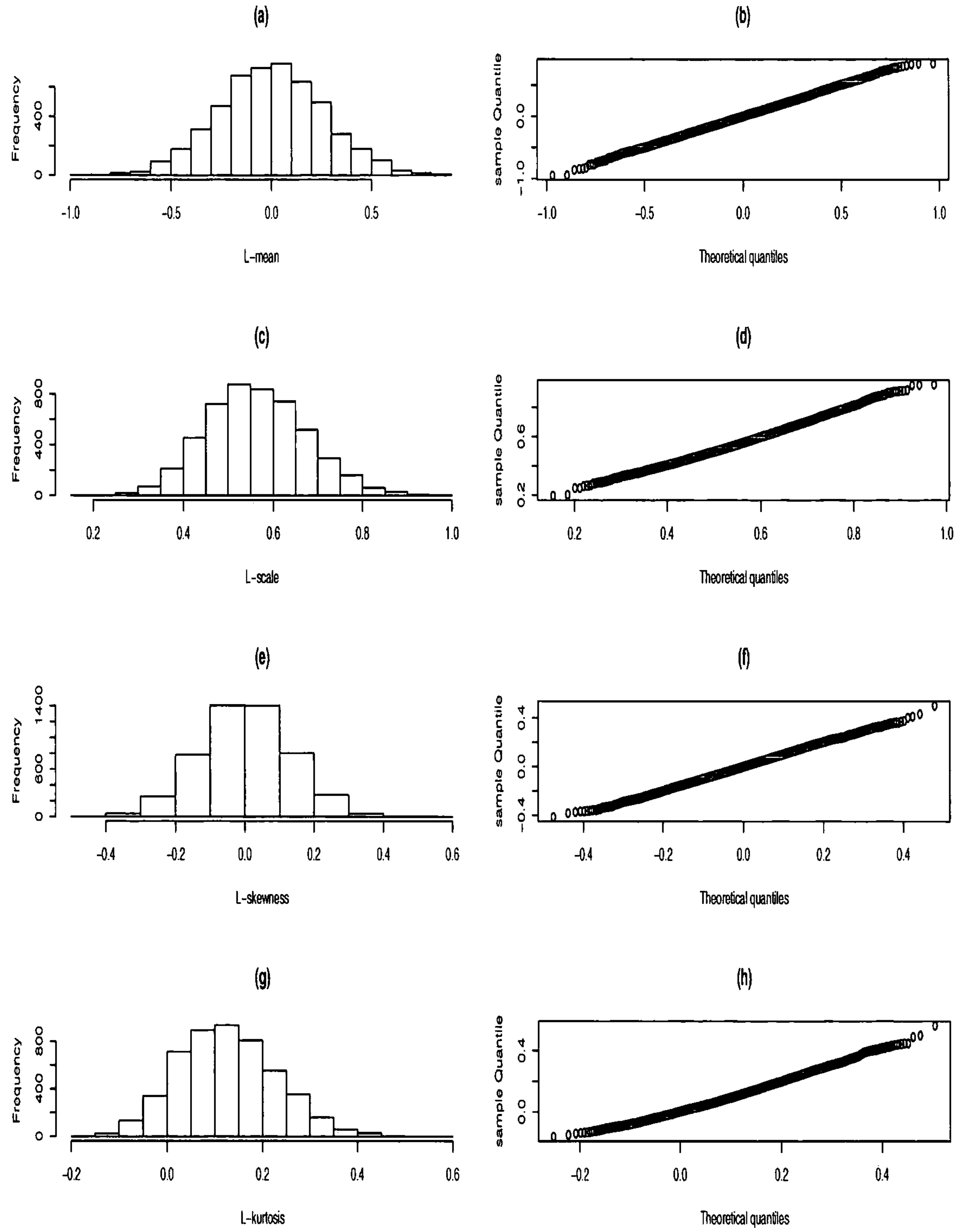

Figure 2.1: The histogram and quantile plots for l_1 in (a) and (b), for l_2 in (c) and (d), for t_3 in (e) and (f) and for t_4 in (g) and (h) when the parent distribution is normal $(0, 1)$, the sample size $n = 15$, and number of replications is 5000.

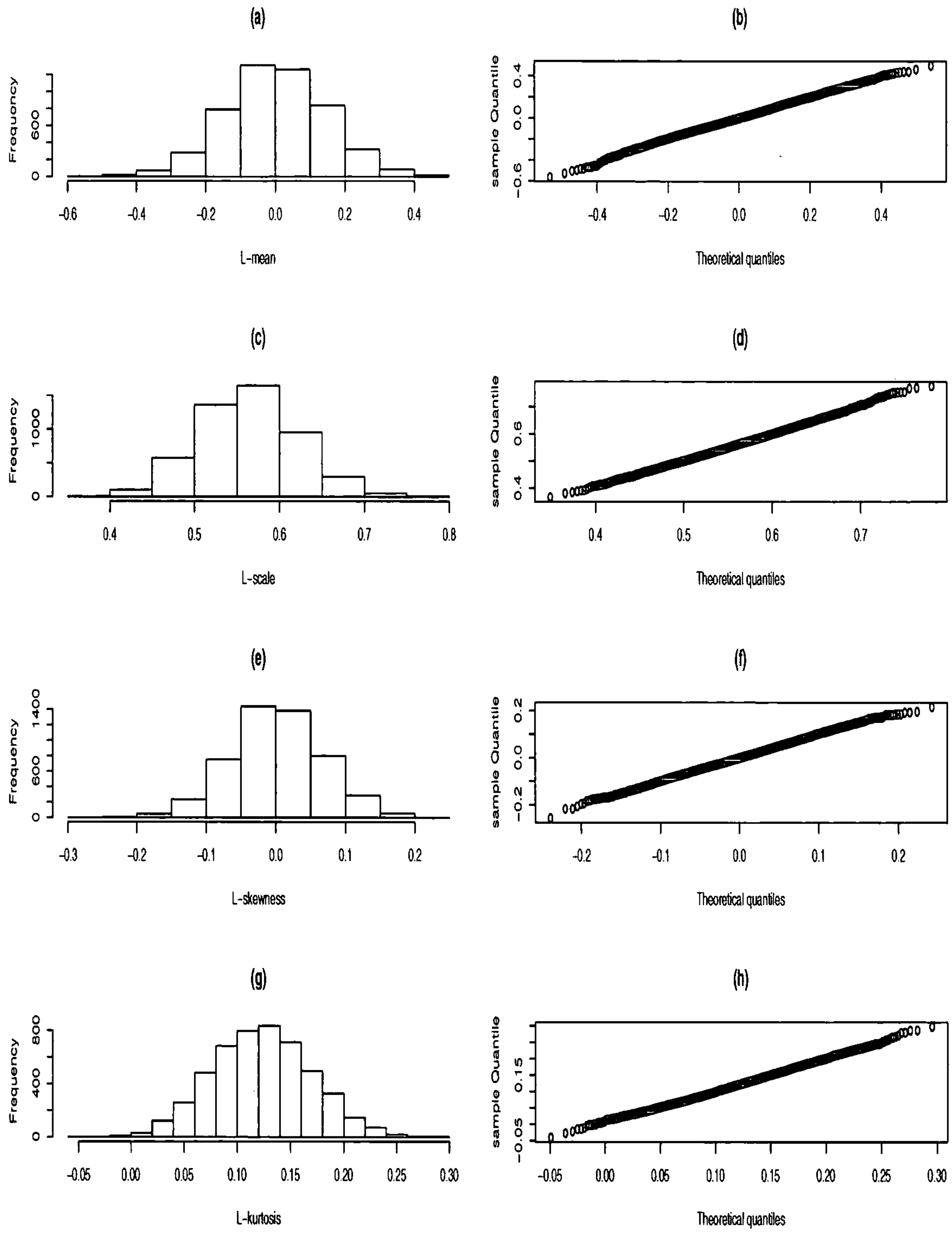

Figure 2.2: The histogram and quantile plots for l_1 in (a) and (b), for l_2 in (c) and (d), for t_3 in (e) and (f) and for t_4 in (g) and (h) when the parent distribution is normal $(0,1)$, the sample size $n=50$, and number of replications is 5000.

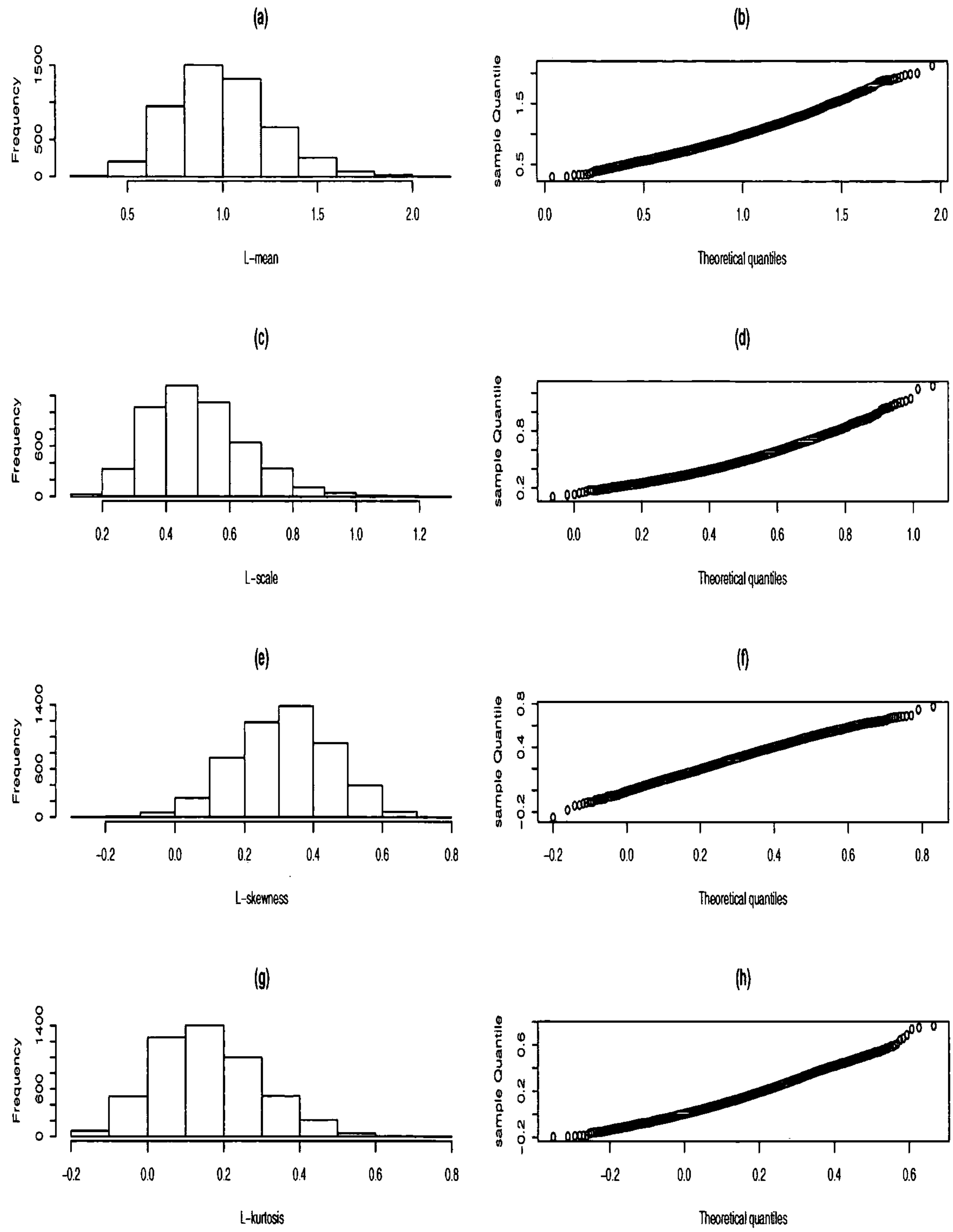

Figure 2.3: The histogram and quantile plots for l_1 in (a) and (b), for l_2 in (c) and (d), for t_3 in (e) and (f) and for t_4 in (g) and (h) when the parent distribution is exponential (1), the sample size $n = 15$, and number of replications is 5000.

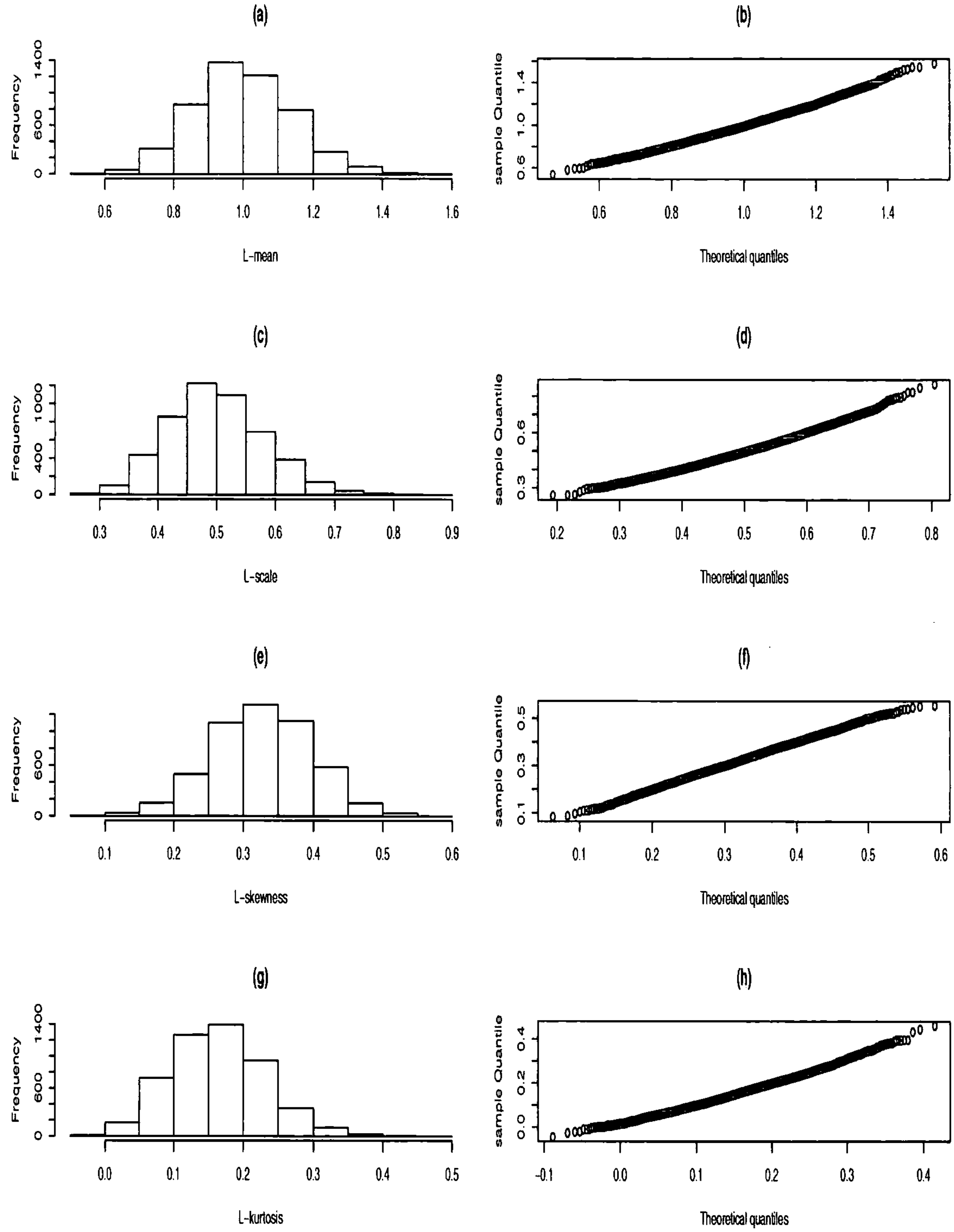

Figure 2.4: The histogram and quantile plots for l_1 in (a) and (b), for l_2 in (c) and (d), for t_3 in (e) and (f) and for t_4 in (g) and (h) when the parent distribution is exponential (1), the sample size $n = 50$, and number of replications is 5000.

Standard calculations show that the density function of $Y_{r:n}$ is

$$\frac{n!}{(r-1)!(n-r)!}F(u)^{r-1}[1-F(u)]^{n-r}\,dF(u)$$

and the joint density function of $Y_{r:n}$ and $Y_{s:n}$ $(r < s)$ can be written

$$\frac{n!}{(r-1)!(s-r-1)!(n-s)!}F(u)^{r-1}[F(v)-F(u)]^{s-r-1}[1-F(v)]^{n-s}\,dF(u)\,dF(v)$$

The following expressions for first and second order moments are well known; see, for example, David (1981):

$$\mathrm{E}\{Y_{r:n}\} = \frac{n!}{(r-1)!(n-r)!}\int_0^1 Q(u)u^{r-1}(1-u)^{n-r}\,du \tag{2.23}$$

$$\mathrm{E}\{Y_{r:n}^2\} = \frac{n!}{(r-1)!(n-r)!}\int_0^1 Q^2(u)u^{r-1}(1-u)^{n-r}\,du \tag{2.24}$$

$$\begin{aligned}\mathrm{E}\{Y_{r:n}Y_{s:n}\} &= \frac{n!}{(r-1)!(s-r-1)!(n-s)!}\int_0^1\int_0^v Q(u)Q(v)u^{r-1} \\ &\quad\times(v-u)^{s-r-1}(1-v)^{n-s}\,dudv\end{aligned} \tag{2.25}$$

where $Q(\cdot)$ is the quantile function for $F(y)$. Var$\{Y_{r:n}\}$ and Cov$\{Y_{r:n}, Y_{s:n}\}$ can be computed in the usual way from these expressions.

To calculate the mean and covariance structure of the sample L-moments $l_1, \ldots, l_n$, we make use of the following identities given by Downton (1966)

$$\sum_{r=1}^{n}(r-1)^{(k)}(n-r)^{(l)}\mathrm{E}\{Y_{r:n}\} = k!\,l!\binom{n}{k+l+1}\mathrm{E}\{Y_{k+1:k+l+1}\} \tag{2.26}$$

$$\sum_{r=1}^{n}(r-1)^{(k)}(n-r)^{(l)}\mathrm{E}\{Y_{r:n}^2\} = k!\,l!\binom{n}{k+l+1}\mathrm{E}\{Y_{k+1:k+l+1}^2\} \tag{2.27}$$

$$\sum_{1 \le r} \sum_{<s \le n} (r-1)^{(k)}(n-s)^{(l)} \mathrm{E}\{Y_{r:n} Y_{s:n}\} = k!\, l! \binom{n}{k+l+2} \mathrm{E}\{Y_{k+1:k+l+2} Y_{k+2:k+l+2}\} \tag{2.28}$$

To evaluate the mean and covariance structure of $l_1, \dots, l_n$, we first evaluate the mean and covariance structure of $b_1, \dots, b_{r-1}$ given in (2.17).

Using the definition of b_k in (2.17) and the identity in (2.26) with $l = 0$, we obtain

$$\mathrm{E}\{b_k\} = \frac{1}{(k+1)} \mathrm{E}\{Y_{k+1:k+1}\} \tag{2.29}$$

It then follows from (2.23) that

$$\mathrm{E}\{Y_{k+1:k+1}\} = (k+1) \int_0^1 Q(u) u^k \, du \tag{2.30}$$

and from (2.15) that

$$\mathrm{E}\{l_r\} = \sum_{k=0}^{r-1} c_{r-1,k} \mathrm{E}\{b_k\} = \sum_{k=0}^{r-1} c_{r-1,k} \int_0^1 Q(u) u^k \, du \tag{2.31}$$

Thus

$$\mathrm{E}\{l_r\} = \int_0^1 Q(u) \sum_{k=0}^{r-1} (-1)^{r-1-k} \binom{r-1}{k} \binom{r-1+k}{k} u^k \, du = \lambda_r$$

which shows that l_r is an *unbiased estimator* of λ_r. Moreover, in view of (2.10), this expectation can be evaluated in terms of the expectations of the order statistics from a sample just of size r. For example,

$$\mathrm{E}\{l_2\} = \frac{1}{2} \left(\mathrm{E}\{Y_{2:2}\} - \mathrm{E}\{Y_{1:2}\} \right) = \lambda_2 \tag{2.32}$$

which either can be evaluated explicitly or there is a once-and-for-all numerical evaluation, as is the case for normal order statistics. Hosking (1990) has given expressions for the first few L-moments for many standard distributions.

We now show that the variance-covariance structure for the sample L-moments can also be evaluated in terms of certain moments of order statistics from small samples with sizes which do not depend on the sample size n. Using the definition of b_k it follows that

$$\mathrm{E}\{b_k b_l\} = \frac{1}{n^{(k+1)}n^{(l+1)}}\mathrm{E}\left(\sum_{r=1}^{n}\sum_{s=1}^{n}(r-1)^{(k)}(s-1)^{(l)}\{Y_{r:n}Y_{s:n}\}\right) \tag{2.33}$$

now

$$\begin{aligned}\mathrm{E}\left[\sum_{r=1}^{n}\sum_{s=1}^{n}(r-1)^{(k)}(s-1)^{(l)}Y_{r:n}Y_{s:n}\right] &= \mathrm{E}\left[\sum_{r=1}^{n}(r-1)^{(k)}(r-1)^{(l)}Y_{r:n}^2\right] \\ &+ \mathrm{E}\left\{\sum_{1\le r}\sum_{<s\le n}\left[(r-1)^{(k)}(s-1)^{(l)}\right.\right. \\ &\left.\left. +(s-1)^{(k)}(r-1)^{(l)}\right]Y_{r:n}Y_{s:n}\right\} \end{aligned} \tag{2.34}$$

Downton (1966) has given the following binomial type relationships

$$(a+b)^{(m)} = \sum_{r=0}^{m}\binom{m}{r}a^{(r)}b^{(m-r)} \tag{2.35}$$

$$(a-b)^{(m)} = \sum_{r=0}^{m}(-1)^r\binom{m}{r}(a-r)^{(m-r)}\,b^{(r)} \tag{2.36}$$

If on the right-hand-side of (2.34) we write $(r-1)^{(k)} = ((r-l-1)+l)^{(k)}$ in the first term and then use the relation (2.35), write $(r-1)^{(k)} = ((n-1)-(n-r))^{(k)}$ or $(r-1)^{(l)} = ((n-1)-(n-r))^{(l)}$ in the second and third terms and then use the relation (2.36), and note further that $n^{(t+s)} = n^{(t)}(n-t)^{(s)}$, the required covariances can be written

$$\mathrm{Cov}\{b_k, b_l\} = \frac{1}{n^{(k+1)}}\sum_{s=0}^{k}A_{kl}^{(s)}(n-l-1)^{(s)} \tag{2.37}$$

where

$$\begin{aligned} A_{kl}^{(s)} &= \frac{k!l!\mathrm{E}\{Y_{s+l+1:s+l+1}^2\}}{(k-s)!(s+l-k)!s!(s+l+1)} + \frac{k!l!(l+1)^{(k+1-s)}}{(k+1-s)!} \\ &\quad \times\sum_{r=0}^{s-1}(-1)^r\frac{\mathrm{E}\{Y_{l+1:l+2+r}Y_{l+2:l+2+r}\}}{(l+2+r)!(s-1-r)!} \\ &- \frac{k!l!(k+1)^{(k+1-s)}}{(k+1-s)!}\sum_{r=0}^{k}(-1)^r\frac{\mathrm{E}\{Y_{s+l-k:s+l-k+1+r}Y_{s+l-k+1:s+l-k+1+r}\}}{(k-r)!(s+l-k+1+r)!} \\ &+ \frac{k!l!\mathrm{E}\{Y_{k+1:k+1}\}\left[\mathrm{E}\{Y_{s+l-k:s+l-k}\}-\mathrm{E}\{Y_{l+1:l+1}\}\right]}{(k+1-s)!(s+l-k)!s!} \end{aligned} \tag{2.38}$$

[Note that this derivation follows from Downton (1966), except that he mistakenly writes $(s+l+1)!$ in the denominator of the first term of $A_{kl}^{(s)}$ in his equivalent equation (3.18), instead of the correct expression, $(s + l + 1)$.]

As noted by Downton (1966),

> ...These coefficients depend only upon diagonal and next-diagonal terms of relatively small variance matrices of ordered observations and upon expected values of the largest observations.

Then, the variances and covariances of l_r can be obtained as follows

$$\text{Cov}\{l_r, l_s\} = \sum_{k=0}^{r-1}\sum_{l=0}^{s-1} c_{r-1,k}c_{s-1,l}\text{Cov}\{b_k, b_l\} \qquad r \leq s \tag{2.39}$$

In matrix notation, noting that equation (2.15) can be written $\mathbf{l} = \mathbf{C}\mathbf{b}$, where $\mathbf{l}^T = (l_1, \ldots, l_n)$, $\mathbf{b}^T = (b_0, \ldots, b_{r-1})$ and $\mathbf{C}$ is the $n \times r$ lower triangular matrix with entries c_{rk} given in (2.13). The variance matrix of $\mathbf{l}$ is

$$\text{Var}[\mathbf{l}] = \mathbf{C}\,\Theta\,\mathbf{C}^T \tag{2.40}$$

where $\Theta = \text{Var}\{\mathbf{b}\}$. In the next subsection we obtain explicit expressions for $\text{Var}\{\mathbf{b}\}$ for the first four L-moments for *any* sample of size n in terms of moments of order statistics from conceptual samples of size no more than *seven*.

2.5.1 Exact variances and covariances for the first four sample L-moments

As previously noted, the first four L-moments $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ provide population measures of location, dispersion, skewness and kurtosis, and we now focus on the covariance structure of their unbiased sample estimates l_1, l_2, l_3, l_4.

It follows from (2.40) that the variance matrix for l_1, l_2, l_3, l_4 can be computed from

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 1 & -6 & 6 & 0 \\ -1 & 12 & -30 & 20 \end{pmatrix} \begin{pmatrix} \theta_{00} & \theta_{01} & \theta_{02} & \theta_{03} \\ \theta_{10} & \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{20} & \theta_{21} & \theta_{22} & \theta_{23} \\ \theta_{30} & \theta_{31} & \theta_{32} & \theta_{33} \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 & -1 \\ 0 & 2 & -6 & 12 \\ 0 & 0 & 6 & -30 \\ 0 & 0 & 0 & 20 \end{pmatrix}$$

For example, explicit expressions for $\text{Var}\{l_1\}$, $\text{Var}\{l_2\}$ and $\text{Cov}\{l_1, l_2\}$ are given below

$$\text{Var}\{l_1\} = \theta_{00} = \frac{\text{E}\{Y_{1:1}^2\} - \text{E}^2\{Y_{1:1}\}}{n} \tag{2.41}$$

which is the variance of the sample mean, a well-established and familiar quantity.

$$\mathrm{Var}\{l_2\} = 4\theta_{11} - 4\theta_{01} + \theta_{00}$$

$$\begin{aligned}\mathrm{Var}\{l_2\} &= \Big\{\frac{4}{3}(n-2)\left(\mathrm{E}\{Y_{3:3}^2\} + \mathrm{E}\{Y_{1:3}Y_{2:3}\} + \mathrm{E}\{Y_{2:3}Y_{3:3}\}\right) - 2(n-3)\mathrm{E}\{Y_{1:2}Y_{2:2}\} \\ &- 2(n-2)\mathrm{E}\{Y_{2:2}^2\} + (n-1)\mathrm{E}\{Y_{1:1}^2\} - 2(2n-3)\mathrm{E}^2\{Y_{2:2}\} \\ &+ \mathrm{E}\{Y_{1:1}\}\left(4(2n-3)\mathrm{E}\{Y_{2:2}\} - 5(n-1)\mathrm{E}\{Y_{1:1}\}\right)\Big\}/(n(n-1)) \qquad (2.42)\end{aligned}$$

This is equivalent to the expression given by Nair (1936) for Gini's mean difference g.

$$\mathrm{Cov}\{l_1, l_2\} = 2\theta_{01} - \theta_{00}$$

$$\mathrm{Cov}\{l_1, l_2\} = \frac{(\mathrm{E}\{Y_{2:2}^2\} - \mathrm{E}\{Y_{1:2}Y_{2:2}\} - \mathrm{E}\{Y_{1:1}^2\}) + \mathrm{E}\{Y_{1:1}\}(3\mathrm{E}\{Y_{1:1}\} - 2\mathrm{E}\{Y_{2:2}\})}{n} \qquad (2.43)$$

This expression is new. The variances and covariances for l_1, l_2, l_3 and l_4 are given in the Appendix.

We now show that $\mathrm{Cov}\{l_1, l_2\} = 0$ for any symmetric distribution. From David (1981) we find for any parent distribution that

$$\sum_{r=1}^{n-1} \sum_{s=r+1}^{n} \mathrm{E}\left(Y_{r:n}^k Y_{s:n}^k\right) = \binom{n}{2} \mathrm{E}^2\left(Y_{1:1}^k\right) \qquad (2.44)$$

If the parent distribution is symmetric about the origin, then $\mathrm{E}(Y_{1:1}) = 0$ and from David (1981) we find

$$\mathrm{E}(Y_{r:n}Y_{s:n}) = \mathrm{E}(Y_{n-r+1:n}Y_{n-s+1:n}) \quad r, s = 1, 2, 3, ...n \qquad (2.45)$$

If n is even then

$$\mathrm{E}\left(Y_{\frac{n}{2}:n-1}^k\right) = \begin{cases} \mathrm{E}\left(Y_{\frac{n}{2}:n}^k\right) & k \text{ even} \\ 0 & k \text{ odd} \end{cases} \qquad (2.46)$$

When $n = 2, r = 1, k = 1$ and $s = 2$, we find from (2.44) that

$$\mathrm{E}\left(Y_{1:2}Y_{2:2}\right)=\left[\mathrm{E}\left(Y_{1:1}\right)\right]^2=0$$

With $n=2$, $r=2$ and $s=2$ in (2.45), and $n=2$ and $k=2$ in (2.46) we obtain

$$\mathrm{E}\left(Y_{2:2}^2\right)=\mathrm{E}\left(Y_{1:2}^2\right)\quad\text{and}\quad\mathrm{E}\left(Y_{1:1}^2\right)=\mathrm{E}\left(Y_{1:2}^2\right)$$

Substituting in equation (2.43), gives $\mathrm{Cov}\left(l_1,l_2\right)=0$ for any symmetric distribution.

2.6 Some applications of covariances of sample L-moments

In this section we apply the results of the previous section to illustrate how to obtain distribution-free unbiased estimators of the variances and covariances of sample L-moments, to two examples and to approximate means and variances of functions of sample L-moments, such as L-skewness and L-kurtosis .

2.6.1 Distribution-free estimators of variances and covariances of sample L-moments

Here we illustrate how to obtain distribution-free unbiased estimators of the variances and covariances of sample L-moments using the following expression due to Downton (1966)

$$\begin{aligned}
&\sum_{r=1}^{n}\sum_{s=1}^{n}(r-1)^{(i)}(s-1)^{(j)}\,\mathrm{E}\left(Y_{r:n}Y_{s:n}\right)= \\
&\sum_{t=j+1}^{i+j+1} n^{(t)}\left\{\frac{i!j!\mathrm{E}\left(Y_{t::t}^2\right)}{(i+j+1-t)!\,(t-i-1)!\,(t-j-1)!(t)}\right. \\
&+\frac{i!j!\,(j+1)^{(i+j+2-t)}}{(i+j+2-t)!}\sum_{r=0}^{t-j-2}(-1)^r\,\frac{\mathrm{E}\left(Y_{j+1:j+2+r}Y_{j+2:j+2+r}\right)}{(j+r+2)!\,(t-j-2-r)!} \\
&-\frac{i!j!\,(i+1)^{(i+j+2-t)}}{(i+j+2-t)!}\sum_{r=0}^{i}(-1)^r\,\frac{\mathrm{E}\left(Y_{t-i-1:t-i+r}Y_{t-i:t-i+r}\right)}{(i-r)!\,(t-i+r)!} \\
&\left.+\frac{i!j!\mathrm{E}\left(Y_{i+1:i+1}\right)\mathrm{E}\left(Y_{t-i-1:t-i-1}\right)}{(i+j+2-t)!\,(t-i-1)!\,(t-j-1)!}\right\} \\
&+\frac{n^{(i+j+2)}\mathrm{E}\left(Y_{i+1:i+1}\right)\mathrm{E}\left(Y_{j+1:j+1}\right)}{(i+1)\,(j+1)}
\end{aligned}\tag{2.47}$$

and the identity

$$\sum_{i=1}^{n} (i-1)^{(r)} Y_{i:n} \sum_{i=1}^{n} (i-1)^{(s)} Y_{i:n} = \sum_{i=1}^{n} \sum_{j=1}^{n} (i-1)^{(r)} (j-1)^{(s)} Y_{i:n} Y_{j:n} \tag{2.48}$$

From equations (2.26), (2.27) and (2.28) we find that

$$A(k,l) = \sum_{i=1}^{n} \frac{\binom{i-1}{k}\binom{n-i}{l}}{\binom{n}{k+l+1}} Y_{i:n} \tag{2.49}$$

$$W(k,l) = \sum_{i=1}^{n} \frac{\binom{i-1}{k}\binom{n-i}{l}}{\binom{n}{k+l+1}} Y_{i:n}^2 \tag{2.50}$$

$$W12(k,l) = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{\binom{i-1}{k}\binom{n-j}{l}}{\binom{n}{k+l+2}} Y_{i:n} Y_{j:n} \tag{2.51}$$

are unbiased estimators of $\mathrm{E}\left(Y_{k+1:k+l+1}\right)$, $\mathrm{E}\left(Y_{k+1:k+l+1}^2\right)$ and $\mathrm{E}\left(Y_{k+1:k+l+2} Y_{k+2:k+l+2}\right)$, respectively.

For example, distribution-free unbiased estimator $\widehat{\mathrm{Var}}\left(l_1\right)$ of $\mathrm{Var}\left(l_1\right)$ can be written as

$$\widehat{\mathrm{Var}}\left(l_1\right) = \frac{W(0,0) - [A(0,0)]^2}{n-1} = \frac{s^2}{n} \tag{2.52}$$

where s^2 is the usual unbiased estimator of σ^2.

Similarly, the following expressions are unbiased estimators of $\mathrm{Var}\left(l_2\right)$ and $\mathrm{Cov}\left(l_1, l_2\right)$, respectively.

$$\begin{aligned}\widehat{\text{Var}}(l_2) &= \Big\{\frac{4}{3}(n-2)\left[W(2,0)+W12(0,1)+W12(1,0)\right]-2(n-3)\,W12(0,0) \\ &\quad -2(n-2)\,W(1,0)+\left[\frac{n^2-n+2}{n-1}\right]W(0,0)-2(2n-3)\,A^2(1,0)+A(0,0) \\ &\quad \left[4(2n-3)\,A(1,0)-\left[\frac{5n^2-9n+6}{n-1}\right]A(0,0)\right]\Big\}/((n-2)(n-3)) \qquad (2.53)\end{aligned}$$

and

$$\begin{aligned}\widehat{\text{Cov}}(l_1,l_2) &= \Big\{W(1,0)-W12(0,0)-\frac{n}{n-1}W(0,0)-2A(0,0)A(1,0) \\ &+ \frac{(3n-2)A^2(0,0)}{n-1}\Big\}/(n-2) \qquad (2.54)\end{aligned}$$

To obtain these particular expressions, and in general, we use (2.47) to correct the bias in using $A(k,l)\times A(k',l')$ as an estimator of $\text{E}(Y_{k+1:k+l+1})\times\text{E}(Y_{k'+1:k'+l'+1})$.

2.6.2 Examples

We now give two illustrations of exact variances and covariances of sample L-moments.

Example 1.

In this example we give exact variances and covariances of the first four sample L-moments for four standard distributions: the standard normal distribution (using tables from Teichroew (1954)), the uniform distribution (with pdf $f(x)=\sqrt{3}/6$ on $-3/\sqrt{3}<x<3/\sqrt{3}$), the Gumbel distribution (with pdf $f(x)=e^{-x}e^{-e^{-x}}$ on $-\infty<x<\infty$) and the exponential distribution (with pdf $f(x)=e^{-x}$ on $x>0$). The results are given in Tables 2.1 and 2.2 and all four distributions have variance $\sigma^2=1$.

In Table 2.3 we compare asymptotic and exact variances for different sample sizes from the Gumbel (Gum) and the Normal (Nor) distributions. As we see, the asymptotic variances of l_2 and l_3 are good approximations when $n\geq 20$ but underestimates for small n. Note that, the asymptotic variances of l_2 and l_3 are from Hosking (1986).

Example 2

We give a numerical example from Rice (1995). The data are given in the Table 2.4 and Figure 2.5 shows the quantile normal plot which Rice has used to support the normality assumption for these data. We may estimate the parameters using

	Normal	Uniform	Divisor
Var(l_1)	1	1	$n^{(1)}$
Cov(l_1, l_2)	0	0	$n^{(1)}$
Cov(l_1, l_3)	0	$-\frac{1}{5}$	$n^{(1)}$
Cov(l_1, l_4)	0	0	$n^{(1)}$
Var(l_2)	$0.163n + 0.038$	$\frac{n+3}{15}$	$n^{(2)}$
Cov(l_2, l_3)	0	0	$n^{(2)}$
Cov(l_2, l_4)	$0.011n + 0.0002$	$-\frac{n+3}{35}$	$n^{(2)}$
Var(l_3)	$0.059n^2 + 0.049n + 0.010$	$\frac{2n^2+10}{35}$	$n^{(3)}$
Cov(l_3, l_4)	0	0	$n^{(3)}$
Var(l_4)	$0.028n^3 + 0.056n^2 + 0.055n + 0.014$	$\frac{2n^3+6n^2+22n+66}{105}$	$n^{(4)}$

Table 2.1: Exact variances and covariances of the first four sample L-moments from normal and uniform distribution, both with $\sigma^2 = 1$.

	Gumbel	exponential	Divisor
Var(l_1)	1	1	$n^{(1)}$
Cov(l_1, l_2)	$\frac{0.281}{n}$	$\frac{1}{2}$	$n^{(1)}$
Cov(l_1, l_3)	0.075	$\frac{1}{6}$	$n^{(1)}$
Cov(l_1, l_4)	0.028	$\frac{1}{12}$	$n^{(1)}$
Var(l_2)	$(0.228n - 0.045)$	$\frac{2n-1}{6}$	$n^{(2)}$
Cov(l_2, l_3)	$(0.081n + 0.021)$	$\frac{n}{6}$	$n^{(2)}$
Cov(l_2, l_4)	$(0.039n + 0.006)$	$\frac{n}{12}$	$n^{(2)}$
Var(l_3)	$(0.086n^2 - 0.0025n - 0.035)$	$\frac{4n^2-3n-2}{30}$	$n^{(3)}$
Cov(l_3, l_4)	$(0.038n^2 + 0.027n + 0.003)$	$\frac{n^2}{12}$	$n^{(3)}$
Var(l_4)	$(0.043n^3 + 0.014n^2 - 0.008n - 0.033)$	$\frac{3n^3+693n^2-2n-3}{42}$	$n^{(4)}$

Table 2.2: Exact variances and covariances of the first four sample L-moments from Gumbel and exponential distribution, both with $\sigma^2 = 1$.

	Var(l_2)				Var(l_3)			
	Nor		Gum		Nor		Gum	
n	Asy.	Ex.	Asy.	Ex.	Asy.	Ex.	Asy.	Ex.
5	0.032	0.042	0.047	0.055	0.012	0.043	0.019	0.035
8	0.020	0.024	0.029	0.032	0.007	0.015	0.012	0.016
10	0.016	0.018	0.023	0.025	0.006	.0.01	0.009	0.012
15	0.011	0.012	0.015	0.016	0.004	0.005	0.006	0.007
20	0.008	0.008	0.012	0.012	0.003	0.004	0.005	0.005
25	0.007	0.007	0.009	0.009	0.002	0.003	0.004	0.004
35	0.005	0.005	0.007	0.007	0.002	0.002	0.003	0.003
50	0.003	0.003	0.005	0.005	0.001	0.001	0.002	0.002

Table 2.3: Comparison between asymptotic and exact variances for different sample sizes from the Gumbel (Gum) and the Normal (Nor) distributions, both with $\sigma^2 = 1$.

850	960	930	940	880	780	890	830	890	740
940	880	880	1000	810	890	840	850	900	960
880	810	840	760	780	740	1070	850	940	860
820	810	720	840	930	810	880	810	720	800
760	850	950	850	800	800	720	650	770	810
950	850	960	860	620	840	760	880	790	980
880	860	800	810	740	850	810	840	980	900
970	750	810	1000	820	870	880	840	840	950
760	850	870	790	1000	810	830	800	880	910
870	980	880	840	790	760	910	960	920	870

Table 2.4: Michelson's determinations of the velocity of the light made from June 5, 1879 to July 2, 1879.

L-moments with unbiased standard error. Assuming normality, we find from Hosking (1990) that $\lambda_1 = \mu$ and $\lambda_2 = \sigma/\sqrt{\pi}$. Then, the unbiased estimators of μ and σ are

$$\widehat{\mu} = l_1 = 852.4 \quad \text{and} \quad \widehat{\sigma} = \sqrt{\pi} l_2 = 78.5$$

We find from Table 2.1 that

$$\text{Var}\{\widehat{\mu}\} = \frac{\sigma^2}{n} \text{ and } \text{Var}\{\widehat{\sigma}\} = \pi \text{Var}\{l_2\} = \left(\frac{0.51n + 0.12}{n(n-1)}\right)\sigma^2$$

replacing σ by its estimate $\widehat{\sigma} = 78.5$ in these expressions, we find

$$\widehat{\text{Var}}\,(\widehat{\mu}) = 61.6, \quad \widehat{\text{Var}}\,(\widehat{\sigma}) = 31.9 \quad \text{and} \quad \text{Cov}\,(\widehat{\mu}, \widehat{\sigma}) = 0$$

If we use the distribution-free unbiased estimators of $\widehat{\text{Var}}\,(l_1)$, $\widehat{\text{Var}}\,(l_2)$ and $\widehat{\text{Cov}}\,(l_1, l_2)$, we find that

$$\widehat{\text{Var}}\{\widehat{\mu}\} = 62.4, \ \text{Var}\{\widehat{\sigma}\} = \pi \widehat{\text{Var}}\{l_2\} = 36.7 \quad \text{and} \quad \widehat{\text{Cov}}\,(\widehat{\mu}, \widehat{\sigma}) = \sqrt{\pi}\ \widehat{\text{Cov}}\,(l_1, l_2) = 1.26$$

2.6.3 Approximate mean and variance of L-skewness and L-kurtosis

We estimate the skewness and kurtosis measures $\tau_3 = \lambda_3/\lambda_2$ and $\tau_4 = \lambda_4/\lambda_2$ by $t_3 = l_3/l_2$ and $t_4 = l_4/l_2$. To obtain approximate variances for t_3 and t_4, we use the following approximation for the variance of a ratio $R = U/V$ of two random quantities

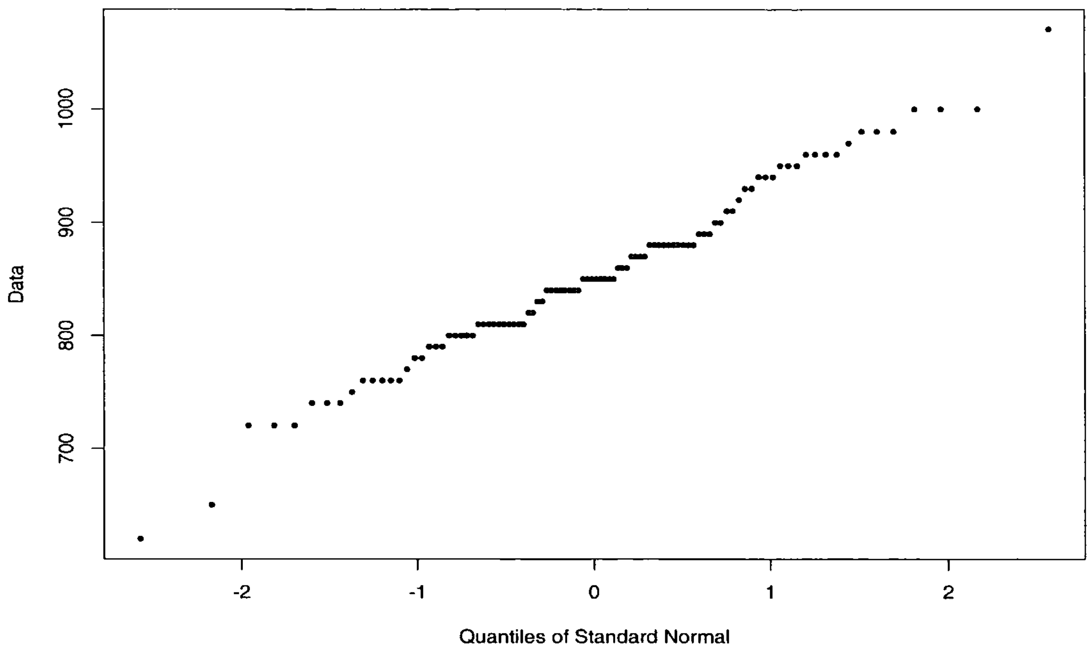

Figure 2.5: Normal probability plot of Michelson's data.

U and V given, for example, by Rice (1995)

$$\mathrm{E}(R) \simeq \frac{\mathrm{E}(U)}{\mathrm{E}(V)} + \frac{\mathrm{E}(U)}{\mathrm{E}^3(V)}\mathrm{Var}(V) - \frac{\mathrm{Cov}(U,V)}{\mathrm{E}^2(V)} \tag{2.55}$$

$$\mathrm{Var}\,(R) \simeq \left\{\mathrm{Var}\,(V)\,\frac{\mathrm{E}^2\,(U)}{\mathrm{E}^4\,(V)} + \frac{\mathrm{Var}\,(U)}{\mathrm{E}^2\,(V)} - \frac{2\mathrm{Cov}\,(U,V)\,\mathrm{E}\,(U)}{\mathrm{E}^3\,(V)}\right\} \tag{2.56}$$

The choices $R = l_3/l_2$ and $R = l_4/l_2$ give approximate expected values and variances for t_3 and t_4 and we give some examples for the distributions which have been defined in example 1.

When sampling is from the normal distribution, $\tau_3 = 0$ and $\tau_4 = 0.123$, we obtain the approximations

$$\begin{aligned} \mathrm{E}\,(t_3) &\simeq 0 \\ \mathrm{Var}\,(t_3) &\simeq \frac{0.1866n^2 + 0.1541n + 0.03261}{n\,(n-1)\,(n-2)} \end{aligned}$$

and

$$\begin{aligned} \mathrm{E}(t_4) &\simeq \frac{0.1227\,(n-0.2184)\,(n-0.5514)}{n\,(n-1)} \\ \mathrm{Var}(t_4) &\simeq \frac{0.086\,(n+1.06)\,(n+0.91)\,(n+0.25)}{n\,(n-1)\,(n-2)\,(n-3)} \end{aligned}$$

For Michelson's data we find that $(t_3, t_4) = (0.019, 0.147)$ with approximate standard errors $(0.044, 0.030)$, respectively, suggesting that the theoretical values $\tau_3 = 0$ and $\tau_4 = 0.123$ for normal distribution are supported by these data at confidence level 0.95.

When sampling is from the uniform distribution, $\tau_3 = 0$ and $\tau_4 = 0$, we obtain the approximations

$$\begin{aligned} \mathrm{E}(t_3) &\simeq 0 \\ \mathrm{Var}(t_3) &\simeq \frac{(12n^2+60)}{70n\,(n-1)\,(n-2)} \end{aligned}$$

and

$$\begin{aligned} \mathrm{E}(t_4) &\simeq \frac{3\,(n+3)}{35n\,(n-1)} \\ \mathrm{Var}(t_4) &\simeq \frac{2(n^2+11)(n+3)}{35n\,(n-1)\,(n-2)\,(n-3)} \end{aligned}$$

When sampling is from the Gumbel distribution, $\tau_3 = 0.1699$ and $\tau_4 = 0.1504$, we obtain the approximations

$$\begin{aligned} \mathrm{E}(t_3) &\simeq \frac{0.1699(n+0.275)(n-2.123)}{n(n-1)} \\ \mathrm{Var}(t_3) &\simeq \frac{0.221(n+0.817)(n-0.344)}{n\,(n-1)\,(n-2)} \end{aligned}$$

and

$$\mathrm{E}(t_4) \simeq \frac{0.150(n+0.228)(n-1.338)}{n\,(n-1)}$$

$$\mathrm{Var}\,(t_4) \simeq \frac{0.126(n-1.022)(n^2+2.209n+1.356)}{n\,(n-1)\,(n-2)\,(n-3)}$$

When sampling is from the exponential distribution, $\tau_3 = 1/3$ and $\tau_4 = 1/6$, we obtain the approximations

$$\mathrm{E}\,(t_3) \simeq \frac{0.331(n+0.210)(n-1.987)}{n(n-1)}$$

$$\mathrm{Var}\,(t_3) \simeq \frac{0.331n^2+0.941n+2.205}{n\,(n-1)\,(n-2)}$$

and

$$\mathrm{E}\,(t_4) \simeq \frac{0.167(n+1.606)(n-2.410)}{n\,(n-1)}$$

$$\mathrm{Var}\,(t_4) \simeq \frac{(80n^3+7081n^2-205n-150)/378}{378n\,(n-1)\,(n-2)\,(n-3)}$$

2.7 Characterisation of the normal distribution based on sample L-moments

In most cases of statistical evaluation, decisions based on a set of observations on a random variable Y depend on the assumption that the distribution function $F(y)$ of Y is known. Knowing $F(y)$ may mean to the applied scientist just a subjective choice that may be supported by the data by using some empirical method such as a quantile-quantile plot or a goodness of fit test.

Many contributions have been made to the problem of characterising the normal distribution such as the independence of the sample mean and sample variance; see, for example, Lin and Mudholker (1980). In this section, we use the covariance structure of sample L-moments to characterise the normal distribution in the class of distributions having finite second moments.

We use the results from Theorem 2.1 below due to Govindarajulu (1966b), which characterises the normal distribution in terms of second moments of the order statistics $Y_{1:n}, Y_{2:n}, ..., Y_{n:n}$.

Theorem 2.1 *Let $Y_1, Y_2, ...$ be a sequence of independent and identically distributed nontrivial random variables with zero mean, unit variance and distribution function $F(y)$. Then, for $i = 1, 2, ..., n$*

$$\sum_{j=1}^{n} E\left(Y_{i:n}Y_{j:n}\right) = 1, \qquad n = 2, 3, \ldots \tag{2.57}$$

if and only if $F(y) = \Phi(y) = \int_{-\infty}^{y} (2\pi)^{-1/2} e^{-t^2/2}dt$, *for* $y \in (-\infty, \infty)$.

Theorem 2.2 *Let* $Y_1, Y_2, \ldots$ *be a sequence of independent and identically distributed nontrivial random variables with zero mean, unit variance with distribution function* $F(y)$*, and for each* $n = 2, 3, \ldots$ *let* $l_1, \ldots, l_n$ *denote the sample L-moments based on* $Y_{1:n} \leq Y_{2:n} \leq \ldots \leq Y_{n:n}$*. Then, for* $i = 1, 2, \ldots, n$ *and* $n = 2, 3, \ldots$

$$Cov\left(l_1, l_r\right) = E\left(l_1 l_r\right) = \begin{cases} \frac{1}{n} & if \quad r = 1 \\ 0 & if \quad r > 1 \end{cases} \tag{2.58}$$

if and only if $F(y) = \Phi(y) = \int_{-\infty}^{y} (2\pi)^{-1/2} e^{-t^2/2}dt$, *for* $y \in (-\infty, \infty)$.

Proof.

First note that (2.57) in Theorem 2.1 can be written

$$\mathrm{E}\left(l_1 Y_{i:n}\right) = \frac{1}{n}, \quad i = 1, \ldots, n \tag{2.59}$$

Using the expression for l_r in (2.15) and (2.17), we may write

$$\mathrm{E}\left(l_1 l_r\right) = \sum_{k=0}^{r-1} \frac{c_{r-1,k}}{n^{(k+1)}} \sum_{i=1}^{n} (i-1)^{(k)} \, \mathrm{E}\left(l_1 Y_{i:n}\right)$$

Thus, when the distribution is normal, we have from (2.59) that

$$\mathrm{E}\left(l_1 l_r\right) = \frac{1}{n} \sum_{k=0}^{r-1} \frac{c_{r-1,k}}{n^{(k+1)}} \sum_{i=1}^{n} (i-1)^{(k)} = \begin{cases} \frac{1}{n} & \text{if} \quad r = 1 \\ 0 & \text{if} \quad r > 1 \end{cases} \tag{2.60}$$

which proves the first part, after noting from Downton (1966) that

$$\sum_{i=1}^{n} (i-1)^{(k)} = \frac{n^{(k+1)}}{k+1}, \tag{2.61}$$

and the identity

$$\sum_{k=0}^{r-1} \frac{c_{r-1,k}}{k+1} = \begin{cases} 1 & \text{if} \quad r = 1 \\ 0 & \text{if} \quad r > 1 \end{cases} \tag{2.62}$$

which follows by noting that

$$r\binom{r-1}{k} = (k+1)\binom{r}{k+1}, \quad \binom{r-1+k}{k} = (-1)^k \binom{-r}{k}$$

and

$$\sum_{j=0}^{n} \binom{a}{j} \binom{b}{n-j} = \binom{a+b}{n}$$

with $n = r - 1$, $j = k$, $a = -r$ and $b = r$.

When (2.58) holds

$$\mathrm{E}\left(l_1 l_r\right) = \sum_{k=0}^{r-1} \frac{c_{r-1,k}}{n^{(k+1)}} \sum_{i=1}^{n} (i-1)^{(k)} \, \mathrm{E}\left(l_1 Y_{i:n}\right) = \begin{cases} \frac{1}{n} & \text{if} \quad r = 1 \\ 0 & \text{if} \quad r = 2, 3, ..., n \end{cases} \tag{2.63}$$

which we can write it as

$$\mathbf{Az} = \mathbf{c} \tag{2.64}$$

where

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots. & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}, \quad \mathbf{z} = \begin{pmatrix} \mathrm{E}\left(l_1 Y_{1:n}\right) \\ \mathrm{E}\left(l_1 Y_{2:n}\right) \\ \vdots \\ \mathrm{E}\left(l_1 Y_{n:n}\right) \end{pmatrix} \quad \text{and} \quad \mathbf{c} = \begin{pmatrix} 1/n \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Because of the non singularity of $\mathbf{A}$ (Sillito, 1969), we find that the unique solution of (2.64) is $\mathbf{z}^T = (1/n, 1/n, ..., 1/n)$ which follows from (2.60). Since this is true for all $n = 2, 3, ...$ we have by Theorem 2.1 that the distribution is normal. This completes the proof.

The non-singularity of $\mathbf{A}$ follows from the following properties which are given by Sillito (1969)

(1) $\mathbf{A}$ can be written as follows

$$\mathbf{A} = \mathbf{LU} \tag{2.65}$$

where $\mathbf{L}$ is a lower triangular matrix in which the (r, c)th element is

$$(-1)^{r+c} \binom{r-1}{c-1} \binom{r+c-2}{r-1} \quad \text{for} \quad r \geq c$$

and $\mathbf{U}$ is an upper triangular matrix in which the (r, c)th element is

$$\binom{c-1}{r-1}\Big/ r\binom{n}{r} \quad \text{for} \quad c \geq r$$

(2) **A** is non-singular matrix with determinant

$$\frac{0!2!4!...(2n-2)!}{(n!)^n} \tag{2.66}$$

For example, when $n = 3$, we have

$$\mathbf{L} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 2 & 0 \\ 1 & -6 & 6 \end{pmatrix}, \qquad \mathbf{U} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{6} & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} \end{pmatrix}$$

and

$$\mathbf{A} = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ -1/3 & 0 & 1/3 \\ 1/3 & -2/3 & 1/3 \end{pmatrix}$$

Thus, we find that

$$\begin{pmatrix} \mathrm{E}\left(l_1^2\right) \\ \mathrm{E}\left(l_1 l_2\right) \\ \mathrm{E}\left(l_1 l_3\right) \end{pmatrix} = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ -1/3 & 0 & 1/3 \\ 1/3 & -2/3 & 1/3 \end{pmatrix} \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix} = \begin{pmatrix} 1/3 \\ 0 \\ 0 \end{pmatrix}$$

and

$$\begin{pmatrix} \mathrm{E}\left(l_1 Y_{1:3}\right) \\ \mathrm{E}\left(l_1 Y_{2:3}\right) \\ \mathrm{E}\left(l_1 Y_{3:3}\right) \end{pmatrix} = \begin{pmatrix} 1 & 1 & 3 \\ 1 & 0 & 4 \\ 1 & 4 & 3 \end{pmatrix} \begin{pmatrix} 1/3 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix}$$

Because the correlation coefficient is location and scale free, we can reformulate Theorem 2.2 as follows

Corollary 2.1

$$Corr\left(l_1, l_r\right) = \begin{cases} 1 & \text{if} \quad r = 1 \\ 0 & \text{if} \quad r = 2, 3, ..., n \end{cases} \tag{2.67}$$

for all $n = 2, 3, ...$ *if and only if* $F(y)$ *is normal.*

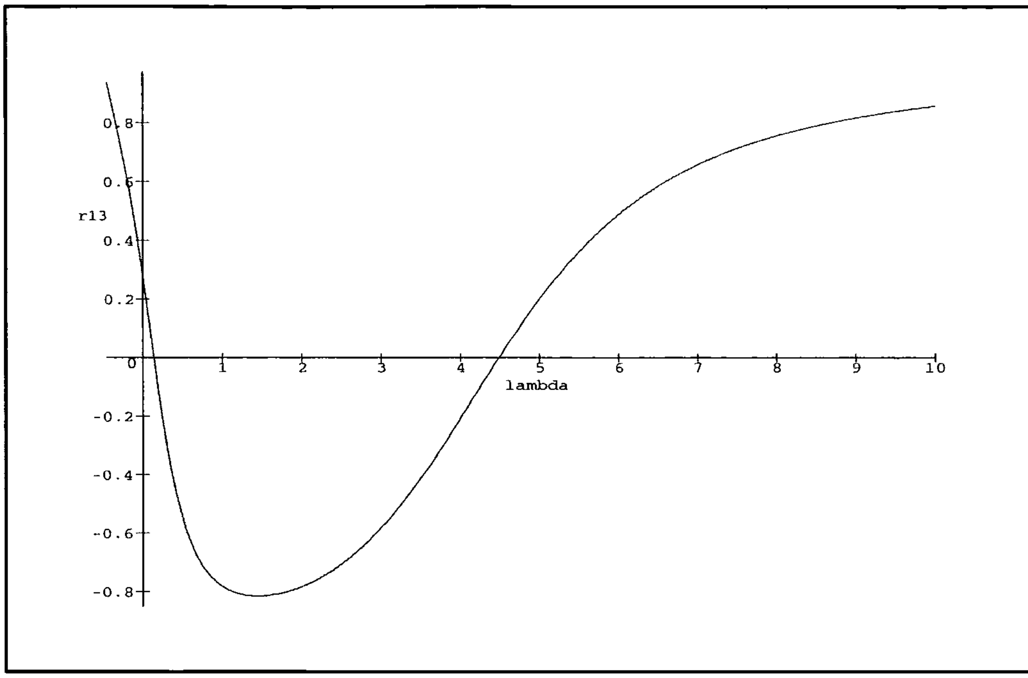

Figure 2.6: Correlation between l_1 and l_3 for the one-parameter symmetric lambda distribution.

We see from Figure 2.6, $\mathrm{Corr}(l_1, l_3) = 0$ for two values of λ (0.144 and 4.452) for the symmetric lambda distribution, which the distribution is a good approximation to the N $(0, 1)$ distribution; see Chapter 4, for details about the lambda distribution.

When $\mathrm{E}\,(Y) = 0$, from (2.39), we can write $\mathrm{E}\,(l_1 l_r)$ as

$$\mathrm{E}\,(l_1 l_r) = \frac{1}{n}\sum_{k=0}^{r-1} \frac{c_{r-1,k}}{(k+1)} \left[\mathrm{E}\left(Y_{k+1:k+1}^2\right) - \mathrm{E}\left(Y_{k:k+1} Y_{k+1:k+1}\right)\right] \tag{2.68}$$

where $c_{r,k}$ is given in (2.13).

Theorem 2.3 *A distribution-free unbiased estimator of* $E\,(l_1 l_r)$ *is given by*

$$\widehat{E}\,(l_1 l_r) = \frac{1}{n}\sum_{k=0}^{r-1} \frac{c_{r-1,k}}{(k+1)\binom{n}{k+1}} \left[\sum_{i=1}^{n} \binom{i-1}{k} Y_{i:n}^2 - \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \binom{i-1}{k-1} Y_{i:n} Y_{j:n}\right] \tag{2.69}$$

Proof:

The proof is straightforward from equations (2.27) and (2.28).

2.8 Probability-weighted moments

Probability-weighted moments (PWM) were defined by Greenwood et al. (1979) as follows

$$M_{p,r,s} = E\left\{X^p\left[F(X)\right]^r\left[1-F(X)\right]^s\right\}, \tag{2.70}$$

where p, r and s are real numbers.

The definition of $M_{p,r,s}$ is valid for both continuous and discrete random variables and we are interested in former case.

When the random variable X is continuous, we can write (2.70) as follows

$$M_{p,r,s} = \int x^p F(x)^r \left[1-F(x)\right]^s dF(x) \tag{2.71}$$

Probability-weighted moments are likely to be most useful when the inverse distribution function $Q(u)$ can be written in closed form, which is the case for the lambda distribution (see Chapter 4). We may write

$$M_{p,r,s} = \int_0^1 \left[Q(u)\right]^p u^r \left[1-u\right]^s du \tag{2.72}$$

The quantities $M_{p,0,0}$ $(p = 1, 2, ...)$ are non-central moments of X. The moments $M_{1,r,s}$ may be preferable for estimating the parameters of the distribution of X, in the sense that the relationship between parameters and moments often takes a simpler form; in this case, linear combinations of order statistics. Two cases of particular interest, because the estimators will be in linear form, are

$$\alpha_s = M_{1,0,s} = \frac{1}{(s+1)}\mathrm{E}\left(X_{1:s+1}\right) = \mathrm{E}\left\{X\left[1-F(X)\right]^s\right\}, \qquad s = 0,1,2,... \tag{2.73}$$

and

$$\beta_r = M_{1,r,0} = \frac{1}{(r+1)}\mathrm{E}\left(X_{r+1:r+1}\right) = \mathrm{E}\left\{X\left[F(X)\right]^r\right\}, \qquad r = 0,1,2,... \tag{2.74}$$

Given a random sample of size n from a distribution F, estimation of β_r and α_s is most conveniently based on the statistics

$$a_s = n^{-1}\sum_{j=1}^{n}\frac{(n-j)(n-j-1)...(n-j-s+1)}{(n-1)(n-2)...(n-s)}X_{j:n}, \qquad s = 0,1,2,... \tag{2.75}$$

and

$$b_r = n^{-1} \sum_{j=1}^{n} \frac{(j-1)(j-2)...(j-r)}{(n-1)(n-2)...(n-r)} X_{j:n} \qquad r = 0, 1, 2, ... \tag{2.76}$$

which are unbiased estimator of β_r and α_s, respectively; see, Landweher et al. (1979) .

The quantities $M_{p,r.s}$ may be used to describe and characterise probability distributions. The characterisations of a distribution by the β_r and by the α_s are interchangeable, because the β_r and α_s are functions of each other; we have in general, see for example Hosking (1986),

$$\alpha_r = \sum_{k=0}^{r} (-1)^k \binom{r}{k} \beta_k, \quad \beta_r = \sum_{k=0}^{r} (-1)^k \binom{r}{k} \alpha_k \tag{2.77}$$

and in particular,

$$\begin{aligned} \alpha_0 &= \beta_0 \\ \alpha_1 &= \beta_0 - \beta_1 \\ \alpha_2 &= \beta_0 - 2\beta_1 + \beta_2 \\ \alpha_3 &= \beta_0 - 3\beta_1 + 3\beta_2 - \beta_3 \end{aligned}$$

and

$$\begin{aligned} \beta_0 &= \alpha_0 \\ \beta_1 &= \alpha_0 - \alpha_1 \\ \beta_2 &= \alpha_0 - 2\alpha_1 + \alpha_2 \\ \beta_3 &= \alpha_0 - 3\alpha_1 + 3\alpha_2 - \alpha_3 \end{aligned}$$

Moreover, we can express the L-moment λ_r linearly in terms of β_r and α_r as follows

$$\lambda_{r+1} = (-1)^r \sum_{k=0}^{r} c_{r,k} \alpha_k = \sum_{k=0}^{r} c_{r,k} \beta_k, \quad r = 0, 1, ..., \tag{2.78}$$

where $c_{r,k}$ is given in (2.13).

In particular

$$\begin{aligned} \lambda_1 &= \alpha_0 && = \beta_0 \\ \lambda_2 &= \alpha_0 - 2\alpha_1 && = 2\beta_1 - \beta_0 \end{aligned}$$

$$\begin{aligned}\lambda_3 = &\quad \alpha_0 - 6\alpha_1 + 6\alpha_2 &&= 6\beta_2 - 6\beta_1 + \beta_0 \\ \lambda_4 = &\quad \alpha_0 - 12\alpha_1 + 30\alpha_2 - 20\alpha_3 &&= 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0\end{aligned}$$

Another proof of the unbiasedness of b_k and a_s which is simpler than that is given in Landweher et al. (1979), follows from equation (2.29); namely

$$\begin{aligned}\mathrm{E}\{b_r\} &= \frac{\mathrm{E}\left(X_{r+1:r+1}\right)}{(r+1)} \\ &= \int_0^1 Q(u)u^r dF = \beta_r\end{aligned}$$

Similarly, a_s is an unbiased estimator of α_s.

2.8.1 Exact variances and covariances of sample probability weighted moments

The exact variances and covariances of b_k are given in (2.37)

$$\text{Cov}\left(b_k, b_l\right) = \theta_{kl} = \frac{1}{n^{(k+1)}} \sum_{g=0}^{k} A_{kl}^{(g)} (n - l - 1)^{(g)} \tag{2.79}$$

It is also possible to determine the first four moments of a_s using the relationship between b_r and a_s, established by Hosking (1986); namely, which is given by,

$$a_s = \sum_{k=0}^{s} (-1)^k \begin{pmatrix} s \\ k \end{pmatrix} b_k$$

It is also possible to determine the first four moments of b_s using the following relationships between l_r and b_s

$$\begin{aligned}b_0 &= l_1 \\ b_1 &= \left(l_2 + l_1\right)/2 \\ b_2 &= \left(l_3 + 3l_2 + 2l_1\right)/6 \\ b_3 &= \left(l_4 + 5l_3 + 9l_2 + 5l_1\right)/20\end{aligned}$$

For example,

$$\text{Var}(b_1) = \frac{1}{4}\text{Var}(l_2) + \frac{1}{4}\text{Var}(l_1) + \frac{1}{2}\text{Cov}(l_1, l_2)$$

2.9 Conclusions

In this chapter, we have defined population and sample L-moments and we have derived the exact variances and covariances of sample L-moment in terms of moments of the order statistics from small samples. Also, we characterise the normal distribution in terms of covariances between l_1 and l_r.

We have investigated the asymptotic and exact variances of the sample L-moments from the Normal and Gumbel distributions and conclude that asymptotic variances have essentially the same values as the exact variances for large samples ($n > 25$), but different values when the sample size is small ($n < 25$). Also, we have derived distribution-free unbiased estimators of the variances and covariances of sample L-moments. We have also discussed probability weighted moments and their relation to L-moments and have obtained their exact variances and covariances. These results are also illustrated in two examples.

Chapter 3

Generalisations of L-moments

3.1 Introduction

Consider the problem of estimating the parameters of a distribution F. Classical estimation methods (e.g, the method of moments, least squares, and maximum likelihood) work well, for example, in cases where the distribution belongs to the exponential family. However, it is recognised that outliers, which arise from heavy-tailed distributions or gross errors of measurement, have undue influence on such methods; for example, $\overline{X}$ which is an unbiased estimator of the mean μ of the normal distribution based on the method of moments, least squares and maximum likelihood, is a non-robust estimator; see, for example, Ali and Luceno (1997). Therefore, if there is concern about extreme observations which having undue influence, one should use a robust method of estimation which has been developed to reduce the influence of outliers on the final estimates. In recent years, a great deal of attention has been focused on robust estimation methods; methods produce estimates that are resistant to the presence of outliers; see, for example, Barnett and Lewis (1994), Hampel et al. (1986), Hawkins (1980) and Rousseuw and Leroy (1987).

As discussed in Chapter 2, Hosking (1990) unified analysis and estimation of distributions using linear combinations of order statistics and used their ratios as new measures of skewness and kurtosis to relate L-moments to the method of moments. Royston (1992) and Vogel and Fennessey (1993) discuss the advantages of L-skewness and L-kurtosis over their product-moment counterparts. Hosking and Wallis (1995), Sillito (1951) and Sillito (1969) consider various theoretical aspects and applications of L-moments. Mudholkar and Hutson (1998) introduced LQ-moments using a "quick" measure of the location of the sampling distribution of the order statistics such as the median, the tri-mean and Gastwirth measure (which we call Gastwirth) in place of the mean. There are wide applications for L-moments in engineering, meteorology, and hydrology; see, for example; Gingras and Adamowski (1994), Guttman et al. (1993), Pearson (1993), Pilon and Adamowski (1992) and Sankarasubramanian and Srinivasan (1999).

In this chapter, see also Elamir and Seheult (2001c), an alternative approach which we call trimmed L-moments (TL-moments) is introduced which gives zero weight to extreme observations. TL-moments have advantages over L-moments and the method of moments: they exist whether or not the mean exists (for example, the Cauchy distribution) and they are more robust to the presence of outliers. Trimming refers to the removal of extreme values of a sample. For example, to symmetrically trim a univariate sample size, one removes the k smallest and k largest values for some specified $k < n/2$. For univariate samples the trimmed mean, the mean of the $n-2k$ un-trimmed sample values, is by far the most widely studied trimmed statistic.

In section 3.2 we define LQ-moments and obtain their large sample variances. In section 3.3 we introduce both population trimmed L-moments and their sample counterparts for estimating parameters from any univariate continuous distribution and also obtain their exact variances and covariances. In section 3.4 we develop the trimmed probability weighted moment method (TPWM) and elucidate its relation to TL-moments. In section 3.5 we study the TL-mean as a robust location estimator and apply the method of TL-moments to some symmetric distributions.

3.2 LQ-moments

Mudholkar and Hutson (1998) defined the rth population LQ-moment ζ_r as

$$\zeta_r = r^{-1}\sum_{k=0}^{r-1}(-1)^k \begin{pmatrix} r-1 \\ k \end{pmatrix} \tau_{p,\alpha}\left(X_{r-k:r}\right), \qquad r = 1, 2, 3, ... \tag{3.1}$$

where

$$\tau_{p,\alpha}\left(X_{r-k:r}\right) = pQ_{X_{r-k:r}}(\alpha) + (1-2p)Q_{X_{r-k:r}}(1/2) + pQ_{X_{r-k:r}}(1-\alpha), \tag{3.2}$$

$0 \le \alpha \le 1/2, \quad 0 \le p \le 1/2$ and Q_X denotes the quantile function of a random variable X. Notice that LQ-moments reduce to L-moments when $\tau_{p,\alpha}\left(X_{r-k:r}\right) = \mathrm{E}\left(X_{r-k:r}\right)$.

Possible candidates for $\tau_{p,\alpha}$ are: the median $(p = 0,\ \alpha = 1)$, the tri-mean $\left(p = \frac{1}{3},\ \alpha = \frac{1}{3}\right)$ and Gastwirth $\left(p = 0.3,\ \alpha = \frac{1}{3}\right)$.

The first four LQ-moments are

$$\begin{aligned}
\zeta_1 &= \tau_{p,\alpha}(X_{1:1}) \\
\zeta_2 &= \frac{1}{2}\left[\tau_{p,\alpha}(X_{2:2}) - \tau_{p,\alpha}(X_{1:2})\right] \\
\zeta_3 &= \frac{1}{3}\left[\tau_{p,\alpha}(X_{3:3}) - 2\tau_{p,\alpha}(X_{2:3}) + \tau_{p,\alpha}(X_{1:3})\right] \\
\zeta_4 &= \frac{1}{4}\left[\tau_{p,\alpha}(X_{4:4}) - 3\tau_{p,\alpha}(X_{3:4}) + 3\tau_{p,\alpha}(X_{2:4}) - \tau_{p,\alpha}(X_{1:4})\right]
\end{aligned}$$

these equations have the same form as those for L-moments given in Chapter 2.

Evaluation of LQ-moments for any continuous distribution is simplified using the following relationship

$$\tau_{p,\alpha}(X_{r-k:r}) = pQ_X\left[B^{-1}_{r-k:r}(\alpha)\right] + (1-2p)Q_X\left[B^{-1}_{r-k:r}\left(\frac{1}{2}\right)\right] + pQ_X\left[B^{-1}_{r-k:r}(1-\alpha)\right] \tag{3.3}$$

where $B^{-1}_{r-k:r}(1-\alpha)$ denotes the corresponding α-quantile of a beta distribution with parameters $r-k$ and $k+1$.

Mudholkar and Hutson (1998) estimate population LQ-moments by the following sample LQ-moments

$$\widehat{\zeta}_r = r^{-1}\sum_{k=0}^{r-1}(-1)^k \binom{r-1}{k} \widehat{\tau}_{p,\alpha}(X_{r-k:r}), \tag{3.4}$$

where they take the estimator $\widehat{\tau}_{p,\alpha}(X_{r-k:r})$ of the location of the order statistics $X_{r-k:r}$ to be

$$\widehat{\tau}_{p,\alpha}(X_{r-k:r}) = p\widehat{Q}_X\left[B^{-1}_{r-k:r}(\alpha)\right] + (1-2p)\widehat{Q}_X\left[B^{-1}_{r-k:r}\left(\frac{1}{2}\right)\right] + p\widehat{Q}_X\left[B^{-1}_{r-k:r}(1-\alpha)\right] \tag{3.5}$$

and $\widehat{Q}_X(.)$ denotes the linear interpolation estimator of $Q_X(.)$ given by

$$\widehat{Q}_X(u) = (1-\epsilon)\,X_{[(n+1)u]:n} + \epsilon X_{[(n+1)u]+1:n} \tag{3.6}$$

with $\epsilon = (n+1)\,u - [(n+1)\,u]$.

Sample LQ-moments depend upon the choice of $\tau_{p,\alpha}(.)$ and the quantile estimator used for estimating it. However, their asymptotic normality follows from the large sample theory of linear functions of order statistics. The asymptotic mean of $\widehat{\zeta}_r$ is ζ_r

	Median	Tri-mean	Gastwirth
ζ_1	μ	μ	μ
ζ_2	0.539σ	0.544σ	0.542σ
ζ_3	0	0	0
ζ_4	0.062σ	0.04σ	0.063σ
η_3	0	0	0
η_4	0.116	0.118	0.117

Table 3.1: Population LQ-moments $(\zeta_1, \zeta_2, \zeta_3, \zeta_4, \eta_3, \eta_4)$ for the normal distribution using (a) median (b) tri-mean (c) Gastwirth

and their asymptotic covariances are given by Mudholkar and Hutson (1998) as follows

$$\begin{aligned} \text{Cov}\left(\widehat{\zeta}_r, \widehat{\zeta}_s\right) &= \frac{1}{rs}\sum_{k=0}^{r-1}\sum_{l=0}^{s-1}(-1)^{k+l}\binom{r-1}{k}\binom{s-1}{l} \\ &\times \text{Cov}\left(\widehat{\tau}_{p,\alpha}\left(X_{r-k:r}\right), \widehat{\tau}_{p,\alpha}\left(X_{s-l:s}\right)\right) \end{aligned} \quad (3.7)$$

where

$$\begin{aligned} &\text{Cov}\left[\widehat{\tau}_{p,\alpha}\left(X_{r-k:r}\right), \widehat{\tau}_{p,\alpha}\left(X_{s-l:s}\right)\right] = p\left\{p\,\text{Cov}\left[\widehat{Q}\left(u_1\right), \widehat{Q}\left(u_2\right)\right]\right. \\ &+p\,\text{Cov}\left[\widehat{Q}\left(u_1\right), \widehat{Q}\left(u_6\right)\right] + p\,\text{Cov}\left[\widehat{Q}\left(u_4\right), \widehat{Q}\left(u_5\right)\right] \\ &+p\,\text{Cov}\left[\widehat{Q}\left(u_4\right), \widehat{Q}\left(u_5\right)\right] + p\,\text{Cov}\left[\widehat{Q}\left(u_2\right), \widehat{Q}\left(u_5\right)\right] \\ &+(1-2p)\,\text{Cov}\left[\widehat{Q}\left(u_3\right), \widehat{Q}\left(u_4\right)\right] + (1-2p)\,\text{Cov}\left[\widehat{Q}\left(u_3\right), \widehat{Q}\left(u_6\right)\right] \\ &\left.+(1-2p)\,\text{Cov}\left[\widehat{Q}\left(u_2\right), \widehat{Q}\left(u_3\right)\right] + (1-2p)\,\text{Cov}\left[\widehat{Q}\left(u_1\right), \widehat{Q}\left(u_4\right)\right]\right\} \end{aligned} \quad (3.8)$$

and

$$\text{Cov}\left[\widehat{Q}\left(u_i\right), \widehat{Q}\left(u_j\right)\right] = u_i\left(1-u_j\right)Q'\left(u_i\right)Q'\left(u_j\right)/n \quad \text{for } i \le j \quad (3.9)$$

note that $\widehat{Q}$ is short hand for $\widehat{Q}_X$.

As an example, Table 3.1 gives the population values of ζ_1, ζ_2, ζ_3 and ζ_4 and their LQ-skewness and LQ-kurtosis (η_3,η_4), based on the median, the tri-mean and Gastwirth functional for a normal distribution.

3.3 Trimmed L-moments (TL-moments)

In this section we generalise L-moments to TL-moments and show that L-moments are a special case of TL-moments. Sample TL-moments which can be expressed as linear combinations of the sample order statistics are unbiased for the corresponding population quantities. We derive expressions for the exact variances and covariances for sample TL-moments in terms of first and second-moments of the order statistics from small samples.

3.3.1 Population TL-moments

We define TL-moments analogously to L-moments (given in Chapter 2) as follows

Definition 3.1 *Let X be a real-valued random variables with distribution function F. We define the rth TL-moment as*

$$\lambda_r^{(t_1 t_2)} = r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} E\left(X_{r+t_1-k:r+t_1+t_2}\right), \quad r = 1, 2, \ldots \tag{3.10}$$

where t_1 and t_2 are the amounts of lower and upper trimming.

We study the special case when $t_1 = t_2 = t$, and write (3.10) as

$$\lambda_r^{(t)} = r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} \mathrm{E}\left(X_{r+t-k:r+2t}\right), \quad r = 1, 2, \ldots \tag{3.11}$$

Clearly TL-moments reduce to L-moments when $t = 0$; in particular, $\lambda_1^{(0)} = \mathrm{E}(X)$. Moreover, in view of (3.11), $\lambda_r^{(t)}$ is evaluated in terms of expectations of order statistics from a sample of size $r + 2t$. For example,

$$\lambda_1^{(t)} = \mathrm{E}\left[X_{1+t:1+2t}\right]$$

which is the population mean of the sample median from a sample of size $1+2t$, which will be zero if the distribution is symmetric about the origin.

While the definition of $\lambda_r^{(t)}$ is valid for both continuous and discrete random variables, we restrict attention to continuous random variables.

As we have seen previously, the expectation of an order statistic may be written as (David (1981))

$$\mathrm{E}\left(X_{i:r}\right) = \frac{r!}{(i-1)!(r-i)!} \int_0^1 Q(u) u^{i-1} (1-u)^{r-i} \, du \tag{3.12}$$

where Q is the quantile function Q_X.

Substituting this expression in (3.11), we have

$$\begin{aligned}\lambda_r^{(t)} &= r^{-1}\sum_{k=0}^{r-1}(-1)^k\begin{pmatrix} r-1 \\ k\end{pmatrix}\frac{(r+2t)!}{(r+t-k-1)!\,(t+k)!} \\ &\times\int_0^1 Q(u)u^{r+t-k-1}(1-u)^{t+k}\,du, \quad r=1,2,... \end{aligned} \tag{3.13}$$

We now examine in more detail the expression for $\lambda_r^{(t)}$ when $t = 0, 1, 2$. When $t = 0$, equation (3.13) gives the ordinary L-moments, described in Chapter 2. When $t = 1$, the first four TL-moments are

$$\begin{aligned}
\lambda_1^{(1)} &= \mathrm{E}\left(X_{2:3}\right) = 6\int_0^1 Q(u)u\,(1-u)\,du \\
\lambda_2^{(1)} &= \frac{1}{2}\mathrm{E}\left(X_{3:4}-X_{2:4}\right) = 6\int_0^1 Q(u)u\,(1-u)\,(2u-1)\,du \\
\lambda_3^{(1)} &= \frac{1}{3}\mathrm{E}\left(X_{4:5}-2X_{3:5}+X_{2:5}\right) = \frac{20}{3}\int_0^1 Q(u)u\,(1-u)\left(5u^2-5u+1\right)dF \\
\lambda_4^{(1)} &= \frac{1}{4}\mathrm{E}\left(X_{5:6}-3X_{4:6}+3X_{3:6}-X_{2:6}\right) = \frac{15}{2}\int_0^1 Q(u)u\,(1-u) \\
&\times\left(14u^3-21u^2+9u-1\right)du
\end{aligned}$$

When $t = 2$, the first four TL-moments are

$$\begin{aligned}
\lambda_1^{(2)} &= \mathrm{E}\left(X_{3:5}\right) = 30\int_0^1 Q(u)u^2\,(1-u)^2\,du \\
\lambda_2^{(2)} &= \frac{1}{2}\mathrm{E}\left(X_{4:6}-X_{3:6}\right) = 30\int_0^1 Q(u)u^2\,(1-u)^2\,(2u-1)\,du \\
\lambda_3^{(2)} &= \frac{1}{3}\mathrm{E}\left(X_{5:7}-2X_{4:7}+X_{3:7}\right) = \frac{35}{3}\int_0^1 Q\left(u\right)u^2\,(1-u)^2\left(5u^2-5u+1\right)du \\
\lambda_4^{(2)} &= \frac{1}{4}\mathrm{E}\left(X_{6:8}-3X_{5:8}+3X_{4:8}-X_{3:8}\right) = \frac{15}{2}\int_0^1 Q\left(u\right)u^2\,(1-u)^2 \\
&\times\left(14u^3-21u^2+9u-1\right)du
\end{aligned}$$

Note that as $t \to \infty$, $\lambda_1^{(t)}$ converges to the population median. Thus, a distribution may be specified by its trimmed L-moments even if some of its L-moments and conventional moments do not exist; for example, the Cauchy distribution.

We can standardise the higher moments $\lambda_r^{(t)}$, $r \geq 3$, so that they are independent of the units of the measurement of X as follows

$$\begin{aligned} \tau_r^{(t)} &= \frac{\lambda_r^{(t)}}{\lambda_2^{(t)}}, \\ \tau^{(t)} &= \frac{\lambda_2^{(t)}}{\lambda_1^{(t)}} \end{aligned} \tag{3.14}$$

for $r = 3, 4, ...$ and $\tau^{(t)}$ is called TL- coefficient of variation.

3.3.2 TL- skewness and TL-kurtosis

Pearson (1895) proposed $(\mu - M)/\sigma$ as a measure of skewness for a univariate distribution with mean μ, mode M and variance σ^2. Three other measures of skewness have been introduced by Bowley (1937) and Yule (1944). These are $(\mu - m)/\sigma$, where m is the median, $\beta_1 = \mu_3/\sigma^3$, where μ_3 is the third central moment, and $(q_u + q_l - 2m)/(q_u - q_l)$, where q_u and q_l are the upper and lower quartiles, respectively. All are based on the criteria that a skewness measure should be scale-free and zero for symmetric distributions. The coefficient of skewness β_1 appears more prominently as the measure of skewness, as illustrated by Doodson (1917). The most common measure of kurtosis is $\beta_2 = \mu_4/\mu_2^2$ where μ_4 is the fourth central moment.

Both classical measures β_1 and β_2 have played an important role in classifying distributions, in model fitting and in parameter estimation. Because of some deficiencies of β_1 and β_2, such as their sensitivity to the extreme tails of a distribution and their nonexistence for some distribution, such as the Cauchy distribution, alternative measures of skewness and kurtosis have been proposed. For an overview and discussion of coefficients of skewness and kurtosis; see, for example, Hinkley (1975), Hogg (1974), Hogg et al. (1975), Groeneveld (1998) and Kendall and Stuart (1987).

Hosking (1990) has defined the ratios τ_3 and τ_4 called L-skewness and L-kurtosis as

$$\tau_3 = \frac{\lambda_3}{\lambda_2} \quad \text{and} \quad \tau_4 = \frac{\lambda_4}{\lambda_2} \tag{3.15}$$

where $-1 < \tau_3 < 1$ and $(5\tau_3^2 - 1)/4 \leq \tau_4 < 1$.

Mudholkar and Hutson (1998) have defined the ratios η_3 and η_4 called LQ-skewness and LQ-kurtosis as

$$\eta_3 = \frac{\zeta_3}{\zeta_2} \quad \text{and} \quad \eta_4 = \frac{\zeta_4}{\zeta_2} \tag{3.16}$$

We define the TL-skewness and TL-kurtosis analogously

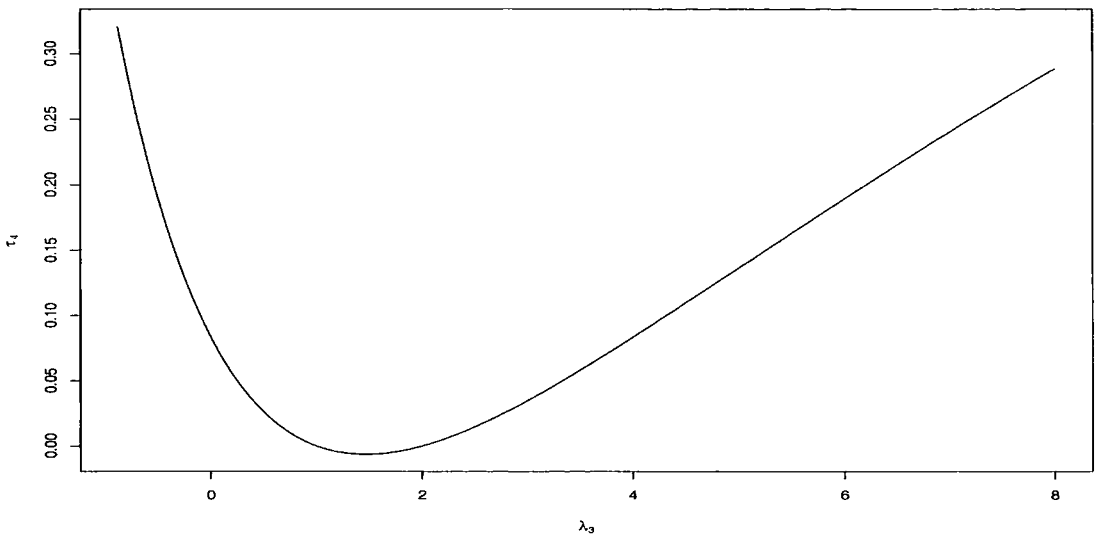

Figure 3.1: Plot of $\tau_4^{(1)}$ as a function of λ for the symmetric lambda distribution.

Definition 3.2 *TL-skewness and TL-kurtosis measures $\tau_3^{(t)}$ and $\tau_4^{(t)}$ are defined as*

$$\tau_3^{(t)} = \frac{\lambda_3^{(t)}}{\lambda_2^{(t)}}, \qquad \tau_4^{(t)} = \frac{\lambda_4^{(t)}}{\lambda_2^{(t)}} \tag{3.17}$$

It should be noted that TL-skewness and TL-kurtosis are location and scale invariant, and exist for the Cauchy distribution ($t > 2$). When the distribution is symmetric TL-skewness is zero and the magnitude of TL-kurtosis increases for the distributions with heavier tails, as illustrated in Figure 3.1 for Tukey's lambda distribution (see Chapter 4) which has quantile function

$$Q(u) = \frac{u^\lambda - (1-u)^\lambda}{\lambda}$$

3.3.3 Sample TL-moments

TL-moments can be estimated in a straightforward manner by estimating the expected value of order statistics given in equation (3.11) from Downton (1966). We define the rth sample TL-moment to be

Definition 3.3 *For a sample of size n, the* rth *sample TL-moment is given by*

$$l_r^{(t_1 t_2)} = r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} \widehat{E}\left(X_{r+t_1-k:r+t_1+t_2}\right), \quad r = 1, 2, ...n-t_1-t_2 \quad (3.18)$$

where

$$\widehat{\mathrm{E}}\left(X_{r+t_1-k:r+t_1+t_2}\right) = \frac{1}{\binom{n}{r+t_1+t_2}} \sum_{i=1}^{n} \left[\binom{i-1}{r+t_1-k-1}\binom{n-i}{t_2+k}\right] X_{i:n} \quad (3.19)$$

and $t_1 = [n\alpha_1]$ and $t_2 = [n\alpha_2]$ are the amounts of trimming from each end of the sample and α_1, α_2 are specified proportions with $0 \leq \alpha_1 < 0.5$ and $0 \leq \alpha_2 < 0.5$.

When $t_1 = t_2 = t = [n\alpha]$, α is a pre-chosen proportion ($0 \leq \alpha < 0.5$), we may write (3.18) as

$$l_r^{(t)} = r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} \widehat{\mathrm{E}}\left(X_{r+t-k:r+2t}\right), \quad r = 1, 2, ...n - 2t \quad (3.20)$$

where

$$\widehat{\mathrm{E}}\left(X_{r+t-k:r+2t}\right) = \frac{1}{\binom{n}{r+2t}} \sum_{i=1}^{n} \left[\binom{i-1}{r+t-k-1}\binom{n-i}{t+k}\right] X_{i:n} \quad (3.21)$$

Clearly sample TL-moments reduce to sample L-moments when $\alpha = 0$ and $l_1^{(0)} = \overline{X}$ is the sample mean.When $\alpha = 0.4\dot{9}$, then $l_1^{([0.4\dot{9} n])} = \widetilde{X}$ is the sample median, where $\dot{9}$ means 9 repeated; for example, when $n = 100$ and 1000, to obtain the median we must choose $\alpha = 0.49$ and $\alpha = 0.499$, respectively.

In particular, the first four sample TL-moments from equation (3.18) when $n = 25$ and $\alpha = 0.05$ are given by

$$l_1^{(1)} = \frac{1}{\binom{n}{3}} \sum_{i=1}^{n} \left[\binom{i-1}{1}\binom{n-i}{1}\right] X_{i:n}$$

$$l_2^{(1)} = \frac{1}{2\binom{n}{4}} \sum_{i=1}^{n} \left[\binom{i-1}{2}\binom{n-i}{1} - \binom{i-1}{1}\binom{n-i}{2} \right] X_{i:n}$$

$$l_3^{(1)} = \frac{1}{3\binom{n}{5}} \sum_{i=1}^{n} \left[\binom{i-1}{3}\binom{n-i}{1} - 2\binom{i-1}{2}\binom{n-i}{2} + \binom{i-1}{1}\binom{n-i}{3} \right] X_{i:n}$$

$$l_4^{(1)} = \frac{1}{4\binom{n}{6}} \sum_{i=1}^{n} \left[\binom{i-1}{4}\binom{n-i}{1} - 3\binom{i-1}{3}\binom{n-i}{2} + 3\binom{i-1}{2}\binom{n-i}{3} - \binom{i-1}{1}\binom{n-i}{4} \right] X_{i:n}$$

and when $n = 25$, $\alpha = 0.10$

$$l_1^{(2)} = \frac{1}{\binom{n}{5}} \sum_{i=1}^{n} \left[\binom{i-1}{2}\binom{n-i}{2} \right] X_{i:n}$$

$$l_2^{(2)} = \frac{1}{2\binom{n}{6}} \sum_{i=1}^{n} \left[\binom{i-1}{3}\binom{n-i}{2} - \binom{i-1}{2}\binom{n-i}{3} \right] X_{i:n}$$

$$l_3^{(2)} = \frac{1}{3\binom{n}{7}} \sum_{i=1}^{n} \left[\binom{i-1}{4}\binom{n-i}{2} - 2\binom{i-1}{3}\binom{n-i}{3} + \binom{i-1}{2}\binom{n-i}{4} \right] X_{i:n}$$

$$l_4^{(2)} = \frac{1}{4\binom{n}{8}} \sum_{i=1}^{n} \left[\binom{i-1}{5}\binom{n-i}{2} - 3\binom{i-1}{4}\binom{n-i}{3} + 3\binom{i-1}{3}\binom{n-i}{4} - \binom{i-1}{2}\binom{n-i}{5} \right] X_{i:n}$$

Also, we can estimate TL-skewness and TL-kurtosis as follows

$$t_3^{(t)} = \frac{l_3^{(t)}}{l_2^{(t)}}, \qquad t_4^{(t)} = \frac{l_4^{(t)}}{l_2^{(t)}}$$

In the following theorem, we show that the $l_r^{(t)}$ is an unbiased estimator of $\lambda_r^{(t)}$.

Theorem 3.1 *$l_r^{(t)}$ is an unbiased estimator of $\lambda_r^{(t)}$.*

Proof:
From Chapter 2 we find that

$$\mathrm{E}\left[\sum_{i=1}^{n} (i-1)^{(k)} (n-i)^{(l)} X_{i:n}\right] = k!l! \begin{pmatrix} n \\ k+l+1 \end{pmatrix} \mathrm{E}\left(X_{k+1:k+l+1}\right)$$

Using this equation and substituting in (3.20) we have

$$\begin{aligned} \mathrm{E}\left(l_r^{(t)}\right) &= \frac{1}{r\begin{pmatrix} n \\ r+2t \end{pmatrix}} \sum_{k=0}^{r-1} (-1)^k \begin{pmatrix} r-1 \\ k \end{pmatrix} \\ &\times \mathrm{E}\left\{\sum_{i=1}^{n}\left[\begin{pmatrix} i-1 \\ r+t-k-1 \end{pmatrix}\begin{pmatrix} n-i \\ t+k \end{pmatrix}\right] X_{i:n}\right\} \end{aligned}$$

then

$$\mathrm{E}\left(l_r^{(t)}\right) = r^{-1}\sum_{k=0}^{r-1} (-1)^k \begin{pmatrix} r-1 \\ k \end{pmatrix} \mathrm{E}\left(X_{r+t-k:2t+r}\right) = \lambda_r^{(t)}$$

Moreover, asymptotic normality of $l_r^{(t)}$ follows from the large sample theory of the linear functions of order statistics; see, for example, Stigler (1974) and Hosking (1986). Simulation results are illustrated in Figures 3.2, 3.3, 3.4 and 3.5. Figure 3.3 shows a slight curvature for a sample size 15 from the exponential distribution while Figures 3.2, 3.4 and 3.5 show good normal approximations.

3.3.4 Exact covariances of sample TL-moments

As the sample TL-moments are linear combinations of order statistics, we can calculate their exact sampling variances and covariances. From Downton (1966), the exact variances and covariances of $l_r^{(t_1 t_2)}$ can be written as follows

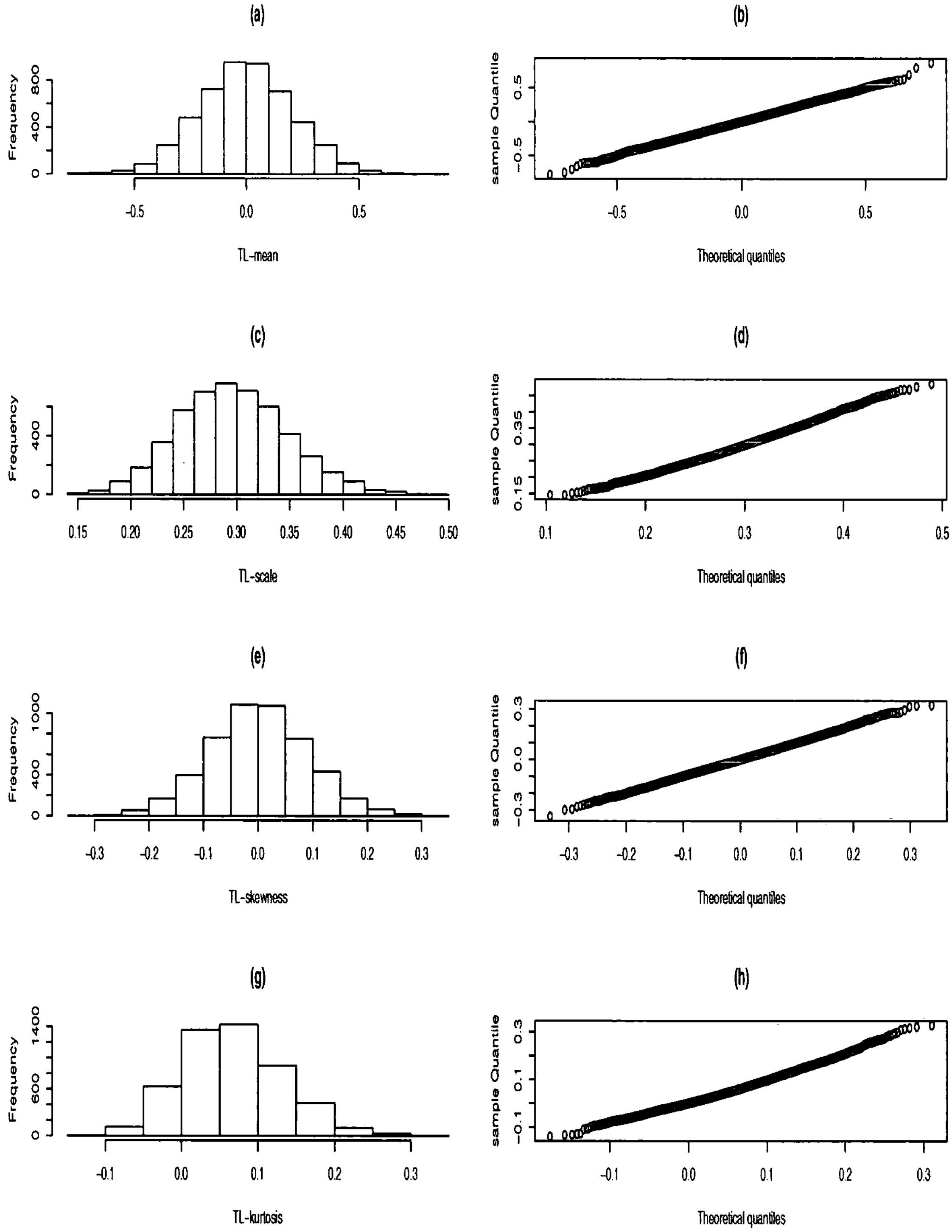

Figure 3.2: The histogram and quantile plots (a) and (b) of $l_1^{(1)}$, (c) and (d) of $l_2^{(1)}$, (e) and (f) of $t_3^{(1)}$ and (g) and (h) of $t_4^{(1)}$. The parent distribution is normal $(0, 1)$, the sample size is 25 and number of replications is 5000.

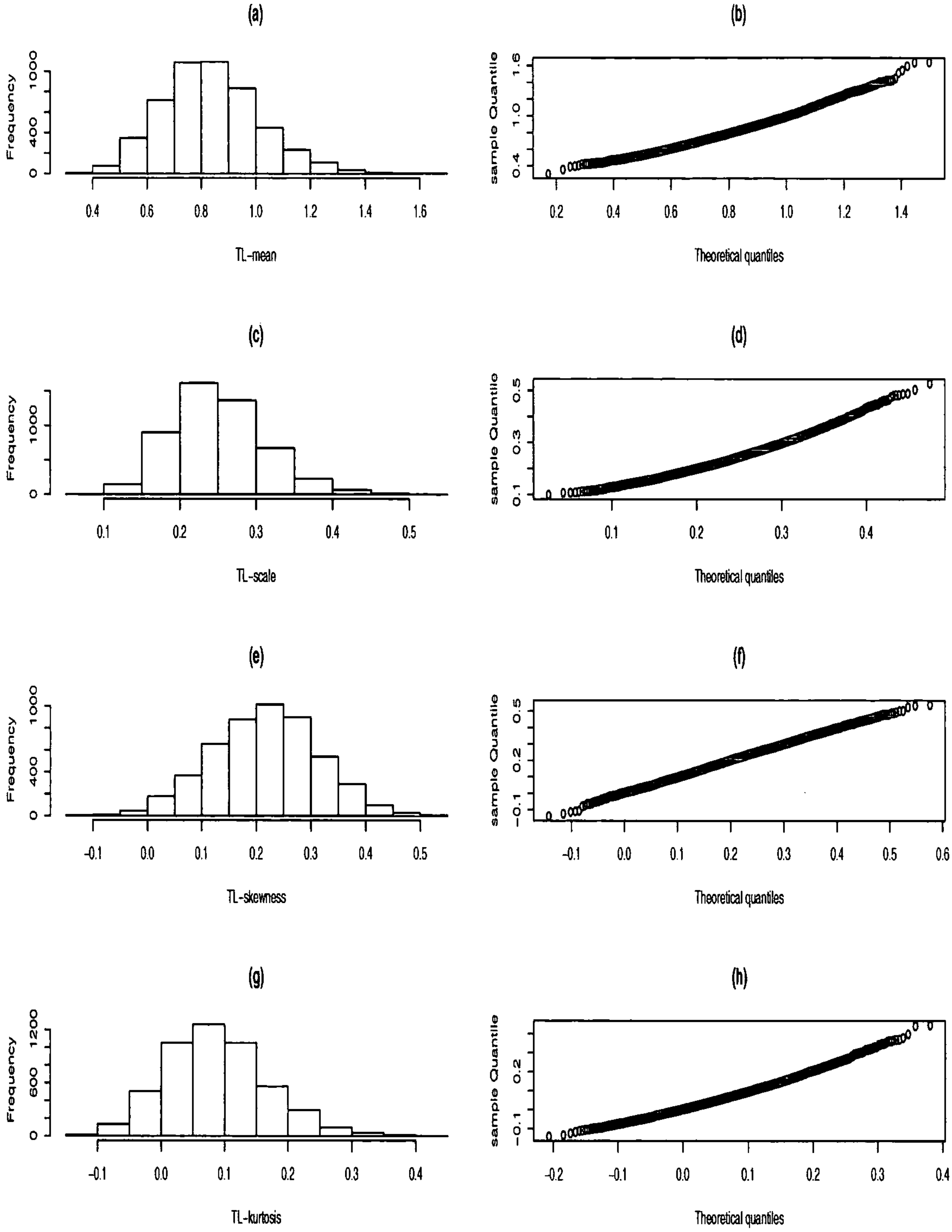

Figure 3.3: The histogram and quantile plots (a) and (b) of $l_1^{(1)}$, (c) and (d) of $l_2^{(1)}$, (e) and (f) of $t_3^{(1)}$ and (g) and (h) $t_4^{(1)}$. The parent distribution is exponential (1), the sample size is 25 and number of replications is 5000.

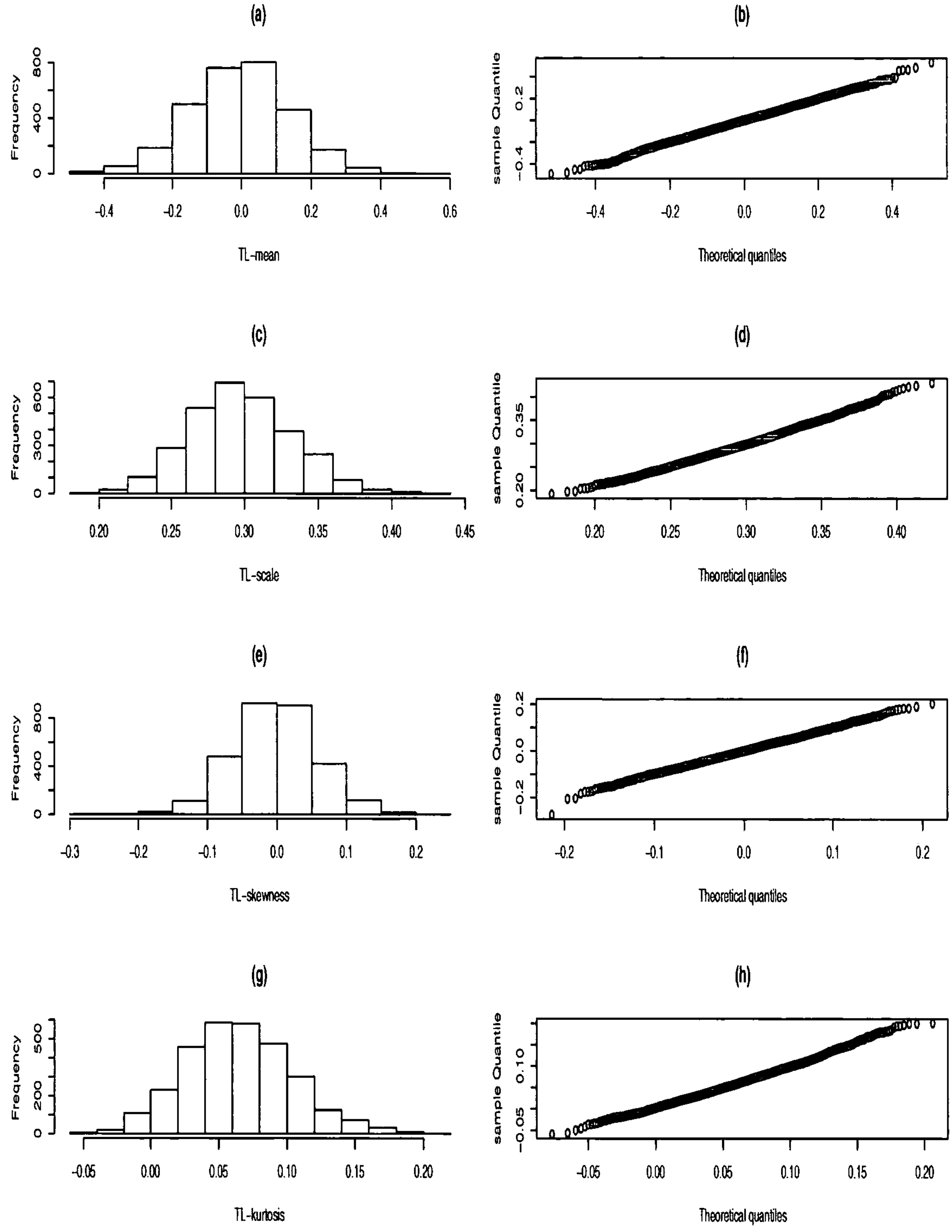

Figure 3.4: The histogram and quantiles plots of (a) and (b) $l_1^{(1)}$, (c) and (d) $l_2^{(1)}$, (e) and (f) $t_3^{(1)}$ and (g) and (h) $t_4^{(1)}$. The parent distribution is normal $(0, 1)$, the sample size is 50 and number of replications is 3000.

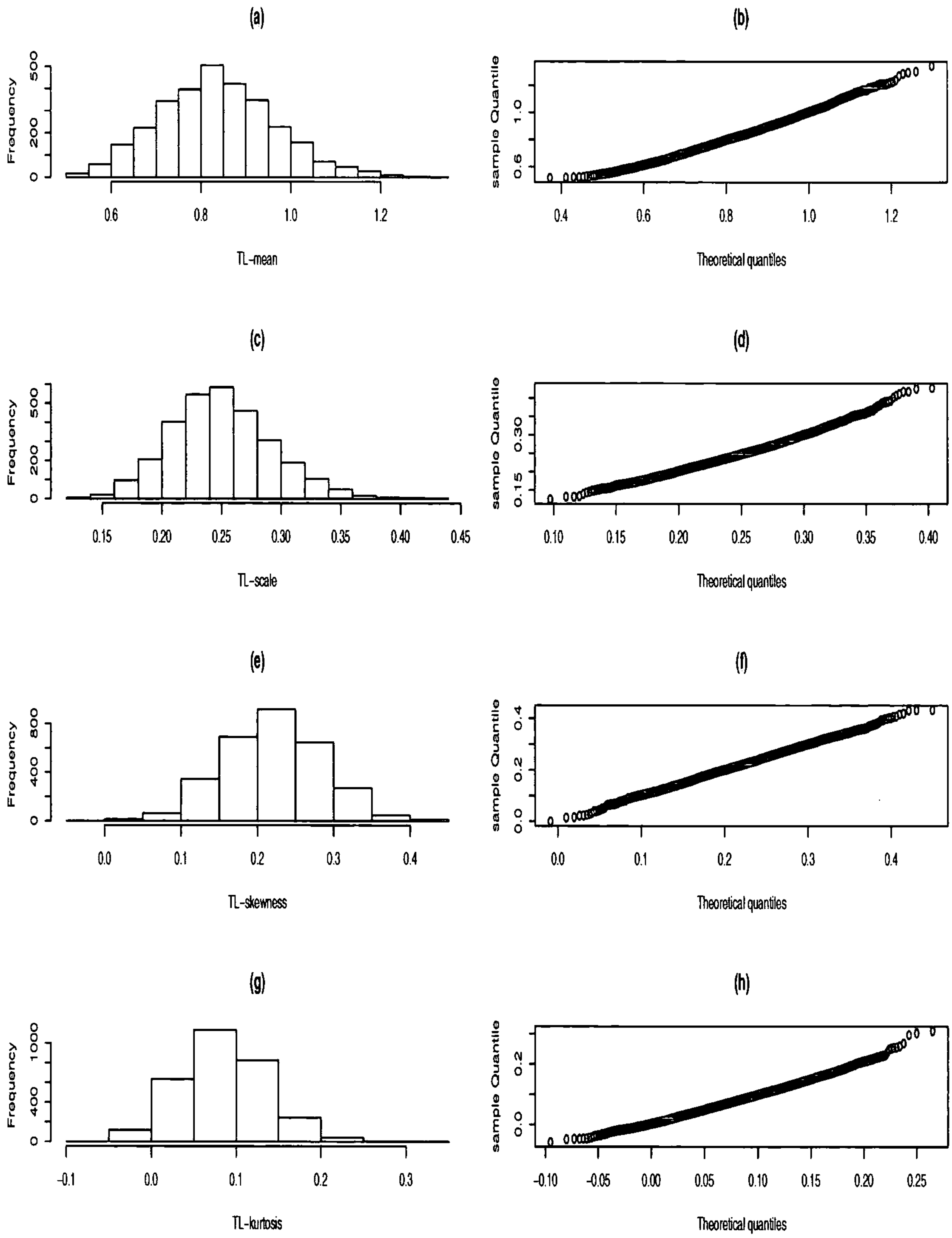

Figure 3.5: The histogram and quantile plots of (a) and (b) $l_1^{(1)}$, (c) and (d) $l_2^{(1)}$, (e) and (f) $t_3^{(1)}$ and (g) and (h) $t_4^{(1)}$. The parent distribution is exponential(1), the sample size is 50 and number of replications is 3000.

$$\begin{aligned} \operatorname{Cov}\left(l_r^{(t_1 t_2)}, l_s^{(t_1 t_2)}\right) &= r^{-1}s^{-1}\sum_{k=0}^{r-1}\sum_{l=0}^{s-1}(-1)^r(-1)^s\binom{r-1}{k}\binom{s-1}{l} \\ &\times\ \Omega\left(r+t_1-k-1, k+t_2, s+t_1-l-1, l+t_2\right) \end{aligned} \quad (3.22)$$

and in the special case when $t_1 = t_2 = t$

$$\begin{aligned} \operatorname{Cov}\left(l_r^{(t)}, l_s^{(t)}\right) &= r^{-1}s^{-1}\sum_{k=0}^{r-1}\sum_{l=0}^{s-1}(-1)^r(-1)^s\binom{r-1}{k}\binom{s-1}{l} \\ &\times\ \Omega\left(r+t-k-1, k+t, s+t-l-1, l+t\right) \end{aligned} \quad (3.23)$$

where $\Omega(k,l,p,q)$ is given by

$$\begin{aligned} \Omega(k,l,p,q) &= \left\{\sum_{r=0}^{p}\sum_{s=0}^{q}\binom{p}{r}\binom{q}{s}(p-r)!\,(q-s)!\binom{k}{p-r}\binom{l}{q-s}(k+r)!\right. \\ &\times (l+s)!\binom{n}{k+l+r+s+1}\mathrm{E}\left(X^2_{k+r+1:k+l+r+s+1}\right) \\ &+ \sum_{r=0}^{l}\sum_{s=0}^{p}(-1)^{r+s}\binom{l}{r}\binom{p}{s}\binom{n-k-r-1}{l-r}\binom{n-q-s-1}{p-s} \\ &\times (l-r)!\,(p-s)!\,(k+r)!\,(q+s)!\binom{n}{k+q+r+s+2} \\ &\times \mathrm{E}\left(X_{k+r+1:k+q+r+s+2}X_{k+r+2:k+q+r+s+2}\right) \\ &+ \sum_{r=0}^{q}\sum_{s=0}^{k}(-1)^{r+s}\binom{q}{r}\binom{k}{s}\binom{n-p-r-1}{q-r}\binom{n-l-s-1}{k-s} \\ &\times (q-r)!\,(q-s)!\,(p+r)!\,(l+s)!\binom{n}{p+l+r+s+2} \\ &\left.\times\ \mathrm{E}\left(X_{p+r+1:p+l+r+s+2}X_{p+r+2:p+l+r+s+2}\right)\right\} \\ &\Big/\left(k!l!p!q!\binom{n}{k+l+1}\binom{n}{p+q+1}\right) \\ &-\ \mathrm{E}\left(X_{k+1:k+l+1}\right)\mathrm{E}\left(X_{p+1:p+q+1}\right) \end{aligned} \quad (3.24)$$

It should be noted here that the amount of computation required to obtain the variances and covariances of sample TL-moments is considerably less than what is required to obtain the variances and covariances in the cases of using a complete sample.

Exact covariances of the first few sample TL-moments

We first give simple example to show how to use equation (3.22). We know that $l_1^{(0)}$ is the sample mean and by using equation (3.23) we find

$$\text{Var}(l_1^{(0)}) = \frac{\text{E}(X_{1:1}^2)}{n} - \frac{\text{E}(X_{1:2}X_{2:2})}{n} + \text{E}(X_{1:2}X_{2:2}) - \text{E}^2(X_{1:1})$$

For any continuous distribution it is known that; see, for example David (1981)

$$\text{E}(X_{1:2}X_{2:2}) - \text{E}^2(X_{1:1}) = 0$$

Hence we have the well-known estimator

$$\text{Var}(l_1^{(0)}) = \frac{\left[\text{E}(X_{1:1}^2) - \text{E}^2(X_{1:1})\right]}{n} = \frac{\sigma^2}{n} \tag{3.25}$$

or

$$\text{Var}(l_1^{(0)}) = \frac{[\text{E}(X_{1:1}^2) - \text{E}(X_{1:2}X_{2:2})]}{n} = \frac{\sigma^2}{n} \tag{3.26}$$

note that $\widehat{\text{var}}\left(l_1^{(0)}\right)$ is an unbiased estimator of var $\left(l_1^{(0)}\right)$ from (3.26) without any correction of the bias in (3.25).

Similarly, the exact variances and covariance of other sample TL-moments follow from (3.23); for example, the variance of TL-mean $l_1^{(1)}$ is given by

$$\begin{aligned}\text{Var}\left(l_1^{(1)}\right) &= \left\{\frac{36}{5}(n-2)^2 \text{E}\left[X_{3:5}X_{4:5} + X_{2:5}X_{3:5}\right] - 6(n-2)(2n-3)\text{E}\left(X_{2:4}X_{3:4}\right)\right. \\ &- \frac{6}{5}\left(3n^2 - 15n + 20\right)\text{E}\left(X_{3:6}X_{4:6}\right) + 6\text{E}\left(X_{2:3}^2\right) + 3(n-3)\text{E}\left[X_{3:4}^2 + X_{2:4}^2\right] \\ &+ \left.\frac{6}{5}(n-3)(n-4)\text{E}\left(X_{3:5}^2\right)\right\} / (n(n-1)(n-2))\end{aligned}$$

The variance of the TL-scale measure $l_2^{(1)}$ can be written

$$\begin{aligned}\text{Var}\left(l_2^{(1)}\right) &= 576\left\{\frac{1}{280}(2n-7)\left(n^2 - 7n + 15\right)\text{E}\left(X_{4:8}X_{5:8}\right) - \frac{3}{140}(n-3)\right. \\ &\times \left(n^2 - 6n + 10\right)\text{E}\left[X_{4:7}X_{5:7} + X_{3:7}X_{4:7}\right] + \frac{1}{120}(n-2)(n-3)(2n-5) \\ &\times \text{E}\left[X_{4:6}X_{5:6} + X_{2:6}X_{3:6}\right] + \frac{3}{160}(n-3)\left(3n^2 - 15n + 20\right)\text{E}\left(X_{3:6}X_{4:6}\right) \\ &- \frac{3}{80}(n-2)^2(n-3)\text{E}\left[X_{3:5}X_{4:5} + X_{2:5}X_{3:5}\right]\end{aligned}$$

$$
\begin{aligned}
&+ \frac{1}{96}(n-2)(n-3)(2n-3)\,\mathrm{E}\left(X_{2:4}X_{3:4}\right) \\
&- \frac{1}{1120}(n-4)(n-5)(n-6)\,\mathrm{E}\left[X_{4:7}^2\right] + \frac{1}{1680}(n-4)(n-5)(n-6) \\
&\times \mathrm{E}\left[X_{5:7}^2 + X_{3:7}^2\right] + \frac{1}{480}(n-4)(n-5)\,\mathrm{E}\left[X_{5:6}^2 + X_{2:6}^2\right] \\
&+ (n-4)\,\mathrm{E}\left[\frac{1}{80}\left(X_{2:5}^2\right) + \frac{1}{80}X_{4:5}^2 - \frac{1}{120}X_{3:5}^2\right] \\
&+ \frac{1}{96}\mathrm{E}\left[X_{3:4}^2 + X_{2:4}^2\right]\Big\} / \left(n(n-1)(n-2)(n-3)\right)
\end{aligned}
$$

The covariance between $l_1^{(1)}$ and $l_2^{(1)}$ can be written

$$
\begin{aligned}
\mathrm{Cov}\left(l_1^{(1)}, l_2^{(1)}\right) &= \Big\{6\mathrm{E}\left[X_{3:4}^2 - X_{2:4}^2\right] + \frac{9(n-4)}{5}\mathrm{E}\left[X_{4:5}^2 - X_{3:6}^2\right] \\
&+ \frac{36(n-2)^2}{5}\mathrm{E}\left[X_{2:5}X_{3:5} - X_{3:5}X_{4:5}\right] \\
&+ \frac{12(2n-5)}{5}\mathrm{E}\left[X_{4:6}X_{5:6} - X_{2:6}X_{3:6}\right] \\
&+ \frac{72(n^2-6n+10)}{35}\mathrm{E}\left[X_{3:7}X_{4:7} - X_{4:7}X_{5:7}\right]\Big\} / \left(n(n-1)(n-2)\right)
\end{aligned}
$$

The variance of $l_1^{(2)}$ is

$$
\begin{aligned}
Var\left(l_1^{(2)}\right) &= \Big\{-\frac{10(5n^4-90n^3+655n^2-2250n+3024)}{7}\mathrm{E}\left(X_{5:10}X_{6:10}\right) + \\
&+ \frac{200(n-4)^2(n^2-8n+21)}{7}\mathrm{E}\left[X_{4:9}X_{5:9} + X_{5:9}X_{6:9}\right] \\
&- \frac{225(n-3)(n-4)(n^2-7n+14)}{7}\mathrm{E}\left[X_{5:8}X_{6:8} + X_{3:8}X_{4:8}\right] \\
&- \frac{360(n-4)(2n-7)(n^2-7n+15)}{7}\mathrm{E}\left(X_{4:8}X_{5:8}\right) \\
&+ \frac{720(n-3)(n-4)(n^2-6n+10)}{7}\mathrm{E}\left[X_{4:7}X_{5:7} + X_{3:7}X_{4:7}\right] \\
&- 30(n-3)(n-4)\left(3n^2-15n+20\right)\mathrm{E}\left(X_{3:6}X_{4:6}\right) \\
&+ \frac{10(n-5)(n-6)(n-7)(n-8)}{7}\mathrm{E}\left(X_{5:9}^2\right) \\
&+ \frac{90(n-5)(n-6)(n-7)}{7}\mathrm{E}\left[X_{5:8}^2 + X_{4:8}^2\right] \\
&+ \frac{720(n-5)(n-6)}{7}\mathrm{E}\left(X_{4:7}^2\right) + \frac{120(n-5)(n-6)}{7}\mathrm{E}\left[X_{5:7}^2 + X_{3:7}^2\right] \\
&+ 120(n-5)\,\mathrm{E}\left[X_{4:6}^2 + X_{3:6}^2\right] + 120\mathrm{E}\left(X_{3:5}^2\right)\Big\} \\
&/ \quad n(n-1)(n-2)(n-3)(n-4)
\end{aligned}
$$

We can find distribution-free unbiased estimators of these variances and covariances using equations (2.49), (2.50) and (2.51) from Chapter 2.

For any symmetric distribution, we prove that Cov $\left(l_1^{(1)}, l_2^{(1)}\right) = 0$.

Theorem 3.2

$$Cov\left(l_1^{(1)}, l_2^{(1)}\right) = 0$$

for any symmetric continuous distribution function.

Proof:

From David (1981) we find for any symmetric continuous distribution function

$$\mathrm{E}\left(X_{r:n}^2\right) = \mathrm{E}\left(X_{n+1-r:n}^2\right)$$

and

$$\mathrm{E}\left(X_{r:n}X_{s:n}\right) = \mathrm{E}\left(X_{n+1-r:n}X_{n+1-s:n}\right)$$

Substituting in Cov $\left(l_1^{(1)}, l_2^{(1)}\right)$, we find that

$$\mathrm{Cov}\left(l_1^{(1)}, l_2^{(1)}\right) = 0$$

3.4 Trimmed probability-weighted moments (TPWM)

In this section we describe the trimmed probability weighted moments method (TPWM) and its relation to the method of TL-moments.

Definition 3.4 *Let X be a real-valued random variable with distribution function F. We define the trimmed probability weighted moments of X to be the quantities*

$$M_{p,r,s}^{(t_1 t_2)} = \frac{r!s!}{(r+s+1)!} E\left(X_{r+t_1+1:r+s+t_1+t_2+1}^p\right) \tag{3.27}$$

The term weight comes from the definition of probability-weighted moment method which is given in Chapter 2. In the symmetric case the definition reduces to

$$M_{p,r,s}^{(t)} = \frac{r!s!}{(r+s+1)!} \mathrm{E}\left(X_{r+t+1:r+s+2t+1}^p\right) \tag{3.28}$$

The definition of $M_{p,r,s}^{(t)}$ is valid for both continuous and discrete random variables. In the continuous case if we use equation (3.12), we can write

$$M_{p,r,s}^{(t)} = \frac{r!s!(r+t+1)!}{(r+s+1)!(r+t)!(s+t)!} \int_0^1 Q^p(u) u^{r+t} (1-u)^{s+t} du \tag{3.29}$$

The quantities $M_{p,r,s}^{(t)}$ may be used to describe and characterise probability distributions. One possible approach is to work with $M_{p,0,0}^{(0)}$, $p = 1, 2, ...$; these are just the conventional non-central moments of X; see for example Kendall and Stuart (1987).

Of particular interest are the moments $M_{1,r,s}^{(t)}$, which their estimators will be in linear form, given by

$$\begin{aligned} M_{1,r,s}^{(t)} &= \frac{r!s!}{(r+s+1)!} \mathrm{E}\left(X_{r+t+1:r+s+2t+1}\right) &(3.30)\\ &= \frac{r!s!(r+t+1)!}{(r+s+1)!(r+t)!(s+t)!} \int_0^1 Q(u) u^{r+t} (1-u)^{s+t} du &(3.31)\end{aligned}$$

As with PWM the two special cases

$$\begin{aligned} \beta_r^{(t)} = M_{1,r,0}^{(t)} &= \frac{1}{(r+1)} \mathrm{E}\left(X_{r+t+1:r+2t+1}\right) &(3.32)\\ &= \frac{r!(r+t+1)!}{(r+1)!(r+t)!(t)!} \int_0^1 Q(u) u^{r+t} (1-u)^t du &(3.33)\end{aligned}$$

$r = 0, 1, ..., n-2t$ and

$$\begin{aligned} \alpha_s^{(t)} = M_{1,0,s}^{(t)} &= \frac{1}{(s+1)} \mathrm{E}\left(X_{t+1:s+2t+1}\right) &(3.34)\\ &= \frac{s!(t+1)!}{(s+1)!(t)!(s+t)!} \int_0^1 Q(u) u^t (1-u)^{s+t} du &(3.35)\end{aligned}$$

are of particular interest where $s = 0, 1, 2, ..., n-2t$.

The characterisation of a distribution by the β_r and α_s are interchangeable because the β_r and α_s are functions of each other. Thus,

$$\beta_r^{(t)} = \sum_{k=0}^{r} (-1)^k \binom{r}{k} \alpha_k^{(t)} \tag{3.36}$$

and

$$\alpha_s^{(t)} = \sum_{k=0}^{s} (-1)^k \binom{s}{k} \beta_k^{(t)} \tag{3.37}$$

and in particular

$$\begin{aligned}
\alpha_0^{(t)} &= \beta_0^{(t)} \\
\alpha_1^{(t)} &= \beta_0^{(t)} - \beta_0^{(t)} \\
\alpha_2^{(t)} &= \beta_0^{(t)} - 2\beta_1^{(t)} + \beta_2^{(t)} \\
\alpha_3^{(t)} &= \beta_0^{(t)} - 3\beta_1^{(t)} + 3\beta_2^{(t)} - \beta_3^{(t)}
\end{aligned}$$

and

$$\begin{aligned}
\beta_0^{(t)} &= \alpha_0^{(t)} \\
\beta_1^{(t)} &= \alpha_0^{(t)} - \alpha_1^{(t)} \\
\beta_2^{(t)} &= \alpha_0^{(t)} - 2\alpha_1^{(t)} + \alpha_2^{(t)} \\
\beta_3^{(t)} &= \alpha_0^{(t)} - 3\alpha_1^{(t)} + 3\alpha_2^{(t)} - \alpha_3^{(t)}
\end{aligned}$$

Note that

$$s\alpha_{s-1}^{(0)} = \mathrm{E}\left(X_{1:s}\right) \quad \text{and} \quad r\beta_{r-1}^{(0)} = \mathrm{E}\left(X_{r:r}\right) \tag{3.38}$$

These relationships will be useful when we estimate the extreme values $\mathrm{E}(X_{n:n})$ and $\mathrm{E}(X_{1:1})$ and the parameters from a uniform distribution.

Although TPWM are useful to characterize a distribution, they have no particular meaning, for example, we have to combine β_1 and β_0 to find definition for the scale. It is useful to define some functions of TPWM, which can be seen as descriptive parameters of location, scale and shape of a probability distribution. TL-moments and the TPWM are related by

$$\begin{aligned}
\lambda_{r+1}^{(t)} &= \sum_{k=0}^{r}(-1)^{r-k}\binom{r}{k}\binom{r+k}{k}\beta_k^{(t)} && (3.39) \\
&= (-1)^r\sum_{k=0}^{r}(-1)^{r-k}\binom{r}{k}\binom{r+k}{k}\alpha_k^{(t)} && (3.40)
\end{aligned}$$

In particular,

$$\begin{aligned}
\lambda_1^{(t)} &= \alpha_0^{(t)} = \beta_0^{(t)} \\
\lambda_2^{(t)} &= \alpha_0^{(t)} - 2\alpha_1^{(t)} = 2\beta_1^{(t)} - \beta_0^{(t)} \\
\lambda_3^{(t)} &= \alpha_0^{(t)} - 6\alpha_1^{(t)} + 6\alpha_2^{(t)} = 6\beta_2^{(t)} - 6\beta_1^{(t)} + \beta_0^{(t)} \\
\lambda_4^{(t)} &= \alpha_0^{(t)} - 12\alpha_1^{(t)} + 30\alpha_2^{(t)} - 20\alpha_3^{(t)} = 20\beta_3^{(t)} - 30\beta_2^{(t)} + 12\beta_1^{(t)} - \beta_0^{(t)}
\end{aligned}$$

3.4.1 Estimation of TPWM moments

The TPWM can be estimated in a straightforward manner by estimating the expected value of order statistics as we have done with TL-moments. We define the rth sample TPWM as follows

Definition 3.5 *For samples of size n, the rth sample TPWM is given by*

$$\widehat{M}_{1,r,s}^{(t_1 t_2)} = \frac{r!s!}{(r+s+1)!}\widehat{E}\left(X_{r+t_1+1:r+s+t_1+t_2+1}\right) \tag{3.41}$$

where

$$\widehat{E}\left(X_{r+t_1+1:r+s+t_1+t_2+1}\right) = \frac{1}{\begin{pmatrix} n \\ r+s+t_1+t_2+1 \end{pmatrix}} \sum_{i=1}^{n} \left[\begin{pmatrix} i-1 \\ r+t_1 \end{pmatrix}\begin{pmatrix} n-i \\ s+t_2 \end{pmatrix}\right] X_{i:n}$$

In the symmetric case this simplifies to

$$\widehat{M}_{1,r,s}^{(t)} = \frac{r!s!}{(r+s+1)!}\widehat{\mathrm{E}}\left(X_{r+t+1:r+s+2t+1}\right) \tag{3.42}$$

where

$$\widehat{\mathrm{E}}\left(X_{r+t+1:r+s+2t+1}\right) = \frac{1}{\begin{pmatrix} n \\ r+s+2t+1 \end{pmatrix}} \sum_{i=1}^{n} \left[\begin{pmatrix} i-1 \\ r+t \end{pmatrix}\begin{pmatrix} n-i \\ s+t \end{pmatrix}\right] X_{i:n}$$

3.5 TL-mean as a robust statistic

Most studies of robust estimators of location have been primarily concerned with their asymptotic properties; see for example, Gastwirth (1966), Gastwirth and Rubin (1969), Siddiqui and Raghunandanan (1967) and Collins (2000). Crow and Siddiqui (1967) and Gastwirth and Cohen (1970) studied the small sample behaviour of some robust estimators of location for normal, double exponential and Cauchy data. The results indicated that a suitably trimmed mean performed quite well.

In this section we compare the TL-mean with some robust location estimators which are defined below and are given in Gastwirth and Cohen (1970).

3.5.1 TL-mean

Suppose $X_1, X_2, ..., X_n$ is a random sample from a distribution $F\left((x-\mu)/\sigma\right)$, where μ is an unknown location parameter to be estimated, and σ is a known scale parameter, which without loss of generality we take to be 1.

We investigate TL-mean as a robust estimator of μ and we define it as follows

$$l_1^{(t)} = \frac{1}{\binom{n}{1+2t}} \sum_{i=1}^{n} \left[\binom{i-1}{t} \binom{n-i}{t} \right] X_{i:n} \qquad (3.43)$$

We shall compare the TL-mean with the mean $\overline{X}$, median $\widetilde{X}$, the α-Winsorized mean $W_n(\alpha)$, the α-trimmed mean $T_n(\alpha)$ and $Y_n(p, a)$ which is a combination of the median and the upper and lower pth sample fractiles such that each of the fractiles receives a weight a and the median $1 - 2a$.

$$T_n(\alpha) = (n - 2[n\alpha])^{-1} \sum_{i=[n\alpha]+1}^{n-[n\alpha]} X_{i:n} \qquad (3.44)$$

$$W_n(\alpha) = \frac{1}{n} \left\{ [n\alpha] \left(X_{[n\alpha]} + X_{n+1-[n\alpha]} \right) + \sum_{[n\alpha]+1}^{n-[n\alpha]} X_{i:n} \right\} \qquad (3.45)$$

$$Y_n(p, a) = a \left(X_{[pn]+1} + X_{n-[np]} \right) + (1 - 2a)\widetilde{X}, \qquad (3.46)$$

where α denotes the fraction trimmed from each end of the sample and $[m]$ denotes the largest integer in m: for more details, including the variances of these estimators; see, for example, Gastwirth and Cohen (1970) and Jaeckel (1971).

We study the behaviour of the estimators based on samples of different sizes from normal, logistic, double-exponential and the contaminated normal distributions $CN(\gamma, K)$ with density

$$f_{\gamma,K}(x) = \frac{1}{\sqrt{2\pi}} \left\{ (1-\gamma)\, e^{-x^2/2} + \frac{\gamma}{K} e^{-x^2/2K^2} \right\} \qquad (3.47)$$

We use the results given in Gastwirth and Cohen (1970) for estimators other than $l_1^{(t)}$and the same value of α of the trimmed mean to compute $l_1^{(t)}$. The results in Table 3.2 and Figures 3.6, 3.7 and 3.8, the symmetric lambda distribution is introduced and discussed further in Chapter 4, show that the ordinary sample mean and the 0.05, 0.10 and 0.125 Winsorized means seem to lack robustness properties for example, at sample size 16 the variances of these are 0.125, 0.125 and 0.1172, while the minimum variance is 0.o849 ($l_1^{(2)}$). The median, on the other hand, incurs too much loss of efficiency

Estimators	Distribution					
	Normal	CN(0.01, 3)	CN(0.05, 3)	CN(0.10, 3)	Logistic	D-Exp
			$n = 8$			
$\overline{X}\left(l_1^{(0)}\right)$	0.125	0.135	0.1750	0.2250	0.4112	0.2500
$\widetilde{X}\left(l_1^{(3)}\right)$	0.168	0.171	0.1825	0.1989	0.4533	0.1873
$T_8\,(0.125)$	0.134	0.137	0.1513	0.1737	0.3906	0.1997
$T_8\,(0.25)$	0.148	0.150	0.1618	0.1784	0.4086	0.1839
$W_8\,(0.125)$	0.125	0.135	0.1750	0.2250	0.4112	0.2500
$W_8\,(0.250)$	0.133	0.140	0.1530	0.1803	0.3996	0.2193
$Y_8\left(\frac{1}{4}, 0.3\right)$	0.146	0.149	0.1605	0.1776	0.4071	0.1868
$Y_8\left(\frac{1}{3}, \frac{1}{3}\right)$	0.146	0.148	0.1603	0.17765	0.4075	0.1894
$Y_8\left(\frac{1}{3}, 0.3\right)$	0.146	0.149	0.1605	0.17763	0.4072	0.1868
$l_1^{(1)}$	0.137	0.140	0.1532	0.1728	0.4113	0.1892
$l_1^{(2)}$	0.151	0.154	0.1651	0.1815	0.3998	0.1816
			$n = 16$			
$\overline{X}\left(l_1^{(0)}\right)$	0.062	0.067	0.0874	0.1125	0.2056	0.125
$\widetilde{X}\left(l_1^{(7)}\right)$	0.090	0.092	0.0972	0.1038	0.2366	0.0855
$T_{16}\,(0.10)$	0.064	0.066	0.0742	0.0874	0.1934	0.1048
$T_{16}\,(0.125)$	0.067	0.068	0.0744	0.0835	0.1919	0.0948
$W_{16}\,(0.10)$	0.062	0.067	0.0874	0.1125	0.2056	0.125
$W_{16}\,(0.125)$	0.064	0.066	0.0764	0.0942	0.2006	0.1172
$Y_{16}\left(\frac{1}{4}, 0.3\right)$	0.073	0.074	0.0795	0.0874	0.2008	0.0902
$Y_{16}\left(\frac{1}{3}, \frac{1}{3}\right)$	0.077	0.078	0.0832	0.0907	0.2067	0.0853
$Y_{16}\left(\frac{1}{3}, 0.3\right)$	0.077	0.078	0.0836	0.0910	0.2072	0.0851
$l_1^{(1)}$	0.067	0.068	0.0747	0.084	0.1994	0.0914
$l_1^{(2)}$	0.072	0.073	0.0783	0.0860	0.1859	0.0849
			$n = 20$			
$\overline{X}\left(l_1^{(0)}\right)$	0.050	0.054	0.0700	0.0900	0.1645	0.1000
$\widetilde{X}\left(l_1^{(9)}\right)$	0.073	0.074	0.0789	0.0851	0.1911	0.0666
$T_{20}\,(0.05)$	0.051	0.051	0.0594	0.0706	0.1551	0.0854
$T_{20}\,(0.10)$	0.053	0.054	0.0589	0.0666	0.1529	0.0778
$W_{20}\,(0.05)$	0.050	0.054	0.0700	0.0900	0.1645	0.1000
$W_{20}\,(0.10)$	0.051	0.051	0.0616	0.0769	0.1610	0.0950
$Y_{20}\left(\frac{1}{4}, 0.3\right)$	0.058	0.059	0.0686	0.0697	0.1604	0.0718
$Y_{20}\left(\frac{1}{3}, \frac{1}{3}\right)$	0.061	0.111	0.0657	0.0715	0.1634	0.0685
$Y_{20}\left(\frac{1}{3}, 0.3\right)$	0.061	0.062	0.0661	0.0718	0.1637	0.0673
$l_1^{(1)}$	0.053	0.054	0.0595	0.0669	0.1586	0.0727
$l_1^{(2)}$	0.057	0.058	0.0621	0.0682	0.1472	0.0672

Table 3.2: Variances of various location estimators for samples of different sizes ($n = 8, 16$ and 20) from different distributions.

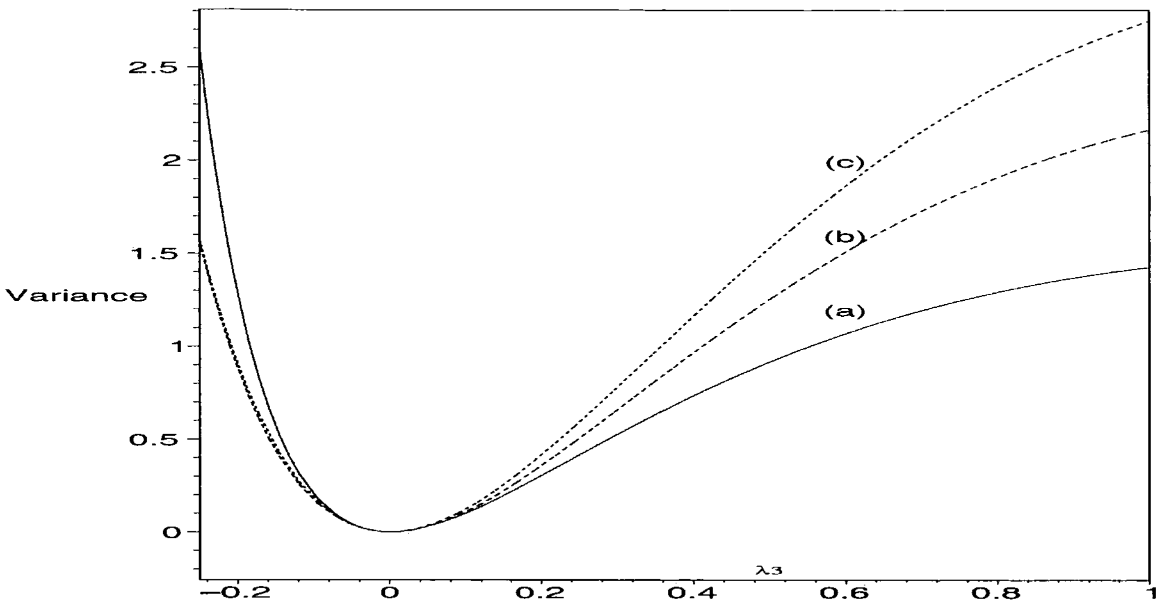

Figure 3.6: Variances (a) of $l_1^{(0)}$, (b) of $l_1^{(1)}$, and (c) of $l_1^{(2)}$ as a function of λ_3 for samples of size 6 from a symmetric lambda distribution for various values of λ_3, where $\lambda_1 = 0$ and $\lambda_2 = 1$.

on normal and near normal data. For all the sample sizes studied, our results indicate that the TL-mean $l_1^{(t)}$ is suitable for general use in small to moderate sample sizes and like the trimmed mean lies between the sample mean $\overline{X}$ and the median $\widetilde{X}$ but the TL-mean gives different weights, which decreases from a maximum weight for the median values to zero for $2t$ the trimmed extremes, to observations and more efficient for suitable choosing for α; for example, $l_1^{(0)}$, $l_1^{(1)}$ and $l_1^{(2)}$ have the smallest variance among other estimators from normal, CN, logistic and double-exponential distributions except from CN$(0.05, 3)$ distribution, note also that the values of variances of $\overline{X}\left(l_1^{(0)}\right)$, $W_8\,(0.125)$, $W_{20}\,(0.05)$ and $W_{16}\,(0.10)$ are the same because the precentage and the sample size are small.

In the contaminated normal densities we used the tables of variances and covariances of order statistics given in Gastwirth and Cohen (1970). From these we find the following variances of $l_1^{(1)}$ and $l_1^{(2)}$

When $\gamma = 0.01$

$$\begin{aligned}
\text{Var}\left(l_1^{(1)}\right) &= \frac{1.0806(n-1.1771)(n-1.6038)}{n(n-1)(n-2)} \\
\text{Var}\left(l_1^{(2)}\right) &= \frac{1.1328(n-3.3672)(n-3.6394)(n^2-2.5705n+1.7722)}{n(n-1)(n-2)(n-3)(n-4)}
\end{aligned}$$

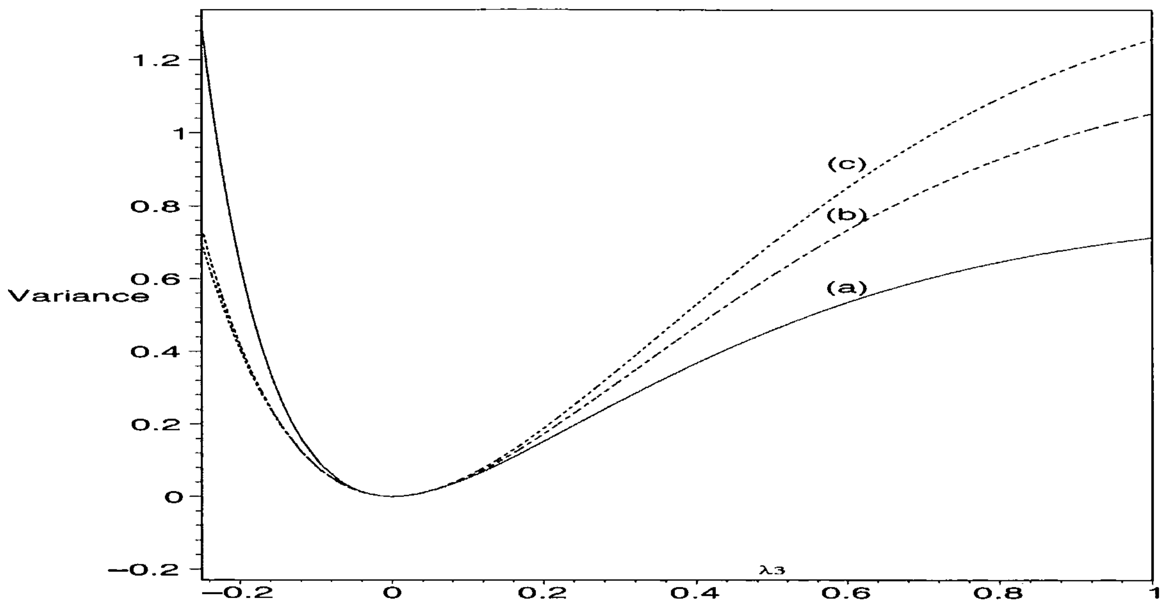

Figure 3.7: Variances (a) of $l_1^{(0)}$, (b) of $l_1^{(1)}$, and (c) of $l_1^{(2)}$ as a function of λ_3 for samples of size 12 from a symmetric lambda distribution for various values of λ_3, where $\lambda_1 = 0$ and $\lambda_2 = 1$.

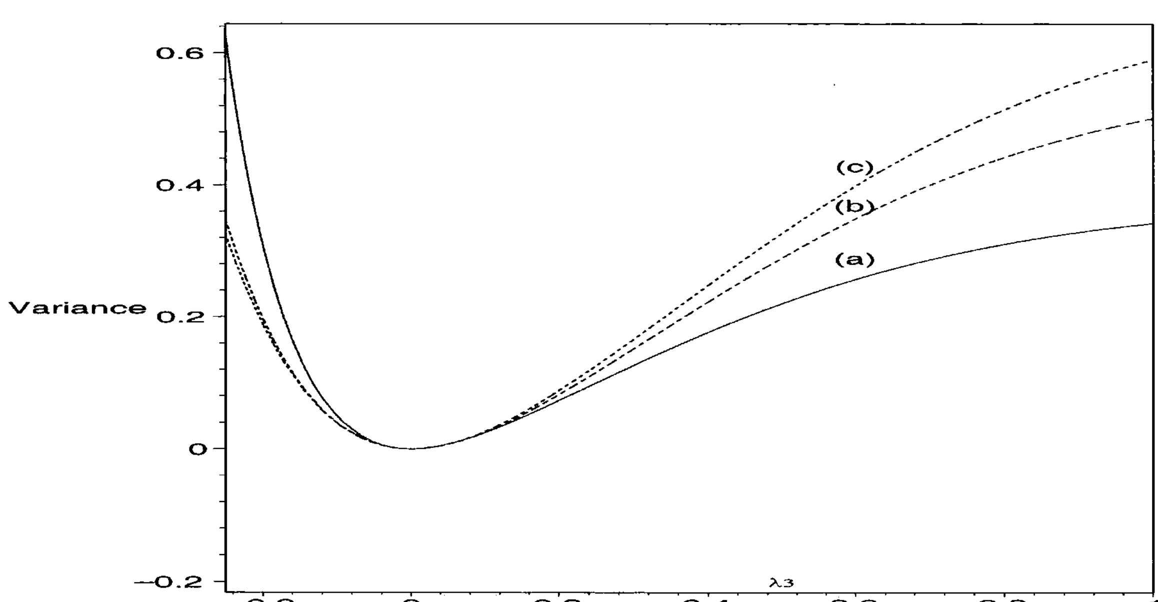

Figure 3.8: Variances of (a) $l_1^{(0)}$, (b) $l_1^{(1)}$, and (c) $l_1^{(2)}$ as a function of λ_3 for samples of size 25 from a symmetric lambda distribution for various values of λ_3, where $\lambda_1 = 0$ and $\lambda_2 = 1$.

When $\gamma = 0.05$

$$\operatorname{Var}\left(l_1^{(1)}\right) = \frac{1.1755(n-1.1202)(n-1.6340)}{n(n-1)(n-2)}$$

$$\operatorname{Var}\left(l_1^{(2)}\right) = \frac{1.2119(n-3.3456)(n-3.6523)(n^2-2.5609n+1.7300)}{n(n-1)(n-2)(n-3)(n-4)}$$

When $\gamma = 0.10$

$$\operatorname{Var}\left(l_1^{(1)}\right) = \frac{1.3186(n-1.0443)(n-1.6681)}{n(n-1)(n-2)}$$

$$\operatorname{Var}\left(l_1^{(2)}\right) = \frac{1.3296(n-3.3104)(n-3.6727)(n^2-2.5593n+1.6970)}{n(n-1)(n-2)(n-3)(n-4)}$$

3.6 Examples of TL-moments for some symmetric distributions

In this section we estimate the parameters of some symmetric probability distributions using TL-moments.

3.6.1 Uniform distribution

The uniform distribution is a short tail distribution with

$$f(x) = \frac{1}{\beta-\alpha}, \quad F(x) = \frac{x-\alpha}{\beta-\alpha} \quad \text{and} \quad Q(F) = \alpha + (\beta-\alpha)F$$

We consider estimating the parameters α and β using the following expression for the TPWM

$$n\alpha_{n-1}^{(0)} = \frac{n\alpha+\beta}{n+1} \quad \text{and} \quad n\beta_{n-1}^{(0)} = \frac{\alpha+n\beta}{n+1}$$

Using (3.38) and solving these equations we obtain the estimators

$$\widehat{\alpha} = \frac{nX_{1:n} - X_{n:n}}{n-1} \quad \text{and} \quad \widehat{\beta} = \frac{nX_{n:n} - X_{1:n}}{n-1}$$

where $\widehat{\alpha}$ and $\widehat{\beta}$ are unbiased estimators of α and β, respectively.

The estimators $\widehat{\alpha}$ and $\widehat{\beta}$ are the modified maximum likelihood and maximum product of spacings estimators, see for example; Cheng and Amin (1983).

3.6.2 Normal distribution

For the normal distribution, when $t = 1$ and using Teichroew (1954) tables we find that

$$\lambda_1^{(1)} = \mu, \quad \lambda_2^{(1)} = 0.297\,\sigma, \quad \lambda_3^{(1)} = 0 \quad \text{and} \quad \lambda_4^{(1)} = 0.0185\,\sigma$$

and

$$\tau_3^{(1)} = 0 \text{ and } \tau_4^{(1)} = 0.0625$$

Thus, unbiased estimators of μ and σ are

$$\widehat{\mu} = l_1^{(1)} \quad \text{and} \quad \widehat{\sigma} = l_2^{(1)}/0.297$$

Where the variances of $l_1^{(1)}$ and $l_1^{(2)}$ are

$$\text{Var}\left(l_1^{(1)}\right) = \frac{1.0594\,(n-1.1875)\,(n-1.5980)}{(n\,(n-1)\,(n-2))}$$

and

$$\text{Var}\left(l_1^{(2)}\right) = \frac{1.1122(n-3.3708)(n-3.6371)(n^2-2.5676n+1.7702)}{n(n-1)(n-2)(n-3)(n-4)}$$

which are used to compute the variances of $l_1^{(1)}$ and $l_1^{(2)}$ in Table 3.2.

3.6.3 Logistic distribution

In the logistic distribution $Q\,(u) = \mu + \sigma\ \left(u/\,(1-u)\right)$ so when $t = 1$ we find that

$$\lambda_1^{(1)} = \mu, \ \lambda_2^{(1)} = 0.5\,\sigma, \ \lambda_3^{(1)} = 0 \text{ and } \lambda_4^{(1)} = 0.0417\,\sigma$$

and

$$\tau_3^{(1)} = 0 \text{ and } \tau_4^{(1)} = 0.0833$$

Therefore, unbiased estimators of μ and σ are

$$\widehat{\mu} = l_1^{(1)} \quad \text{and} \quad \widehat{\sigma} = 2l_2^{(1)}$$

Where the variances of $l_1^{(1)}$ and $l_1^{(2)}$ are

$$\text{Var}\left(l_1^{(1)}\right) = \frac{3.1133(n-0.9476)(n-1.7048)}{n(n-1)(n-2)}$$

and

$$\text{Var}\left(l_1^{(2)}\right) = \frac{2.8265(n-0.3140)(n-2.8918)(n-2.8918)(n-3.7812)}{n(n-1)(n-2)(n-3)(n-4)}$$

which are used to compute the variances of $l_1^{(1)}$ and $l_1^{(2)}$ in Table 3.2.

3.6.4 Double exponential distribution

In the double exponential distribution $Q(u) = \mu + \sigma \log(2u)$ for $u < 0.5$ and $Q(u) = \mu - \sigma \log(2u)$ for $u > 0.5$. When $t = 1$ we find that

$$\lambda_1^{(1)} = \mu, \quad \lambda_2^{(1)} = 0.3438\,\sigma, \quad \lambda_3^{(1)} = 0, \text{ and } \lambda_4^{(1)} = 0.0469\,\sigma$$

and

$$\tau_3^{(1)} = 0, \text{ and } \tau_4^{(1)} = 0.1365$$

Therefore, unbiased estimators of μ and σ are

$$\widehat{\mu} = l_1^{(1)} \text{ and } \widehat{\sigma} = l_2^{(1)}/0.344$$

Where the variances of $l_1^{(1)}$ and $l_1^{(2)}$ are

$$\text{Var}\left(l_1^{(1)}\right) = \frac{1.425\,(n-0.887)\,(n-1.726)}{(n\,(n-1)\,(n-2))}$$

and

$$\text{Var}\left(l_1^{(2)}\right) = \frac{1.2956(n-0.5415)(n-1.9764)(n-3.0517)(n-3.7620)}{n(n-1)(n-2)(n-3)(n-4)}$$

which are used to compute the variances of $l_1^{(1)}$ and $l_1^{(2)}$ in Table 3.2.

3.6.5 Cauchy distribution

In the Cauchy distribution $Q(u) = \mu + \sigma\ (\tan \pi\,(u - 0.5))$. When $t = 1$ we find that

$$\lambda_1^{(1)} = \mu, \quad \lambda_2^{(1)} = 0.698\,\sigma, \quad \lambda_3^{(1)} = 0, \text{ and } \lambda_4^{(1)} = 0.239\,\sigma$$

and

$$\tau_3^{(1)} = 0, \text{ and } \tau_4^{(1)} = 0.343$$

Therefore, unbiased estimators of μ and σ are

$$\widehat{\mu} = l_1^{(1)} \text{ and } \widehat{\sigma} = l_2^{(1)}/0.698$$

Note that, we used the tables of variances and covariances of order statistics given in Govindarajulu (1966a), Barnett (1966), Teichroew (1954) and Gupta (1967) to compute the values of $\lambda_1^{(1)}$, $\lambda_2^{(1)}$, $\lambda_3^{(1)}$, and $\lambda_4^{(1)}$and variances of $l_1^{(1)}$ and $l_1^{(2)}$.

3.7 Conclusions

In this chapter we have discussed LQ-moments based on the median, the tri-mean and Gastwirth location measures, have extended L-moments to trimmed L-moments and shown that the L-moment is a special case of TL-moments.

We have defined both population and sample TL-mean, TL-scale, TL-skewness and TL-kurtosis and have obtained the variances and covariances of sample TL-moments in closed form. We have also investigated the properties of TL-moments in some symmetric distributions, the uniform, normal, logistic, Laplace and Cauchy. Also, we have shown that the TL-mean is a robust measure of location, protects against outliers and gives different weights to the observations. We have described the trimmed probability weighted method (TPWM) and its relation to the TL-moment method.

Chapter 4

Symmetric lambda distribution

4.1 Introduction

Tukey (1962) introduced and discussed subsequently the very useful family of distributions defined by the single parameter quantile function

$$Q(p) = \frac{p^{\lambda} - (1-p)^{\lambda}}{\lambda} \tag{4.1}$$

where $0 < p < 1$.

Random variables with this quantile function are said to be distributed according to a symmetric lambda distribution with parameter λ. Filliben (1969) used this distribution to approximate symmetric distributions with a wide range of tail weights to study location estimators of symmetric distributions. Joiner and Rosenblatt (1971) have given results on the sample range. Chan and Rhodin (1980) used this distribution to study robust estimation of the location parameter based on selected order statistics. Ramberg and Schmeiser (1972) have shown how this distribution can be used to approximate many of the known symmetric distributions and explored its application to Monte Carlo simulation studies. Ramberg and Schmeiser (1974) generalised (4.1) to a four-parameter distribution defined by the quantile function

$$Q(p) = \lambda_1 + \frac{p^{\lambda_3} - (1-p)^{\lambda_4}}{\lambda_2}, \quad 0 < p < 1 \tag{4.2}$$

where λ_1 is a location parameter, λ_2 is a scale parameter and λ_3 and λ_4 are shape parameters.

Although the distribution functions corresponding to (4.1) and (4.2) do not exist in closed form, this should not be of concern to practitioners since the same is true of the normal distribution, whose quantile function is not available in closed form. In this chapter we work with symmetric case $\lambda_3 = \lambda_4$

$$Q(p) = \lambda_1 + \frac{p^{\lambda_3} - (1-p)^{\lambda_3}}{\lambda_2} \tag{4.3}$$

In section 4.2 we give the properties of the symmetric lambda distribution. In section 4.3 we discuss estimating the parameter λ in (4.1) using the maximum likelihood method. Also, we discuss the use of L-moments and LQ-moments for estimating the parameters λ_1, λ_2 and λ_3. In section 4.4 we obtain the asymptotic variances of $\widehat{\lambda}_1$, $\widehat{\lambda}_2$ and $\widehat{\lambda}_3$ from the exact variances and covariances of sample L-moments derived in Chapter 2. In Section 4.5 we use the symmetric lambda distribution to study the effect of the tail of the distribution on the choice of the plotting position for quantile plots.

4.2 Properties of the symmetric lambda distribution

In this section we list some of the properties of the symmetric lambda distribution, most of which follow from the quantile function (4.3), which is a legitimate inverse distribution function whenever λ_2 and λ_3 have the same sign.

The density function of a random variable X with quantile function (4.3) is defined implicitly by

$$f(x) = f(Q(p)) = 1/Q'(p) = \lambda_2 \left[\lambda_3 p^{\lambda_3 - 1} + \lambda_3 (1-p)^{\lambda_3 - 1}\right]^{-1} \tag{4.4}$$

Which can be graphed by letting p range from zero to one and plotting $f(Q(p))$ versus $Q(p)$; see Figure 4.1 for some examples. The density function is symmetric about λ_1. For $\lambda_2, \lambda_3 > 0$ the range of variation of $x = Q(p)$ is positive in the interval $\lambda_1 \pm 1/\lambda_2$ and 0 otherwise; for $\lambda_2, \lambda_3 < 0$ the range of variation is positive on $(-\infty, \infty)$, and the ordinates at the extremes of the range of variation of $x = Q(p)$ are given by

$$f(Q(0)) = f(Q(1)) = \begin{cases} \lambda_2/\lambda_3 & \text{if } \lambda_3 > 1 \\ \lambda_2/2 & \text{if } \lambda_3 = 1 \\ 0 & \text{if } \lambda_3 < 1 \end{cases} \tag{4.5}$$

As Figure 4.2 shows the density of X is U-shaped for $1 < \lambda_3 < 2$ and has a single mode for $\lambda_3 < 1$ or $\lambda_3 > 2$; see Figure 4.1.

Tukey (1962) found that this family of distributions gives useful approximations to the percentage points of the normal and Student's t distributions. An important property

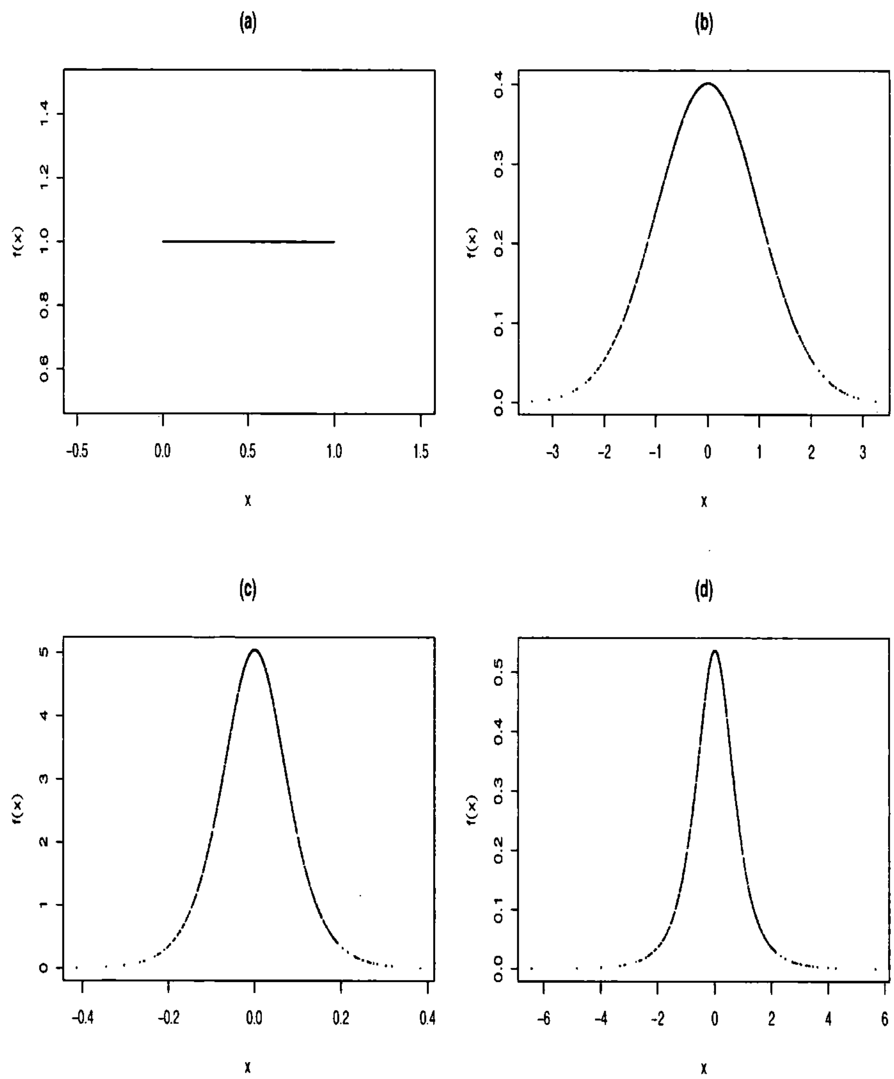

Figure 4.1: Probability density function of some symmetric lambda distributions for different parameter values: (a) $\lambda_1 = 0.5$, $\lambda_2 = 2$ and $\lambda_3 = 1$, (b) $\lambda_1 = 0$, $\lambda_2 = 0.1974$ and $\lambda_3 = 0.1349$, (c) $\lambda_1 = 0$, $\lambda_2 = -0.0870$ and $\lambda_3 = -0.0043$ and (d) $\lambda_1 = 0$, $\lambda_2 = -0.3203$ and $\lambda_3 = -0.1359$.

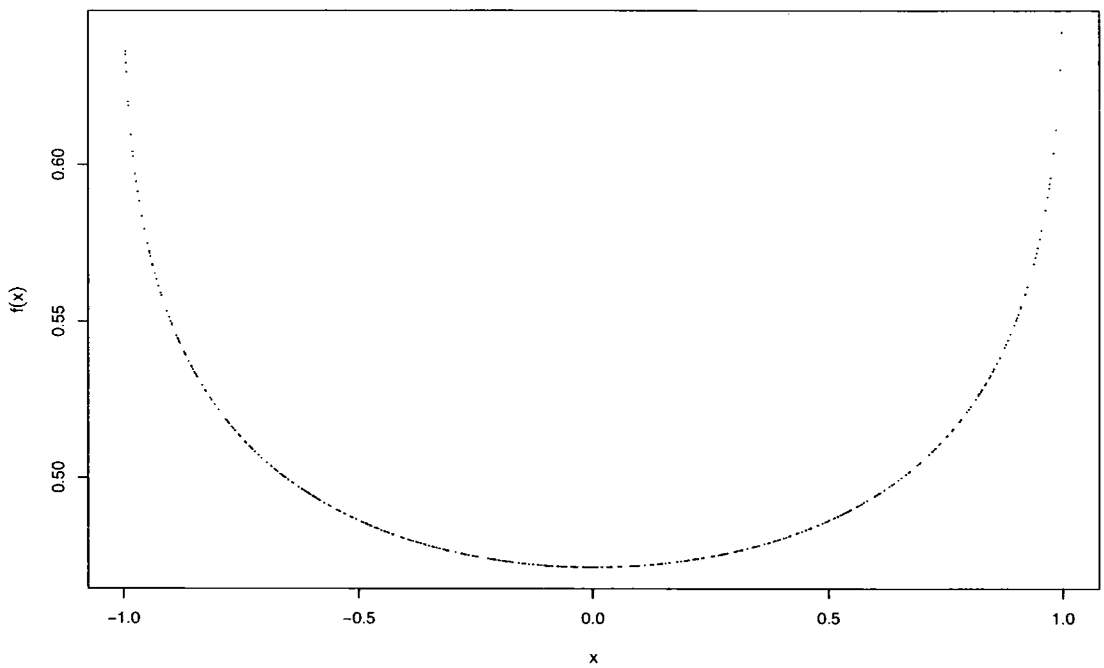

Figure 4.2: Density function of lambda distribution when $\lambda_1 = 0$, $\lambda_2 = 1$ and $\lambda_3 = 1.5$.

of the lambda family is that the percentage points are available directly from equation (4.3). That is, the $100p$ percent point is given by $Q(p)$.

Ramberg et al. (1979) have given the kth moment as

$$\mathrm{E}\left(X^k\right) = \lambda_2^{-k} \sum_{i=0}^{k} \binom{k}{i} (-1)^i \beta\left(\lambda_3 (k-i) + 1, \lambda_3 i + 1\right) \tag{4.6}$$

where $\lambda_1 = 0$ and $\beta(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx$ is the beta function.

The kth moment of the rth order statistic of a random sample of size n can be computed from

$$\begin{aligned} \mathrm{E}\left(X_{r:n}^k\right) &= \frac{r\binom{n}{r}}{\lambda_2^k} \left\{ \sum_{j=0}^{k-1} \left[\sum_{s=0}^{j} (-1)^s \binom{j}{s} \beta\left(r + \lambda_3(j-s), n + s\lambda_3 - r + 1\right)\right] \binom{k}{j} \right. \\ & \left. \times \lambda_1^{k-j} + \sum_{s=0}^{k} (-1)^s \binom{k}{s} \beta\left(r + \lambda_3 (k-s), n + s\lambda_3 - r + 1\right) \right\} \end{aligned} \tag{4.7}$$

Mykytka and Ramberg (1979) gave the proof only for $\lambda_1 = 0$.

The following expressions for the mean and the variance of X are obtained from

(4.6) as

$$\begin{aligned} \mathrm{E}(X) &= \lambda_1 \\ \mathrm{Var}(X) &= 2\left[1/\left(2\lambda_3+1\right) - \beta\left(\lambda_3+1, \lambda_3+1\right)\right]/\lambda_2^2 \end{aligned} \tag{4.8}$$

When $\lambda_3 = 1$ or $\lambda_3 = 2$, $Q(p)$ is linear in p so the distribution is uniform with p.d.f.

$$f(x) = \lambda_2/2 \tag{4.9}$$

A limiting form of (4.1) as $\lambda \to 0$ is the logistic distribution with the density

$$f(x) = \frac{e^x}{(1+e^x)^2} \tag{4.10}$$

There are several advantages of the family of symmetric lambda distributions

I. It includes distributions with short as well as very long tails;

II. It includes the logistic distribution ($\lambda \to 0$) and two uniform distributions ($\lambda = 1, 2$) as particular cases;

III. It provides good approximations to the Cauchy, normal and t-distributions;

IV. The percentage points are easily calculable; and

V. Simulation is immediate from the definition of the quantile function.

4.3 Methods of estimation

In this section we shall consider four methods of estimating the parameters of a symmetric lambda distribution; maximum likelihood in subsection 4.3.1, L-moments in subsection 4.3.2, LQ-moments in subsection 4.3.3 and TL-moments in subsection 4.3.4.

4.3.1 Maximum likelihood

Suppose that random variables $X_1, X_2, ..., X_n$ are independent and identically distributed with probability density function $f(x \mid \theta)$, where $\theta = (\theta_1, \theta_2, \cdots, \theta_k)$ are parameters to be estimated. The likelihood of θ is defined as

$$l(\theta) = \prod_{i=1}^{n} f(X_i|\theta) \tag{4.11}$$

The maximum likelihood estimate of θ is which value of $\widehat{\theta}$ that maximises $l(\theta)$. Rather than maximising the likelihood itself, it is usually easier to maximise its natural logarithm, i.e the log likelihood

$$L(\theta) = \sum_{i=1}^{n} \log f(X_i|\theta) \tag{4.12}$$

The large sample distribution of the maximum likelihood estimates $\widehat{\theta}$ is approximately normal with mean θ_0 and variance $\frac{1}{n}\mathrm{I}^{-1}(\theta_0)$, where the matrix $\mathrm{I}(\theta)$ has the $j-k$th element

$$\mathrm{I}(\Theta) = \mathrm{E}\left(-\frac{\partial^2}{\partial\theta_j\partial\theta_k}\log f(X|\theta)\right) \tag{4.13}$$

In this section we restrict attention to estimating the parameter λ in the one parameter family (4.1) using the maximum likelihood method because it is difficult to work with generalised case.

Limits of lambda

We first give the limits of λ in the following lemma

Lemma 4.1 *The upper and lower limits of λ will be*

$$-\infty < \lambda < L$$

where $L = 1/max(-X_{1:n}, X_{n:n})$

It is easy to show $L > 0$ by using $X_{1:n} < X_{n:n}$ and considering the three cases, both negative, opposite sign and both positive. Writing $l = X_{1:n}$ and $m = X_{n:n}$ we have three cases:

- $l < m < 0$: $-l > 0$ then $\max(-l, m) = -l > 0$
- $l < 0 < m$: $-l > 0$ and $m > 0$ then $\max(-l, m) > 0$
- $0 < l < m$: $-l < 0$ and $m > 0$ then $\max(-l, m) = m > 0$.

In each case $\max(-X_{1:n}, X_{n:n}) > 0$, and therefore, $L > 0$.

To prove Lemma 4.1 we consider the two cases (a) $\lambda > 0$ and (b) $\lambda < 0$.

When $\lambda > 0$, then

$$Q(0) = -\frac{1}{\lambda} \quad \text{and} \quad Q(1) = \frac{1}{\lambda}$$

Hence

$$-\frac{1}{\lambda} < X_{1:n} < X_{2:n} < \cdots < X_{n:n} < \frac{1}{\lambda}$$

So that

$$-X_{1:n} < \frac{1}{\lambda} \quad \text{and} \quad X_{n:n} < \frac{1}{\lambda}$$

therefore

$$\frac{1}{\lambda} > \max(-X_{1:n}, X_{n:n})$$

and because $L > 0$, we have

$$0 < \lambda < L \tag{4.14}$$

When $\lambda < 0$, we can show that

$$Q\,(0) = -\infty \quad \text{and} \quad Q\,(1) = \infty \tag{4.15}$$

and the Lemma follows after noting that the logistic distribution corresponds to $\lambda \to 0$.

Likelihood function

From the density function (4.4) with $\lambda_2 = \lambda_3 = \lambda$, the likelihood function is

$$l(\lambda) = \prod_{i=1}^{n} \left(p_i^{\lambda-1} + (1-p_i)^{\lambda-1}\right)^{-1} \tag{4.16}$$

where p_i is such that $x_i = Q\,(p_i) = \left(p_i^{\lambda} - (1-p_i)^{\lambda}\right)/\lambda$.

The log-likelihood function is

$$L(\lambda) = -\sum_{i=1}^{n} \log\,\left(p_i^{\lambda-1} + (1-p_i)^{\lambda-1}\right) \tag{4.17}$$

Thus

$$\begin{aligned} L'(\lambda) &= -\sum_{i=1}^{n}\left(\frac{p_i^{\lambda-1}\log p_i + (1-p_i)^{\lambda-1}\log(1-p_i)}{p_i^{\lambda-1}+(1-p_i)^{\lambda-1}} - \frac{(\lambda-1)p_i^{\lambda-2} - (\lambda-1)(1-p_i)^{\lambda-2}}{p_i^{\lambda-1}+(1-p_i)^{\lambda-1}}\right. \\ &\quad \left. \times \frac{\lambda(p_i^{\lambda}\log p_i - (1-p_i)^{\lambda}\log(1-p_i)) - (p_i^{\lambda}-(1-p_i)^{\lambda})}{\lambda^2(p_i^{\lambda-1}+(1-p_i)^{\lambda-1})}\right) \end{aligned} \tag{4.18}$$

The maximum likelihood estimator $(\widehat{\lambda})$ is taken to be the value of λ which maximises $L(\lambda)$. Maximising $L(\lambda)$ requires numerical solution. We used the **fsolve** func-

1.546	1.847	-0.484	0.692	0.093	-0.283
1.612	2.244	1.233	0.260	1.062	-0.709
-0.557	-2.066	1.421	0.458	-2.628	-2.203
-0.699	-2.303	-1.243	0.349	1.258	-2.148
-0.340	2.387	-0.307	2.715	3.664	1.489

Table 4.1: Simulated sample of size $n = 30$ from a symmetric lambda distribution with parameter $\lambda = 0.14$.

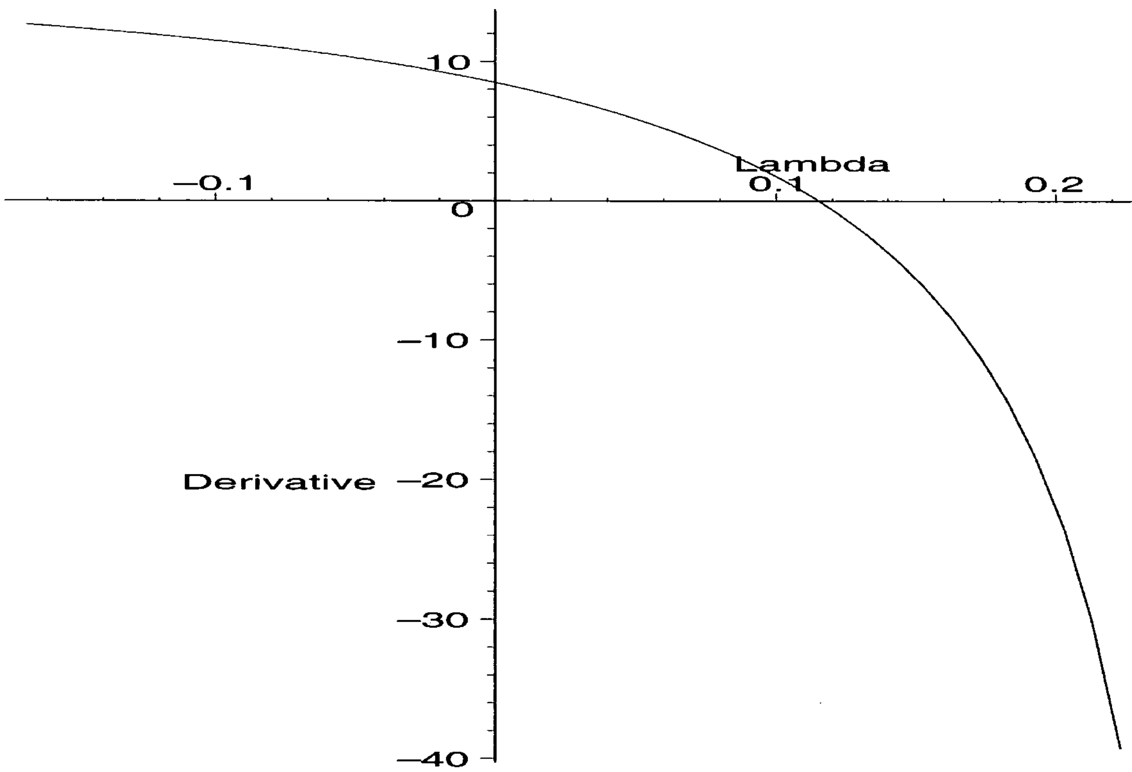

Figure 4.3: Derivative of log likelihood function $L'(\lambda)$ based on a simulated sample of size $n = 30$ from the lambda distribution with $\lambda = 0.14$.

tion in **xmaple 6** to find the numerical solution for equation (4.18) but in many cases **fsolve** required many iterations, a lot of times, a lot of memory and in some cases failed to compute the solution. As an example, we simulate a sample of size $n = 30$ from lambda distribution with $\lambda = 0.14$ (see Table 4.1) and we have plotted $L'(\lambda)$ and $L(\lambda)$ (see Figures 4.3 and 4.4) which give $\widehat{\lambda} \simeq 0.11$.

The likelihood function has simple form for the values $\lambda = 0, 1$ and 2; for example, when $\lambda \to 0$

$$l(0) = \prod_{i=1}^{n} \frac{e^{x_i}}{(1 + e^{x_i})^2} \tag{4.19}$$

which is the likelihood of the logistic distribution.

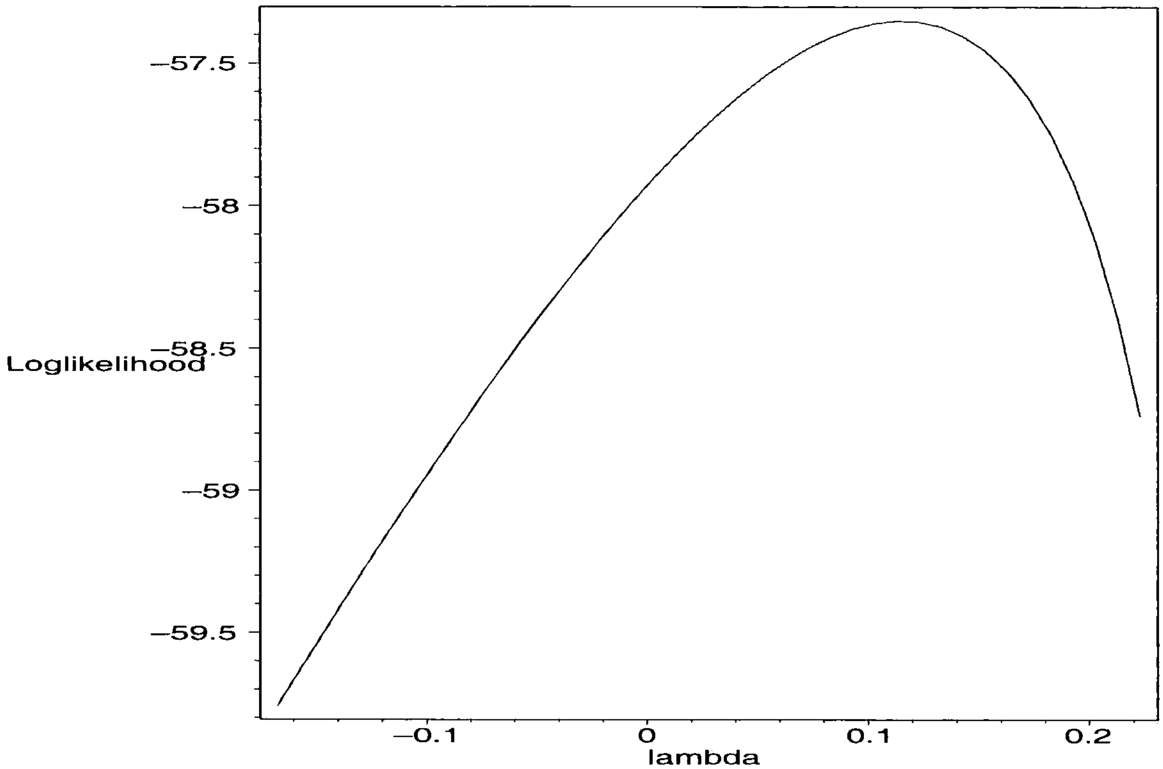

Figure 4.4: Log likelihood function $L(\lambda)$ based on simulated sample of size $n = 30$ from the lambda distribution with $\lambda = 0.14$.

4.3.2 L-moments

In this section we discuss the use of L-moments for estimating the parameters of the symmetric lambda distribution. L-moments are expectations of certain linear combinations of order statistics as defined in Chapter 2.

L-moments for the symmetric lambda distribution

Estimates of the parameters of the symmetric lambda distribution may be obtained by L-moments as follows. From Chapter 2 we have

$$\mathrm{E}(l_r) = r^{-1}\sum_{k=0}^{r-1}(-1)^k \binom{r-1}{k} \mathrm{E}(X_{r-k:r}), \quad r = 1, 2, \ldots \tag{4.20}$$

We have the expected value of order statistics in equation (4.7). From equation (4.20) we conclude that

$$\mathrm{E}(l_1) = \mathrm{E}(X) = \lambda_1$$

		τ_4							
		0	0.070	0.123	0.154	0.167	0.236	0.255	0.350
λ_1		0	0	0	0	0	0	0	0
λ_2	small	$\sqrt{3}/3$	0.396	0.203	0.064	−0.0017	−0.446	−0.608	−2.025
	large	$\sqrt{3}/3$	0.499	0.456	0.434	0.425	0.384	0.374	0.3284
λ_3	small	1	0.368	0.140	0.037	−0.0009	−0.174	−0.214	−0.386
	large	2	3.384	4.262	4.783	5.006	6.263	6.637	8.771

Table 4.2: Some values of λ_1, λ_2 and λ_3 corresponding to the selected values of τ_4 from the standard lambda distributions.

$$
\begin{aligned}
\mathrm{E}(l_2) &= \frac{1}{2}\mathrm{E}(X_{2:2} - X_{1:2}) = \frac{2\lambda_3}{\lambda_2(\lambda_3+1)(\lambda_3+2)} \\
\mathrm{E}(l_3) &= \frac{1}{3}\mathrm{E}(X_{3:3} - 2X_{2:3} + X_{1:3}) = 0 \\
\mathrm{E}(l_4) &= \frac{1}{4}\mathrm{E}(X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}) = \frac{4\lambda_3 - 6\lambda_3^2 + 2\lambda_3^3}{\lambda_2(\lambda_3+1)(\lambda_3+2)(\lambda_3+3)(\lambda_3+4)} \\
\tau_3 &= \frac{\mathrm{E}(l_3)}{\mathrm{E}(l_2)} = 0 \\
\tau_4 &= \frac{\mathrm{E}(l_4)}{\mathrm{E}(l_3)} = \frac{(\lambda_3-1)(\lambda_3-2)}{(\lambda_3+3)(\lambda_3+4)}
\end{aligned}
$$

In the special case when $\lambda_1 = 0$ and $\lambda_2 = \lambda_3 = \lambda$ we have

$$
\begin{aligned}
\mathrm{E}(l_1) &= 0 \\
\mathrm{E}(l_2) &= \frac{2}{(\lambda+1)(\lambda+2)} \\
\mathrm{E}(l_3) &= 0 \\
\mathrm{E}(l_4) &= \frac{2(\lambda-1)(\lambda-2)}{(\lambda+1)(\lambda+2)(\lambda+3)(\lambda+4)}
\end{aligned}
$$

The minimum value of τ_4 for the family of lambda distributions is about -0.01 corresponding approximately to $\lambda_3 \simeq 1.449$, see Figure 4.5. There are two lambda distributions corresponding to every permissible value of τ_4; see equation (4.21) and Table 4.2.

The probability density functions for the two values of lambda corresponding to $\tau_4 = 0.236$ are shown in Figure 4.6.

To estimate λ_1, λ_2 and λ_3 we equate the theoretical L-moments to the computed

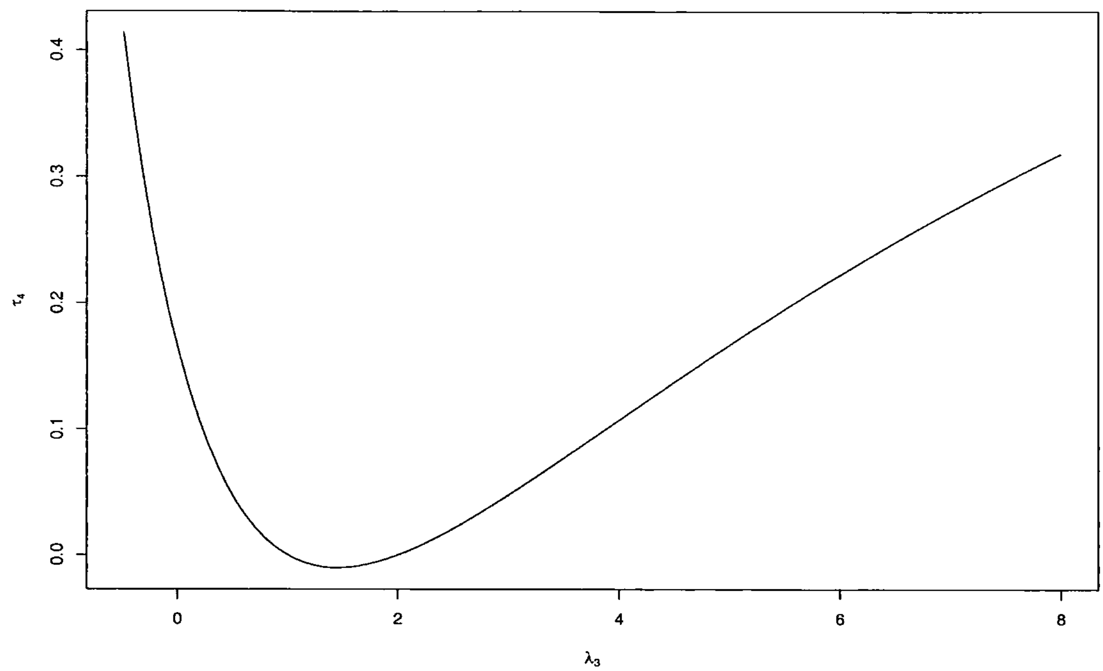

Figure 4.5: τ_4 as a function of λ_3

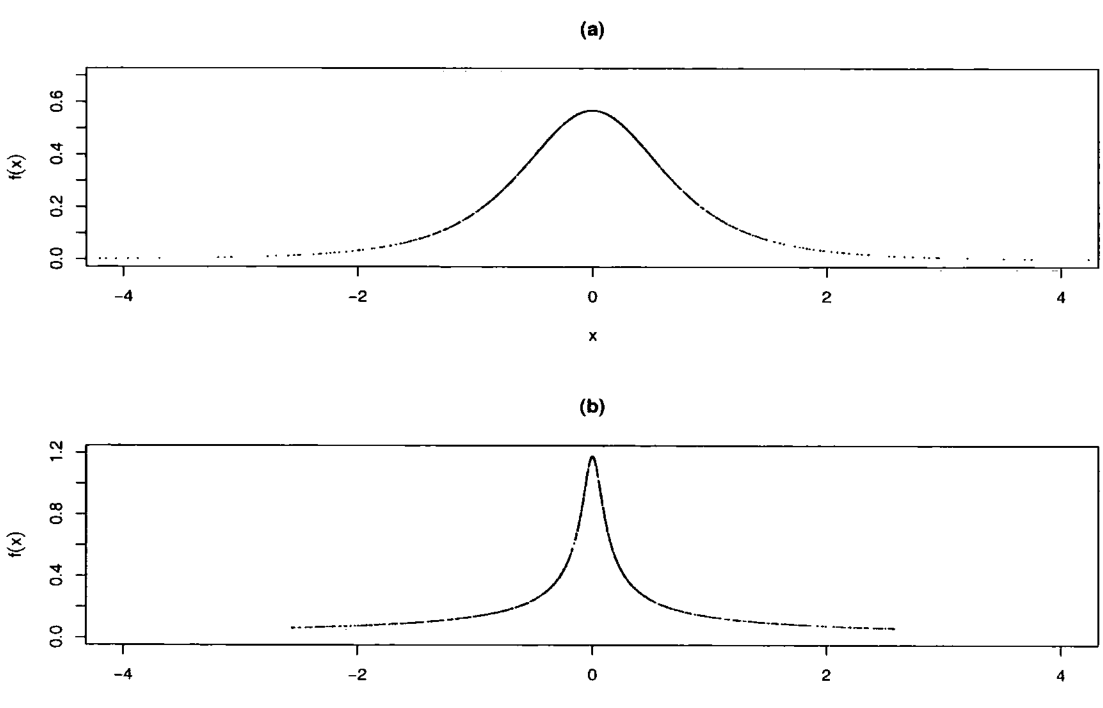

Figure 4.6: Probability density functions of the two lambda distributions having $\tau_3 = 0$ and $\tau_4 = 0.236$ and unit standard deviation.

sample L-moments; thus,

$$\tau_4 = \frac{(\lambda_3 - 1)(\lambda_3 - 2)}{(\lambda_3 + 3)(\lambda_3 + 4)}$$

which can be written

$$(1 - \tau_4)\lambda_3^2 - \lambda_3(3 + 7\tau_4) + 2\,(1 - 6\tau_4) = 0 \tag{4.21}$$

Solving this with respect to λ_3, we obtain the two estimators

$$\widehat{\lambda}_{31} = \frac{3 + 7t_4 - \sqrt{t_4^2 + 98t_4 + 1}}{2\,(1 - t_4)} \tag{4.22}$$

and

$$\widehat{\lambda}_{32} = \frac{3 + 7t_4 + \sqrt{t_4^2 + 98t_4 + 1}}{2\,(1 - t_4)} \tag{4.23}$$

Substituting in E (l_2) we have the following two corresponding estimates of λ_2

$$\widehat{\lambda}_{21} = \frac{2\widehat{\lambda}_{31}}{l_2(\widehat{\lambda}_{31} + 1)(\widehat{\lambda}_{31} + 2)} \tag{4.24}$$

and

$$\widehat{\lambda}_{22} = \frac{2\widehat{\lambda}_{32}}{l_2(\widehat{\lambda}_{32} + 1)(\widehat{\lambda}_{32} + 2)} \tag{4.25}$$

From E (l_1) we obtain

$$\widehat{\lambda}_1 = l_1 = \overline{X} \tag{4.26}$$

4.3.3 LQ-moments

In this section, we discuss the use of the LQ-moments for estimating the parameters of the symmetric lambda distribution. As we have seen in Chapter 3, Mudholkar and Hutson (1998) introduced LQ-moments which are linear functions of location measures (median, tri-mean, and Gastwirth) of the distributions of order statistics.

From Chapter 3 and using tri-mean approximation ($p = .25$ and $\alpha = .25$), we find that

$$\xi_1 = \lambda_1$$

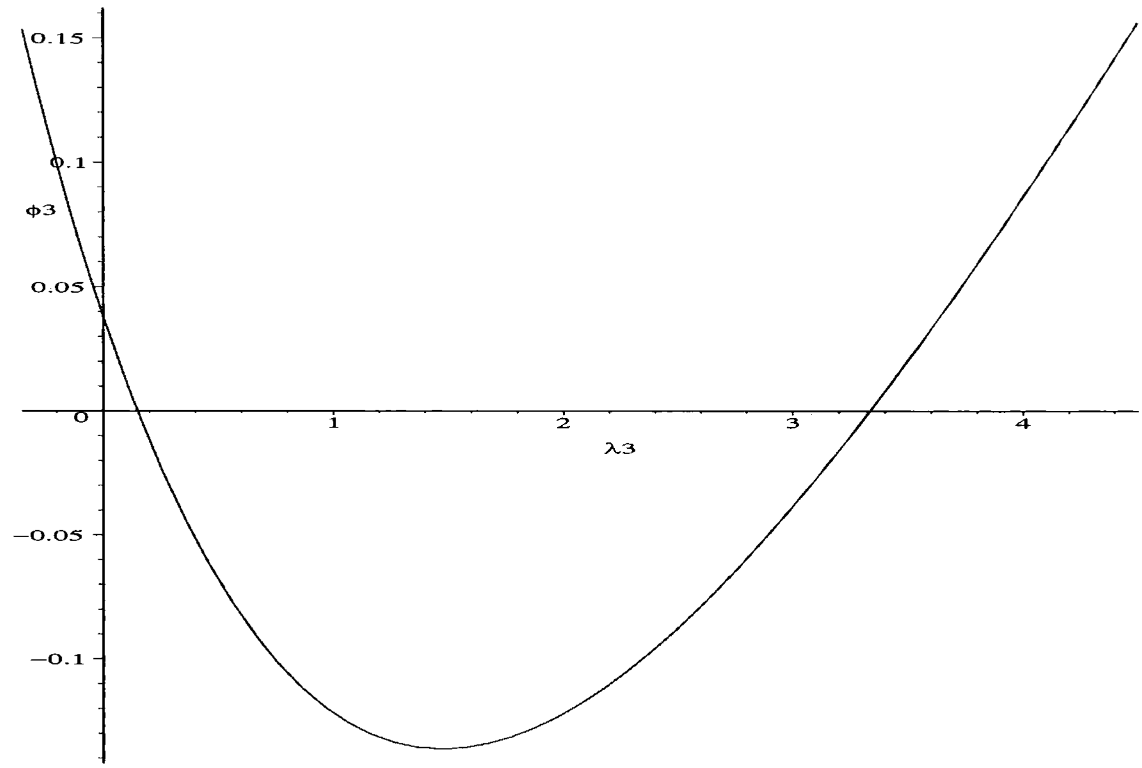

Figure 4.7: Values of λ_3 when η_4 is 0.118

$$\begin{aligned}
\xi_2 &= \left[\frac{2e^{-0.346\lambda_3} - 2e^{-1.228\lambda_3} + e^{-0.144\lambda_3} - e^{-2.010\lambda_3}}{4\lambda_2}\right] \\
\xi_3 &= 0 \qquad (4.27) \\
\xi_4 &= \frac{1}{8\lambda_2}\Big[e^{-0.3465\lambda_3} - e^{-1.228\lambda_3} + 2e^{-0.173\lambda_3} - 2e^{-1.838\lambda_3} + e^{-0.072\lambda_3} - e^{-2.668\lambda_3} \\
&- 3e^{-0.784\lambda_3} + 3e^{-0.609\lambda_3} - 6e^{-0.487\lambda_3} + 6e^{-0.9526\lambda_3} - 3e^{-0.278\lambda_3} + 3e^{-1.414\lambda_3}\Big]
\end{aligned}$$

Then

$$\eta_3 = \xi_3/\xi_2 = 0 \quad \text{and} \quad \eta_4 = \xi_4/\xi_2 = \phi(\lambda_3) \qquad (4.28)$$

We notice that η_4 is a function of λ_3 only , ξ_2 is a function of λ_2 and λ_3 and ξ_1 is a function of λ_1. Solving equation (4.28) with respect to λ_3 using **fsolve** in maple gives two values of λ_3, see Figure 4.7, and substituting in equation (4.27) gives two values of λ_2 and one value of λ_1. For example, the standard normal distribution has $\eta_4 = 0.118$ and this gives the two solutions $\lambda_1 = 0$, $\lambda_2 = 0.212$ and $\lambda_3 = 0.148$ and $\lambda_1 = 0$, $\lambda_2 = 0.502$ and $\lambda_3 = 3.334$; see Figure 4.8.

The minimum value of η_4 for the family of symmetric lambda distributions is about $\eta_4 \simeq -0.010$ corresponding approximately to $\lambda_3 \simeq 1.4$. There is no upper bound on

λ's		η_4				
		$-.004$	0.118	0.156	0.283	0.528
λ_1		0	0	0	0	0
λ_2	small	0.577556	0.212575	0.0000632	1.93423	(-1.97914)
	large	0.57748	0.502218	0.486004	0.44286	(0.16004)
λ_3	small	1.00215	0.148057	0.0000281	-0.38008	-0.880419
	large	1.9973	3.3337	3.63816	4.56318	6.21457

Table 4.3: Some values of λ_1, λ_2 and λ_3 corresponding to selected values of η_4 and using the tri-mean ($p = 0.25$ and $a = 0.25$) in LQ-moments from standard lambda distributions.

possible values of η_4 for this family and, as shown in Figure 4.8 and Table 4.3, there are two lambda distributions corresponding to every permissible value of η_4.

The probability density functions for the two values of λ_2 and λ_3 corresponding to $\eta_4 = 0.118$ are shown in Figure 4.9. The distribution for $\lambda_3 = 3.333$ shown there illustrates the interesting phenomenon that distributions corresponding to larger values of λ_3 have extreme peakedness and short high tails. Thus the class of lambda distributions having large values of λ_3 illustrates that "peakedness", "kurtosis", and "tail length" are not always synonymous. Karian and Dudewicz (2000) have suggested that large values of λ_3 may be of interest to those investigating the properties of truncated distributions, since severely truncated distributions ordinarily have short high tails.

If we used the median ($p = 0.5$ and $\alpha = 0.5$) we find that

$$\begin{aligned}
\xi_1 &= \lambda_1 \\
\xi_2 &= \left[\frac{e^{-0.346\lambda_3} - e^{-1.228\lambda_3}}{\lambda_2}\right] \\
\xi_3 &= 0 \\
\xi_4 &= \left[\frac{e^{-0.173\lambda_3} - e^{-1.838\lambda_3} - 3e^{-0.487\lambda_3} + 3e^{-0.953\lambda_3}}{2\lambda_2}\right]
\end{aligned} \tag{4.29}$$

Then

$$\eta_3 = \xi_3/\xi_2 = 0 \quad \text{and} \quad \eta_4 = \xi_4/\xi_2 = \phi(\lambda_3) \tag{4.30}$$

We notice that η_4 is a function of λ_3 only , ξ_2 is a function of λ_2 and λ_3 and ξ_1 is a function of λ_1. Solving equation (4.30) using **fsolve** in maple gives two values of λ_3 and substituting in equation (4.29) gives two values of λ_2 and we have one value of λ_1. For example, the standard normal distribution has $\eta_4 = 0.116$ and this gives two solutions $\lambda_1 = 0$, $\lambda_2 = 0.216$ and $\lambda_3 = 0.151$ and $\lambda_1 = 0$, $\lambda_2 = 0.516$ and $\lambda_3 = 3.088$;

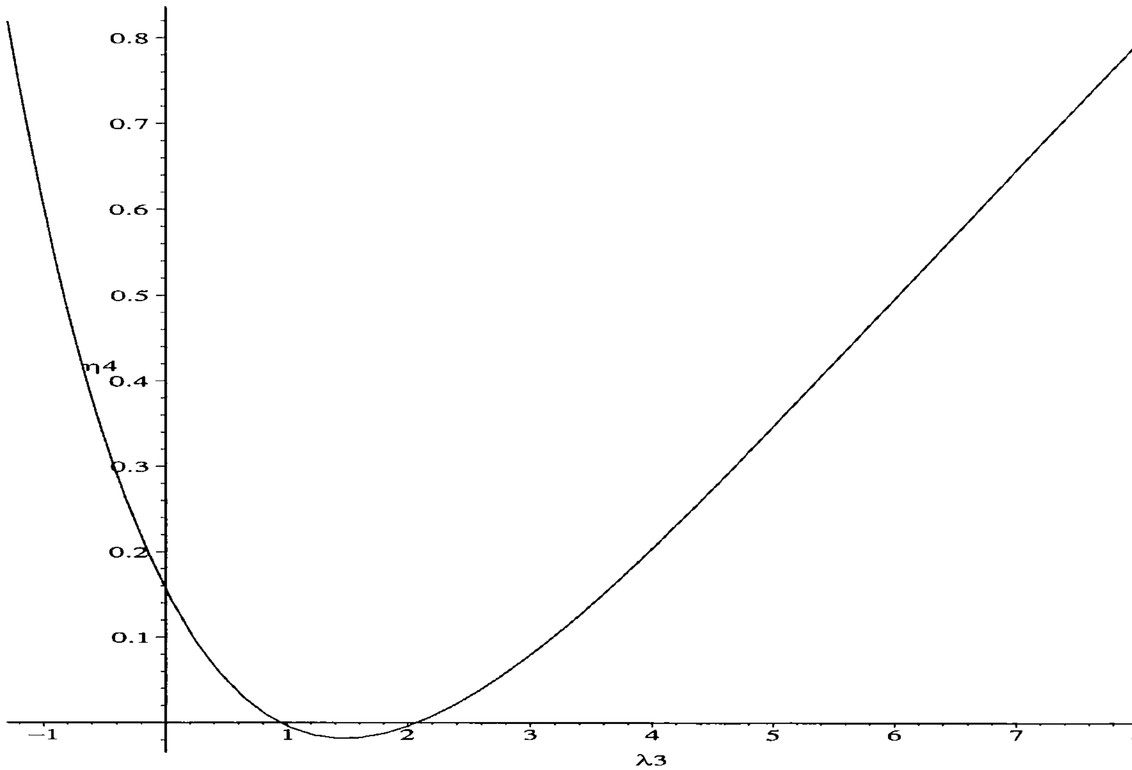

Figure 4.8: η_4 as a function of λ_3 using tri-mean, i.e $p = 0.25$ and $\alpha = 0.25$

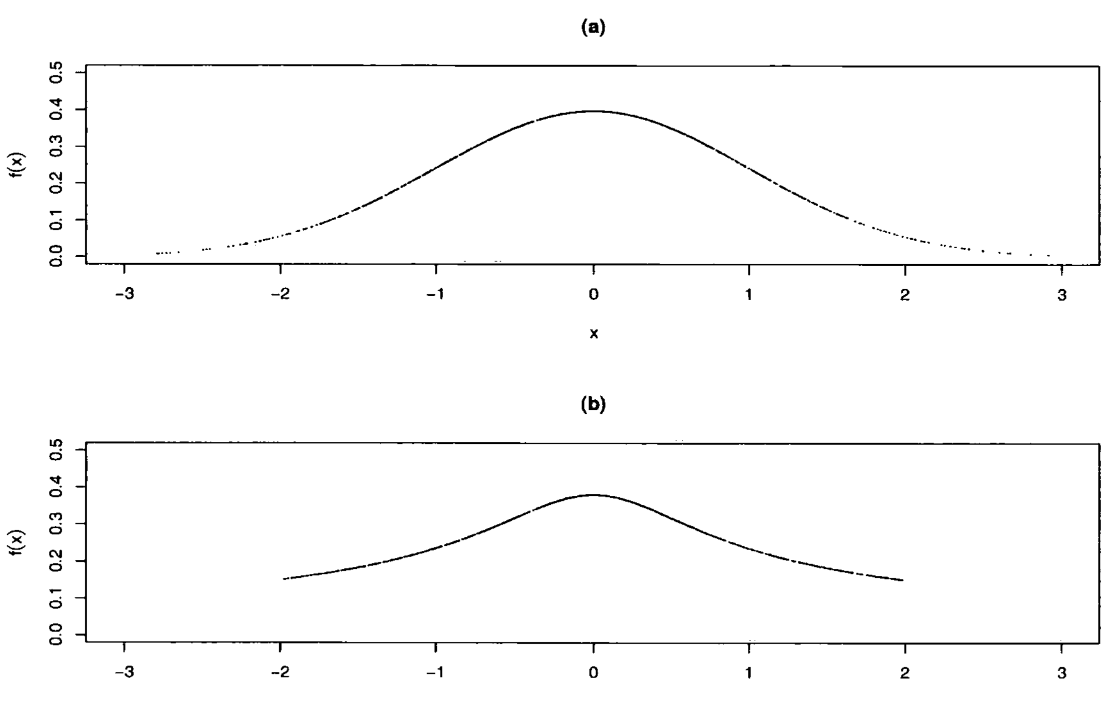

Figure 4.9: Probability density functions of the two lambda distributions having $\eta_3 = 0$ and $\eta_4 = 0.118$ and unit standard deviation

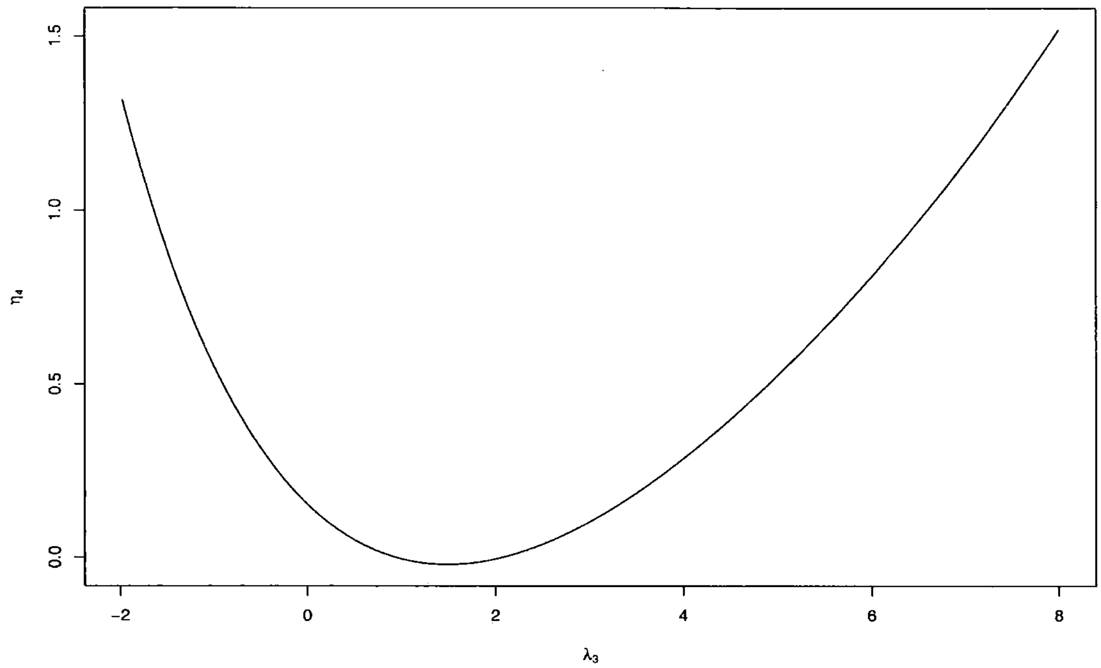

Figure 4.10: η_4 as a function of λ_3 using the median, ($p = 0.50$ and $\alpha = 0.50$).

see Figure 4.10 which shows η_4 plotted as a function of λ_3.

If we use Gastwirth ($p = 1/3$ and $\alpha = 1/3$) we find that

$$\begin{aligned}
\xi_1 &= \lambda_1 \\
\xi_2 &= \left[\frac{e^{-0.549\lambda_3} - e^{-0.8612\lambda_3} + e^{-0.346\lambda_3} - e^{-1.228\lambda_3} + e^{-0.203\lambda_3} - e^{-1.695\lambda_3}}{3\lambda_2}\right] \\
\xi_3 &= 0 \qquad (4.31)\\
\xi_4 &= \frac{1}{\lambda_2}\left\{0.167\left[e^{-0.275\lambda_3} - e^{-1.426\lambda_3} + e^{-0.173\lambda_3} - e^{-1.84\lambda_3} + e^{-0.101\lambda_3} - e^{-2.339\lambda_3}\right]\right. \\
&- \left. 0.5\left[e^{-0.666\lambda_3} - e^{-0.721\lambda_3} + e^{-0.487\lambda_3} - e^{-0.953\lambda_3} + e^{-0.345\lambda_3} - e^{-1.233\lambda_3}\right]\right\}
\end{aligned}$$

Then

$$\eta_3 = \xi_3/\xi_2 = 0 \quad \text{and} \quad \eta_4 = \xi_4/\xi_2 = \phi(\lambda_3) \qquad (4.32)$$

We notice that η_4 is a function of λ_3 only, ξ_2 is a function of λ_2 and λ_3 and ξ_1 is a function in λ_1. Solving equation (4.32) using **fsolve** in maple gives two values of λ_3 and substituting of equation (4.32) gives two values of λ_2 and we have one value of λ_1. For example, the standard normal distribution has $\eta_4 = 0.117$ and this gives two solutions $\lambda_1 = 0$, $\lambda_2 = 0.2146$ and $\lambda_3 = 0.1498$ and $\lambda_1 = 0$, $\lambda_2 = 0.5087$ and $\lambda_3 = 3.216$, see

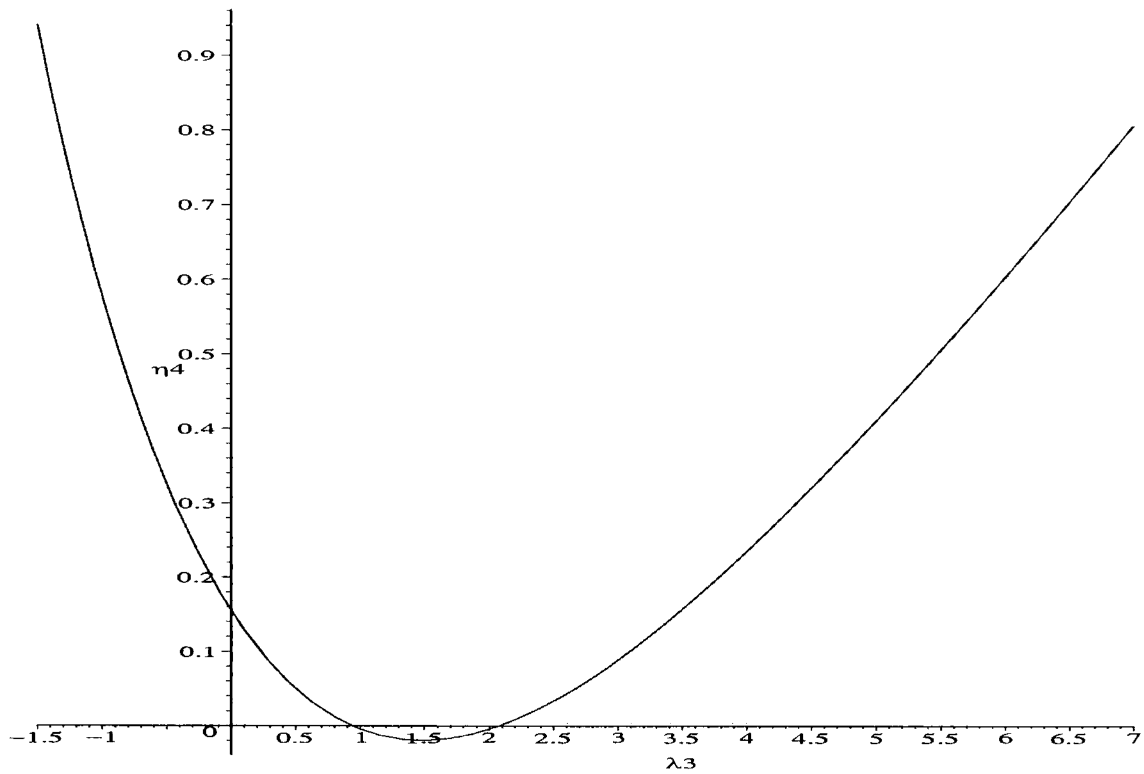

Figure 4.11: η_4 as a function of λ_3 using the Gaswirth ($p = 1/3$ and $\alpha = 1/3$).

Figure 4.11 which shows η_4 plotted as a function of λ_3.

4.3.4 TL-moments

Estimates of the symmetric lambda distribution parameters may be obtained by TL-moments defined in Chapter 3 as

$$\lambda_r^{(t)} = \sum_{k=0}^{r-1}(-1)^k \begin{pmatrix} r-1 \\ k \end{pmatrix} \mathrm{E}\left(X_{r+2t:r+t-k}\right) \tag{4.33}$$

We have the expected values of order statistics in equation (4.7). From equation (4.33) we conclude that when $t = 1$

$$\begin{aligned}
\lambda_1^{(1)} &= \lambda_1 \\
\lambda_2^{(1)} &= \frac{12\lambda_3}{\lambda_2(\lambda_3+2)(\lambda_3+3)(\lambda_3+4)} \\
\lambda_3^{(1)} &= 0 \\
\lambda_4^{(1)} &= \frac{15\lambda_3(\lambda_3-1)(\lambda_3-2)}{\lambda_2(\lambda_3+2)(\lambda_3+3)(\lambda_3+4)(\lambda_3+5)(\lambda_3+6)} \\
\tau_3^{(1)} &= 0
\end{aligned}$$

$$\tau_4^{(1)} = \frac{5(\lambda_3-1)(\lambda_3-2)}{4(\lambda_3+5)(\lambda_3+6)} = \phi(\lambda_3)$$

We notice that τ_4 is a function of λ_3. By solving this equation we have

$$\widehat{\lambda}_{31}^{(1)} = \frac{-15-44t_4^{(1)}-\sqrt{25+3380t_4^{(1)}+16t_4^{(1)2}}}{-10+8t_4^{(1)}} \tag{4.34}$$

and

$$\widehat{\lambda}_{32}^{(1)} = \frac{-15-44t_4^{(1)}+\sqrt{25+3380t_4^{(1)}+16t_4^{(1)2}}}{-10+8t_4^{(1)}} \tag{4.35}$$

Substituting in $\lambda_2^{(1)}$we obtain

$$\widehat{\lambda}_{21}^{(1)} = \frac{\widehat{\lambda}_{31}^{(1)}}{l_2^{(1)}\left(\widehat{\lambda}_{31}+2\right)\left(\widehat{\lambda}_{31}+3\right)\left(\widehat{\lambda}_{31}+4\right)} \tag{4.36}$$

$$\widehat{\lambda}_{22}^{(1)} = \frac{\widehat{\lambda}_{32}^{(1)}}{l_2^{(1)}\left(\widehat{\lambda}_{32}+2\right)\left(\widehat{\lambda}_{32}+3\right)\left(\widehat{\lambda}_{32}+4\right)} \tag{4.37}$$

Finally, we find from $\lambda_1^{(1)}$that

$$\widehat{\lambda}_1^{(1)} = l_1^{(1)} \tag{4.38}$$

The minimum value of $\tau_4^{(1)}$ for the family of symmetric lambda distributions is about -0.0064 corresponding approximately to $\lambda_3 \simeq 1.4495$. There is no upper bound on possible values of $\tau_4^{(1)}$ for this family and, as shown in Figure 4.8, there are two lambda distributions corresponding to every permissible value of $\tau_4^{(1)}$.

4.4 Approximate variances of the estimators

The estimators of the parameters of the symmetric lambda distribution depend upon the choice of the method of estimation; L-moments, LQ-moments or TL-moments. However, their mean and variances are not known exactly. A simple method of approximating mean and variance uses a Taylor expansion as follows. Suppose that we know the expectation and the variance of a random variable X and we are interested in the mean and variance of $Y = g(X)$ for some fixed function g. From Rice (1995) we find that

$$\mathrm{E}(Y) \simeq g(\mathrm{E}(X)) \tag{4.39}$$

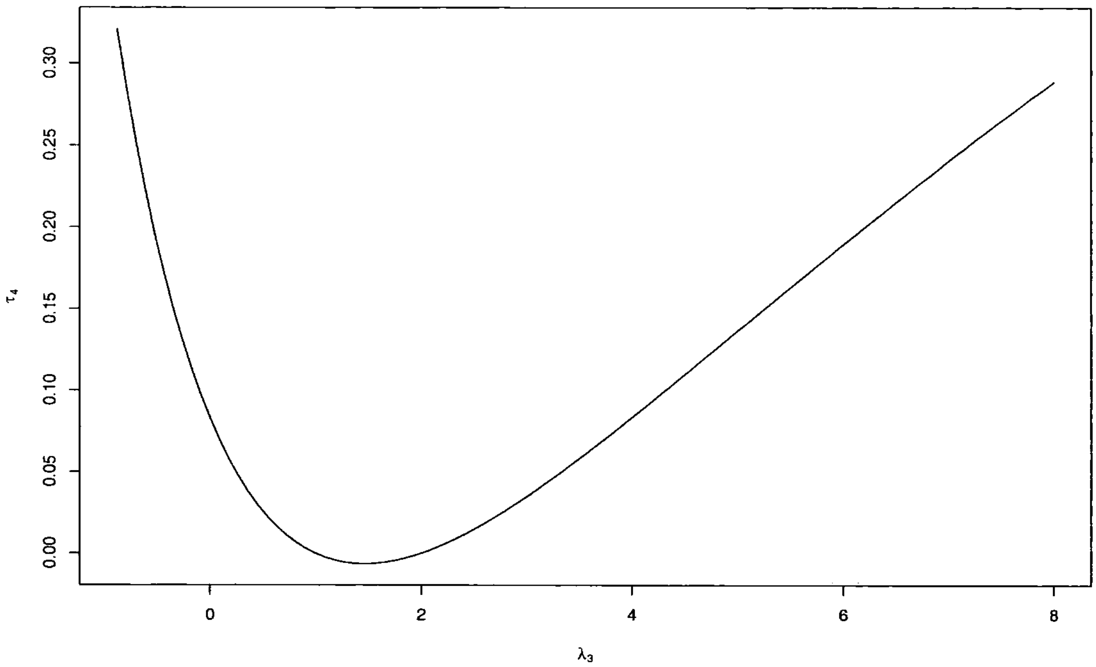

Figure 4.12: $\tau_4^{(1)}$ as a function of λ_3 using T-L moments and the minimum of $\tau_4^{(1)} \simeq -0.00644$ at $\lambda_3 \simeq 1.4495$.

and

$$\mathrm{Var}\,(Y) \simeq \mathrm{Var}\,(\mathrm{X})\,\{\mathrm{g}'\,(\mathrm{E}\,(\mathrm{X}))\}^2 \tag{4.40}$$

For example, in the case of the L-moment method, we apply equations (4.39) and (4.40) to the estimators which are given in equations (4.22), (4.23), (4.24), (4.25) and (4.26) when we consider functions $g(.)$ of the form

$$\begin{aligned} g_1(t_4) &= 3 + 7t_4 - \left(t_4^2 + 98t_4 + 1\right)^{1/2} \\ g_2\,(t_4) &= 2 - 2t_4 \\ g_3\,(t_4) &= 3 + 7t_4 + \left(t_4^2 + 98t_4 + 1\right)^{1/2} \\ g_4\,(l_2) &= C_1 l_2^{-1} \\ g_5\,(l_2) &= C_2 l_2^{-1} \\ g_6\,(l_1) &= l_1 \end{aligned}$$

with derivatives

$$g_1'(t_4) = 7 - (t_4 + 49)\left(t_4^2 + 98t_4 + 1\right)^{-1/2}$$

$$
\begin{aligned}
g_2'(t_4) &= -2 \\
g_3'(t_4) &= 7 + (t_4 + 49)\left(t_4^2 + 98t_4 + 1\right)^{-1/2} \\
g_4'(l_2) &= -C_1 l_2^{-2} \\
g_5'(l_2) &= -C_2 l_2^{-2} \\
g_6'(l_1) &= 1
\end{aligned}
$$

where

$$
C_1 = \frac{2\widehat{\lambda}_{31}}{\left(\widehat{\lambda}_{31}+1\right)\left(\widehat{\lambda}_{31}+2\right)}
$$

and

$$
C_2 = \frac{2\widehat{\lambda}_{32}}{\left(\widehat{\lambda}_{32}+1\right)\left(\widehat{\lambda}_{32}+2\right)}
$$

The approximate variances of $\widehat{\lambda}_1$, $\widehat{\lambda}_{21}$, $\widehat{\lambda}_{22}$, $\widehat{\lambda}_{31}$ and $\widehat{\lambda}_{33}$ are given by

$$
\begin{aligned}
\text{Var}\left(\widehat{\lambda}_1\right) &= \text{Var}(l_1) \\
\text{Var}\left(\widehat{\lambda}_{21}\right) &\simeq \left[g_4'(\lambda_2)\right]^2 \text{Var}(l_2) \\
\text{Var}\left(\widehat{\lambda}_{22}\right) &\simeq \left[g_5'(\lambda_2)\right]^2 \text{Var}(l_2) \\
\text{Var}\left(\widehat{\lambda}_{31}\right) &\simeq \left[\frac{g_2(\tau_4)g_1'(\tau_4) - g_1(\tau_4)g_2'(\tau_4)}{g_2^2(\tau_4)}\right]^2 \text{Var}(t_3) \\
\text{Var}\left(\widehat{\lambda}_{32}\right) &\simeq \left[\frac{g_2(\tau_4)g_3'(\tau_4) - g_3(\tau_4)g_2'(\tau_4)}{g_2^2(\tau_4)}\right]^2 \text{Var}(t_4)
\end{aligned}
$$

Note that the variances of t_3 and t_4 are approximated. Similarly, we can find the approximate means and variances of the estimators based on LQ-moments and TL-moment methods.

4.5 Application to estimating plotting position for quantile plots

In this section we use the symmetric lambda distribution family as a basis for estimating plotting position for quantile plots and make comparison with other values of plotting positions; see, for example, Kimball (1960) and Harter (1984).

We mean by a plotting positions $(p_{i:n})$ a distribution-free estimator of $F(x_{i:n})$, the nonexceedance probability of the ith order statistic from a sample size n. We consider formulae of the form

$$p_{i:n} = (i + \gamma)/(n + \delta) \quad \text{for } \delta > \gamma > -1 \tag{4.41}$$

and we restrict attention to the symmetric plotting positions $p_{i:n} = 1 - p_{n+1-i:n}$ when $\delta = 1 + 2\gamma$

Given a series of observations ordered from smallest to largest, each observation may be assigned plotting positions which is its cumulative probability. The cumulative empirical distribution function of a sample of size n is usually defined as a step function which jumps from $(i-1)/n$ to i/n at the ith order statistic of the sample. If the plotting position i/n is used, the largest value cannot be plotted, while if $(i-1)/n$ is used , the smallest value cannot be plotted, since the probabilities 1 and 0 are off the scale of the probability paper. Hazen (1914) proposed the compromise position $(i - 1/2)/n$, the value midway through the jump from $(i - 1)/n$ to i/n. Gumbel (1941) showed that the most probable (modal) position is $(i - 1)/(n - 1)$ and the mean position is $i/(n+1)$. Chernoff and Lieberman (1954) studied the use of normal probability paper, with special attention to plotting positions. They showed that the estimate of σ, in the normal distribution, based on the plotting position $i/(n + 1)$ is much less efficient than that based on the position $(i - 1/2)/n$; see also Chernoff and Lieberman (1956). Blom (1958) applied what he called " the α, β-correction" to the plotting position $i/(n + 1)$, obtaining $(i - \alpha)/(n - \alpha - \beta + 1)$. In the symmetric case, $(\alpha = \beta)$ this becomes $(i - \alpha)/(n - 2\alpha + 1)$ and he used the plotting position $(i - 3/8)/(n + 1/4)$. Tukey (1962) used the plotting position $(i - 1/3)/(n + 1/3)$.

Assume we have a model of the form

$$F(x_{i:n}) = \frac{i + \gamma}{n + 1 + 2\gamma} + \epsilon \tag{4.42}$$

We consider estimating γ by minimising the least squares criterion

$$Q = \sum_{i=1}^{n} \left[F(x_{i:n}) - \frac{i + \gamma}{n + 1 + 2\gamma} \right]^2$$

Solving

$$\frac{dQ}{d\gamma} = 2\sum_{i=1}^{n} \left[F(x_{i:n}) - \frac{i + \gamma}{n + 2\gamma + 1} \right] \left[-\frac{(n + 2\gamma + 1) - 2(i + \gamma)}{(n + 2\gamma + 1)^2} \right] = 0$$

we obtain

sample size	τ_4							
n	0	0.070	0.123	0.154	0.167	0.236	0.255	0.350
5	0	−0.230	−0.354	−0.415	−0.438	−0.542	−0.567	−0.671
8	0	−0.239	−0.366	−0.426	−0.449	−0.554	−0.579	−0.685
12	0	−0.247	−0.375	−0.436	−0.459	−0.565	−0.589	−0.697
15	0	−0.251	−0.3799	−0.441	−0.464	−0.570	−0.595	−0.703
20	0	−0.255	−0.3852	−0.446	−0.470	−0.576	−0.602	−0.710
25	0	−0.258	−0.3889	−0.450	−0.474	−0.581	−0.606	−0.715
30	0	−0.260	−0.3916	−0.453	−0.477	−0.584	−0.609	−0.718
35	0	−0.262	−0.3937	−0.456	−0.479	−0.586	−0.612	−0.721
45	0	−0.264	−0.3968	−0.459	−0.482	−0.590	−0.616	−0.725
50	0	−0.265	−0.3980	−0.460	−0.484	−0.592	−0.617	−0.727
75	0	−0.269	−0.402	−0.464	−0.488	−0.596	−0.622	−0.732
100	0	−0.270	−0.404	−0.467	−0.490	−0.599	−0.6249	−0.735
200	0	−0.273	−0.407	−0.471	−0.495	−0.603	−0.629	−0.739

Table 4.4: Estimates of γ using least squares for selected values of τ_4 of the symmetric lambda distribution for different sample sizes.

$$\widehat{\gamma} = \frac{\sum_{i=1}^{n} i\,(n+1-2i) - (n+1)\sum_{i=1}^{n}(n+1-2i)\,F\,(x_{i:n})}{2\sum_{i=1}^{n}(n+1-2i)\,F\,(x_{i:n})} \tag{4.43}$$

Note that $\widehat{\gamma}$ cannot be evaluated unless we approximate $F\,(x_{i:n})$. We do this here using the estimate $F\,[\mathrm{E}\,(X_{i:n})]$, where F is a symmetric lambda distribution; that is, we solve the equation

$$E\,(X_{i:n}) = \lambda_1 + \frac{P_{i:n}^{\lambda_3} - \left(1 - P_{i:n}^{\lambda_3}\right)}{\lambda_2}$$

with respect to $P_{i:n}$ for $i = 1, 2, ..., n$ and $\mathrm{E}\,(X_{i:n})$ is given in (4.7).

Thus

$$\widehat{\gamma} = \frac{\sum_{i=1}^{n} i\,(n+1-2i) - (n+1)\sum_{i=1}^{n}(n+1-2i)\,P_{i:n}}{2\sum_{i=1}^{n}(n+1-2i)\,P_{i:n}} \tag{4.44}$$

Table 4.4 shows values of $\widehat{\gamma}$ for selected values of τ_4 for a symmetric lambda distribution with $\mathrm{E}\,(X) = 0$ and $\mathrm{Var}\,(X) = 1$.

Since $\mathrm{E}\,(X_{i:n})$ depends on the form of the parent distribution, the choice of plotting position must be performed separately for each distribution. This is done for several plotting position formulae, see Table 4.5 where some values of τ_4 has been used, including the uniform distribution when $\tau_4 = 0$, the logistic distribution when $\tau_4 = 0.167$ and approximate normal distribution when $\tau_4 = 0.123$, and also the plots of the symmetric lambda distribution in Figures 4.13, 4.14, 4.15, 4.16 and 4.17 for sample sizes

Proponent	$p_{i:n} = \frac{i+\gamma}{n+2\gamma+1}$	γ
Weibull (1939)	$i/(n+1)$	0
Hazen (1914)	$(i-1/2)/n$	$-1/2$
Blom (1958)	$(i-3/8)/(n+1/4)$	$-3/8$
Tukey (1962)	$(i-1/3)/(n+1/3)$	$-1/3$

Table 4.5: Some common plotting position formulae

15 and 35. The general conclusions to be drawn from the plots are

I. All of them perform quite well (almost lies on the line at the tails) in cases $\tau_4 = 0$, 0.123 and 0.167.

II. On the other hand, when $\tau_4 = 0.25$ and 0.35, the least squares estimate $\widehat{\gamma}$ outperforms the other methods, especially in the tails of the plot.

III. The Hazen, Blom, Tukey and least squares formulae are quite good in most cases while the Weibull formula is not nearly as good as the others, for example when $\tau_4 = 0.35$ we find Weibull formula does not fit the line at the tails.

IV. In symmetric distributions, we recommend to obtain t_3 and t_4 from the data and using Table 4.4 to estimate γ.

4.6 Conclusions

In this chapter we have described four methods of estimating the parameters of the symmetric lambda distribution: maximum likelihood in the case of a single parameter and L-moments, LQ-moments and TL-moments in the case of three parameters.

We have also shown that the estimators are simple and in explicit form when we are using the methods of L-moments and TL-moments, while we need numerical methods for the other two methods, maximum likelihood and LQ-moments. A wide variety of curve shapes are possible with the symmetric lambda distribution as indicated by the Figures in Section 4.2. Because of this flexibility and the simplicity of the distribution it is useful as a fit to data when, as is often the case, the underlying distribution is unknown. The definition of distributions leads to a simple algorithm for generating random variates as is discussed in Section 4.2. Also, we have studied the symmetric plotting position for quantile plots based on the symmetric lambda distribution and conclude that the choice of γ depends upon the shape of the distribution, and support for this claim is borne out in our empiricalk results.

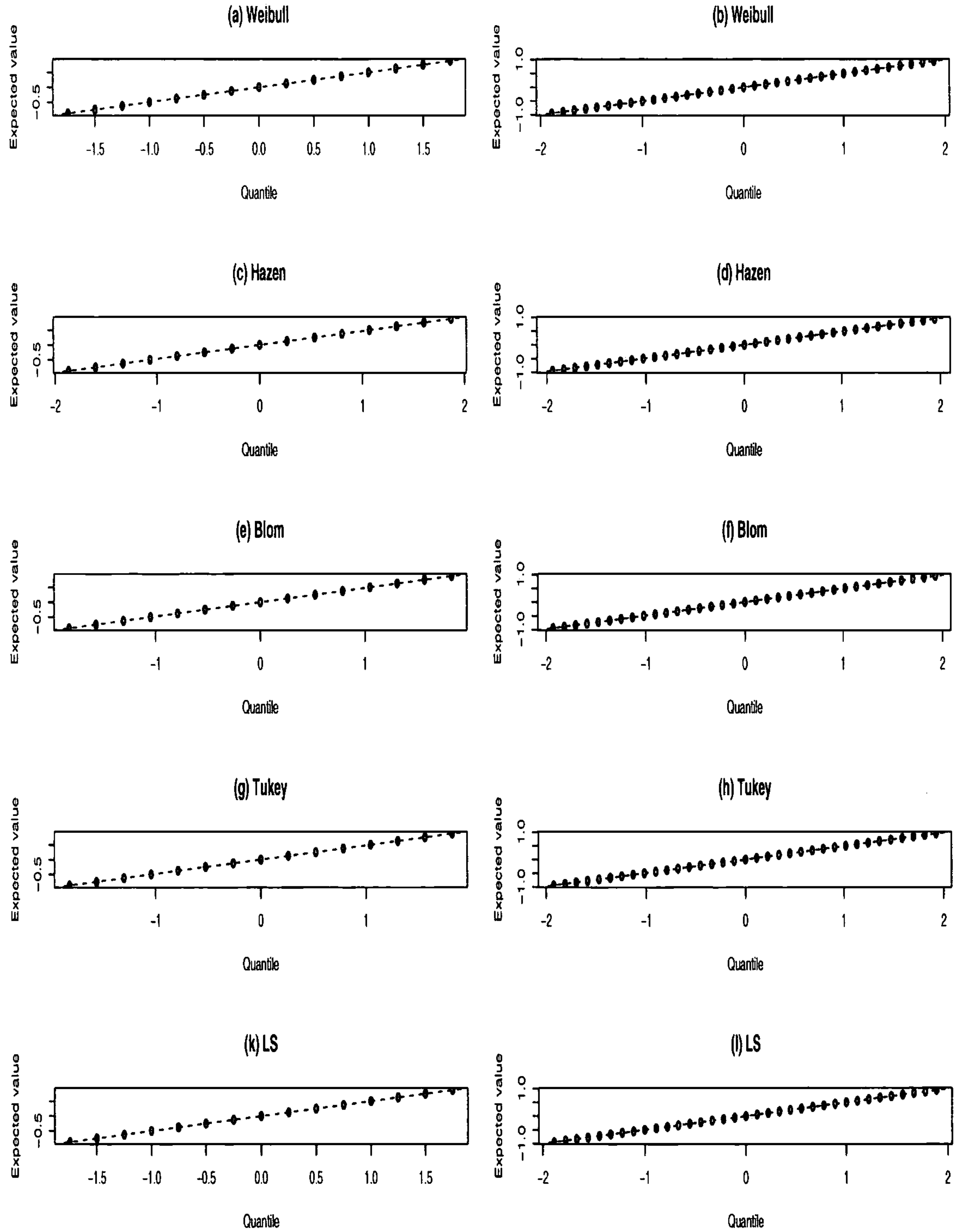

Figure 4.13: Comparison of plotting positions based on different plotting formulae when sampling from a symmetric lambda distribution with $\lambda_1 = 0$, $\lambda_2 = 0.5$, and $\lambda_3 = 1$ ($\tau_3 = 0$, $\tau_4 = 0$): (a), (c), (e), (g) and (k) for $n = 15$ and (b), (d), (f), (h) and (l) for $n = 35$.

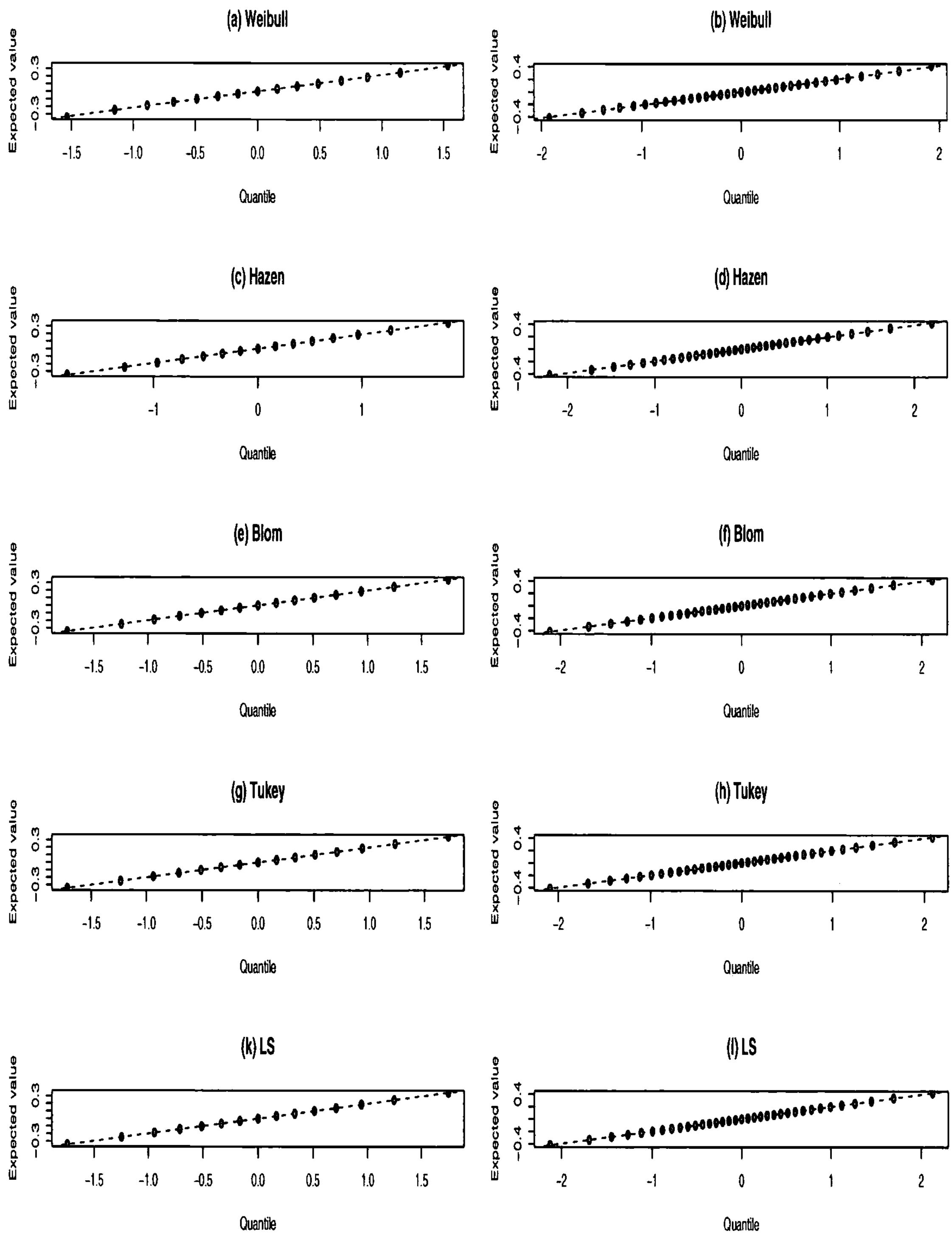

Figure 4.14: Comparison of plotting positions based on different plotting formulae when sampling from a symmetric lambda distribution with $\lambda_1 = 0$, $\lambda_2 = 0.20$, and $\lambda_3 = 0.14$ ($\tau_3 = 0$, $\tau_4 = 0.123$): (a), (c), (e), (g) and (k) for $n = 15$ and (b), (d), (f), (h) and (l) for $n = 35$.

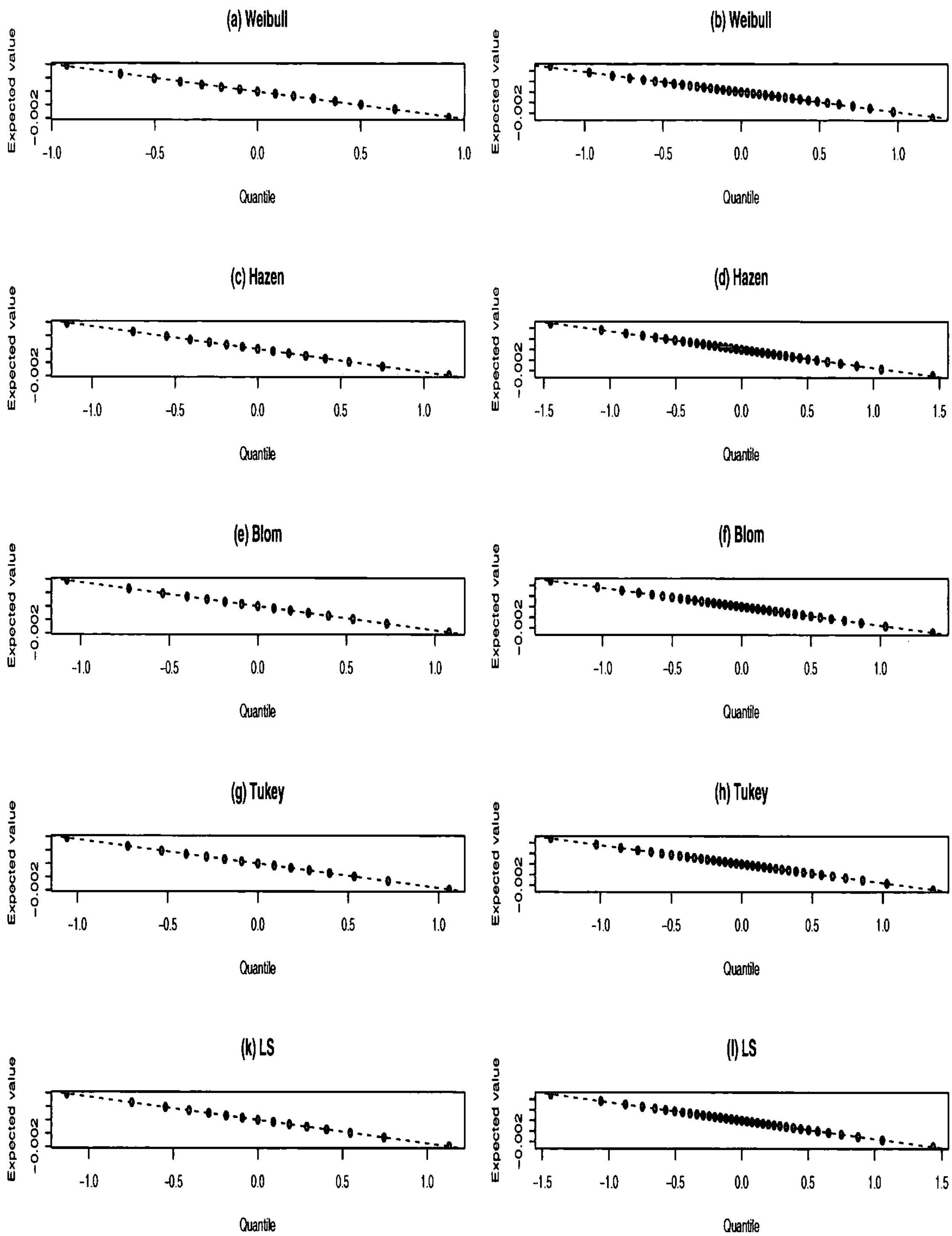

Figure 4.15: Comparison of plotting positions based on different plotting formulae when sampling from a symmetric lambda distribution with $\lambda_1 = 0$, $\lambda_2 = -0.00174$, and $\lambda_3 = -0.00059$ ($\tau_3 = 0$, $\tau_4 = 0.167$): (a), (c), (e), (g) and (k) for $n = 15$ and (b), (d), (f), (h) and (l) for $n = 35$.

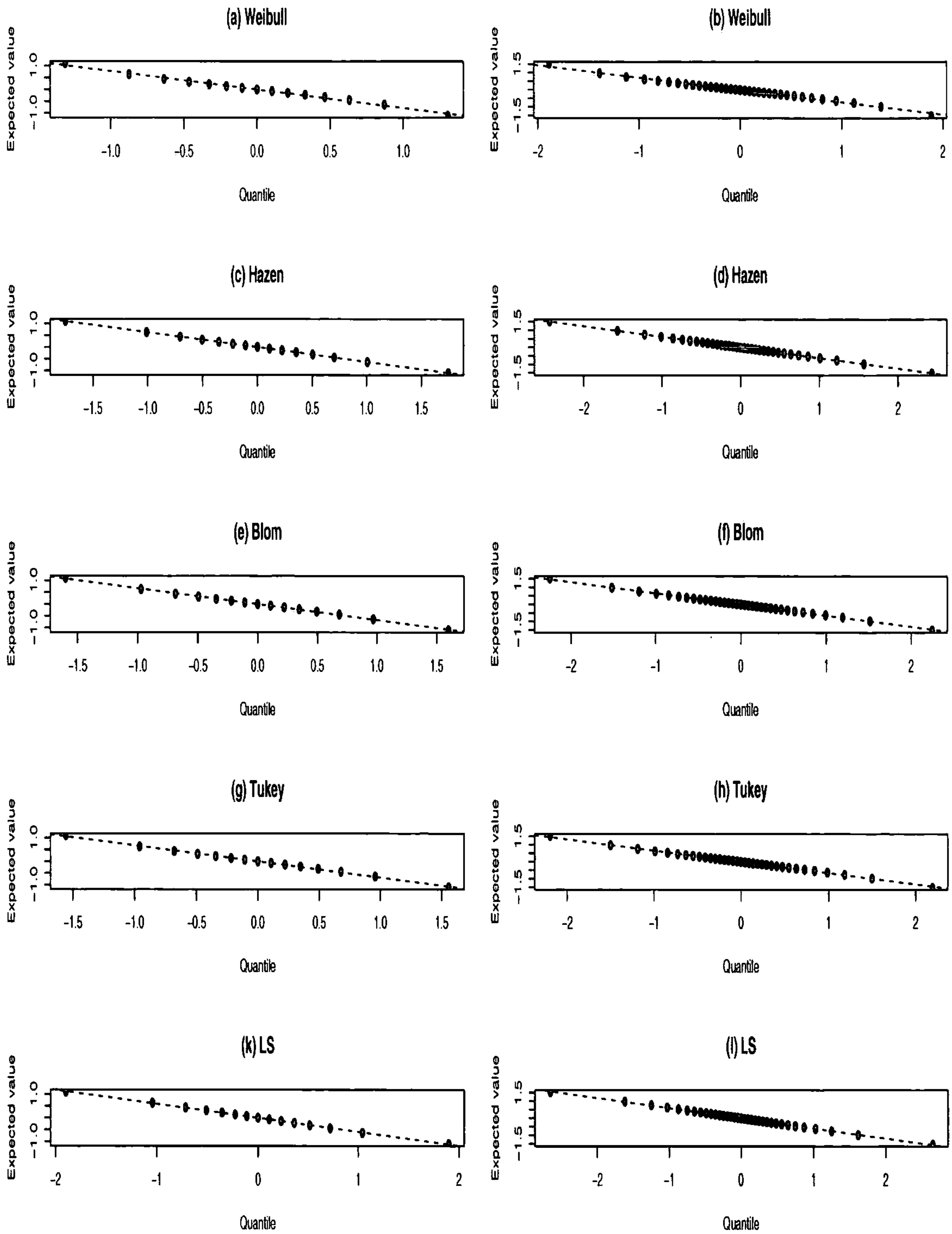

Figure 4.16: Comparison of plotting positions based on different plotting formulae when sampling from a symmetric lambda distribution with $\lambda_1 = 0$, $\lambda_2 = -0.608$, and $\lambda_3 = -0.214$ ($\tau_3 = 0$, $\tau_4 = 0.25$): (a), (c), (e), (g) and (k) for $n = 15$ and (b), (d), (f), (h) and (l) for $n = 35$.

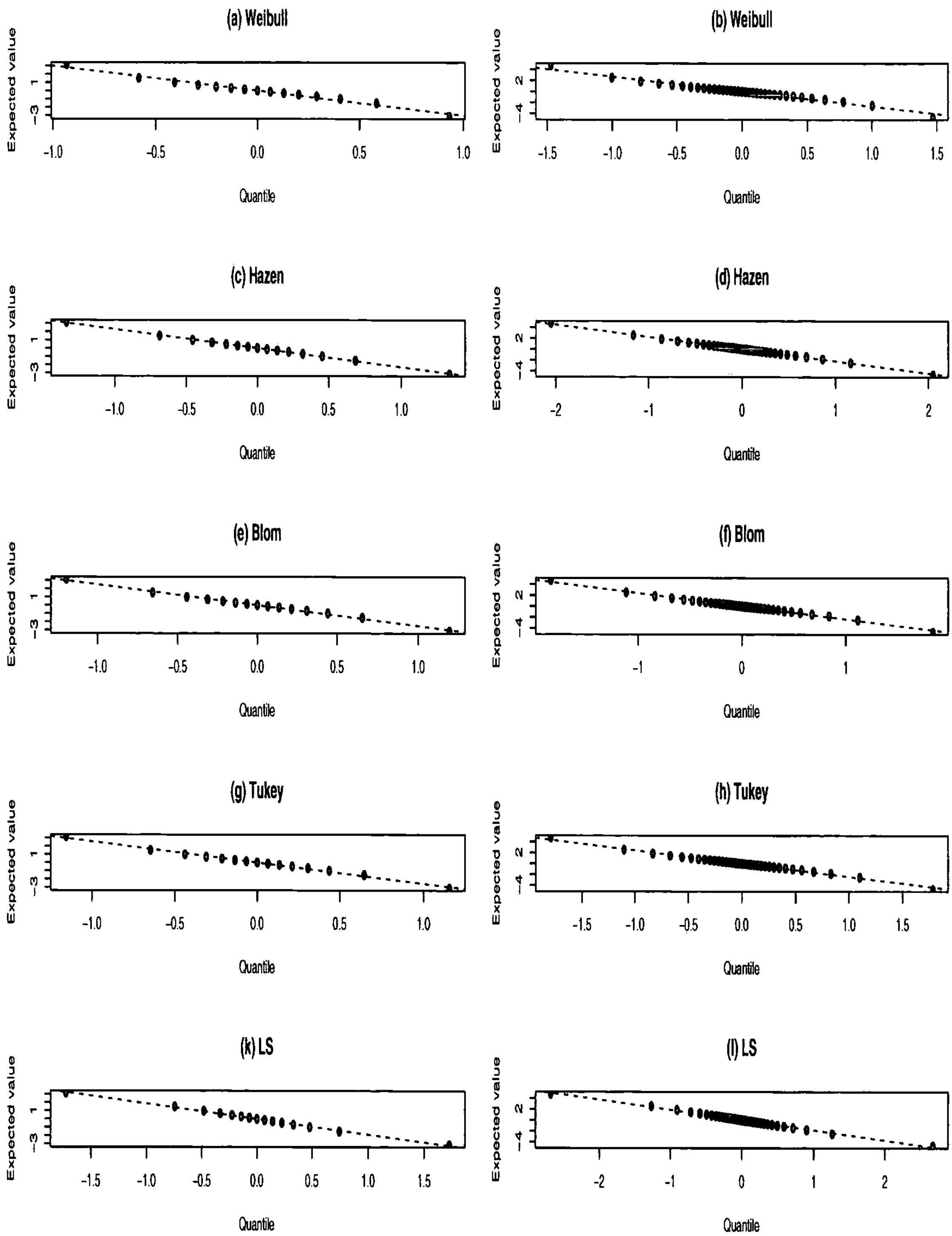

Figure 4.17: Comparison of plotting positions based on different plotting formulae when sampling from a symmetric lambda distribution with $\lambda_1 = 0$, $\lambda_2 = -2.025$ and $\lambda_3 = -0.386$ ($\tau_3 = 0$, $\tau_4 = 0.35$): (a), (c), (e), (g) and (k) for $n = 15$ and (b), (d), (f), (h) and (l) for $n = 35$.

Chapter 5

Control charts based on sample L-moments

5.1 Introduction

The usual practice in using control charts to monitor a process is to take samples from the process at fixed-length sampling intervals and plot some sample statistics on the chart. A point outside the control limits is taken as an indication that something, called "assignable cause", has happened to change the process. Since Shewhart introduced control charts in 1924, they have found widespread application in improving the quality of manufacturing processes. Another popular control procedure is the cumulative sum (CUSUM) control chart which was introduced by Page (1954). There has also been a renewed interest in the exponentially weighted moving average (EWMA) control charts, introduced by Roberts (1959) who called it a geometric moving average chart. It is known that Shewhart-type charts are relatively inefficient in detecting small changes in the process parameters; see, for example, Hunter (1986) and Montgomery (1996). On the other hand, EWMA charts have been shown to be more efficient than Shewhart-type charts in detecting small shifts in the process mean; see, for example, Ng and Case (1989), Crowder (1989), Lucas and Saccucci (1990), Amin and Searcy (1991) and Wetherill and Brown (1991). In fact, the EWMA control chart has become popular for monitoring the process mean; see Hunter (1986) for a good discussion. More recently, EWMA charts have been developed for monitoring process variability; see, for example, Macgregor and Harris (1993), Amin and Wollf (1995) and Gan (1995).

Like the Shewhart control chart, an EWMA control chart is easy to implement and interpret. It is based on the statistics

$$Z_i = \lambda X_i + (1-\lambda)Z_{i-1} \tag{5.1}$$

where X_i is the current observation, Z_0 is a starting value, such as the overall sample mean, and $0 < \lambda \leq 1$ is a constant that determines the "depth of memory" of the EWMA: The value $\lambda = 1$ gives the classical charts, such as the $\overline{X}$ chart. While the choice of λ can be left to the judgement of the quality control analyst. Experience with econometric data suggests values between 0.1 and 0.3 when it is desirable to detect small changes in whatever process characteristic is being monitored; see, for example, Hunter (1986).

Both Lucas and Saccucci (1990) and Box and Luceno (1997) give the representation

$$Z_i = \lambda \sum_{j=0}^{i-1} (1-\lambda)^j X_{i-j} + (1-\lambda)^i Z_0 \tag{5.2}$$

for an EWMA process. Thus, Z_i can be regarded as a moving average of the current and past values of the control statistics, where the weights on past data fall off exponentially as in a geometric series; and the smaller the value of λ, the greater is the influence of the past values. When the X_i are independent and identically distributed with common variance σ^2, the variance of the control statistics is given by

$$\text{Var}(Z_i) = \left\{\left[1-(1-\lambda)^{2i}\right]\lambda/(2-\lambda)\right\}\sigma^2 \tag{5.3}$$

The effect of the starting point soon dissipates and the variance increases quickly to its asymptotic value $[\lambda/(2-\lambda)]\,\sigma^2$ as i increases. Control limits are usually based on this asymptotic variance.

The presence of outliers tends to reduce the sensitivity of control chart procedures because the control limits become stretched so that the detection of outliers themselves becomes more difficult; see, for example, Rocke (1989), Tatum (1997) and Langenberg and Iglewicz (1986).

In this chapter, see also Elamir and Seheult (2001a), we propose EWMA control charts to monitor the process mean and dispersion using the Gini's mean difference and the sample mean, and also charts based on trimmed versions of the same statistics. The proposed control charts limits are less influenced by extreme observations than classical EWMA control charts, and lead to tighter limits in the presence of out-of-control observations. Specifically, these control charts and their acronyms are:

- **EWMAM:** EWMA of the sample mean to monitor the process mean, using

Gini's mean difference to estimate the process standard deviation.

- **EWMAG:** EWMA of the sample Gini's mean difference to monitor process standard deviation.
- **EWMATM**: EWMA of the sample mean to monitor the process mean, using a trimmed mean of the sample means to estimate the process mean and Gini's mean difference to estimate the process standard deviation.
- **EWMATG:** EWMA of the sample Gini's mean difference to monitor the process standard deviation using a trimmed mean of the sample Gini's mean differences to estimate the process standard deviation.

5.2 Gini's mean difference

Gini's mean difference scale estimate g is defined to be

$$g = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} |X_j - X_i| \tag{5.4}$$

where $X_1, X_2, ..., X_n$ is a random sample from a continuous distribution. It is interesting to note that g can be written alternatively as either a linear combination of the order statistics $X_{1:n} \leq X_{2:n} \leq ... \leq X_{n:n}$ or as a linear combination of the $[n/2]$ "sub-ranges" $X_{n-i+1:n} - X_{i:n}$ as follows:

$$g = \frac{2}{n(n-1)} \sum_{i=1}^{n} (2i - n - 1)\, X_{i:n} = \frac{2}{n(n-1)} \sum_{i=1}^{[n/2]} (n - 2i + 1)\,(X_{n-i+1:n} - X_{i:n}) \tag{5.5}$$

where $[x]$ is the greatest integer less than or equal to x; for example, when $n = 5$

$$g = 0.40\,(X_{5:5} - X_{1:5}) + 0.20\,(X_{4:5} - X_{2:5})$$

Downton (1966) and Barnett et al. (1967) show that, when sampling is from N (μ, σ^2)

$$G = \sqrt{\pi} l_2 \tag{5.6}$$

is an unbiased estimator of σ with variance

$$\text{Var}\,(G) = \frac{0.51n + 0.12}{n(n-1)} \sigma^2 \tag{5.7}$$

sample size	Var(G)/var(R)	Var(LS)/Var(G)
4	98.7	99,94
5	96.9	99.83
7	92.8	99.12
10	86.7	98.64

Table 5.1: Efficiency of the range and the least squares estimator of σ with respect to G for different sample size when sampling from a normal distribution.

sample size	mean	variance	skewness	kurtosis
4	1	0.180	0.254	3.145
5	1	0.134	0.1797	3.089
6	1	0.106	0.138	3.061
7	1	0.088	0.112	3.045
8	1	0.075	0.094	3.035
9	1	0.065	0.080	3.028
10	1	0.058	0.071	3.024
15	1	0.037	0.043	3.012
20	1	0.027	0.032	3.008

Table 5.2: Various moments of the sampling distribution of the standardised G/σ for different sample sizes when sampling from a normal distribution with standard deviation σ.

where $l_2 = g/2$.

There are many advantages of G:

(i) No table of coefficients of the $X_{i:n}$, such as d_n for the range, is required;

(ii) Its efficiency relative to best linear unbiased estimator of σ is very close to 100%, in small samples, decreasing to 98% asymptotically; see Table 5.1 and Downton (1966);

(iii) It is not so influenced by outliers as is either the range or the sample standard deviation;

(v) Like the range, it is simple to calculate but uses more information from the data;

(vi) Its sampling distribution is asymptotically normal as discussed in Chapter 2, as indicated in the results of simulation experiments summarised in Figures 5.1, 5.2 and in Table 5.2.

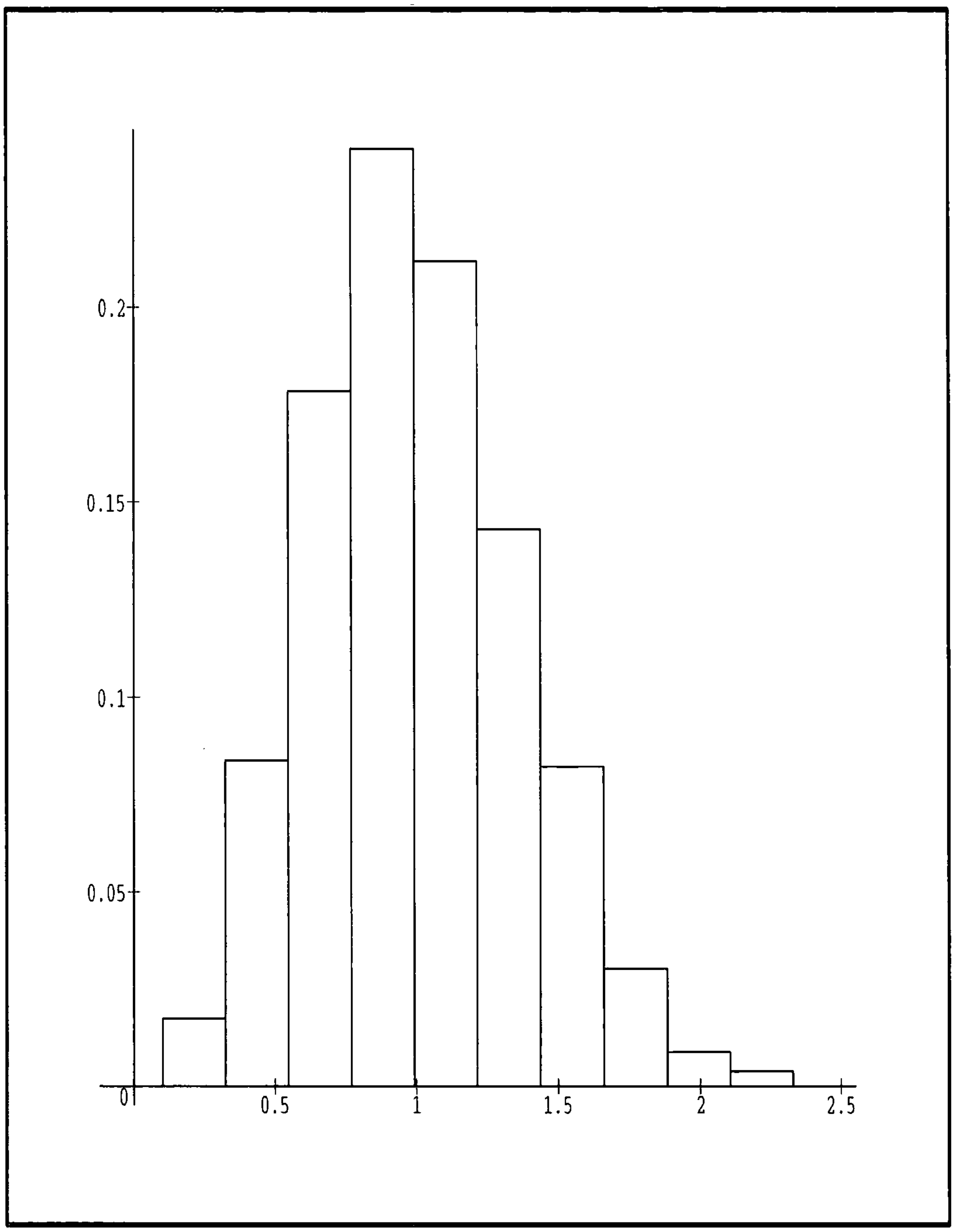

Figure 5.1: Histogram of the standardised G/σ for 5000 replications of the sample size $n = 5$ from $\mathrm{N}(0, \sigma^2)$.

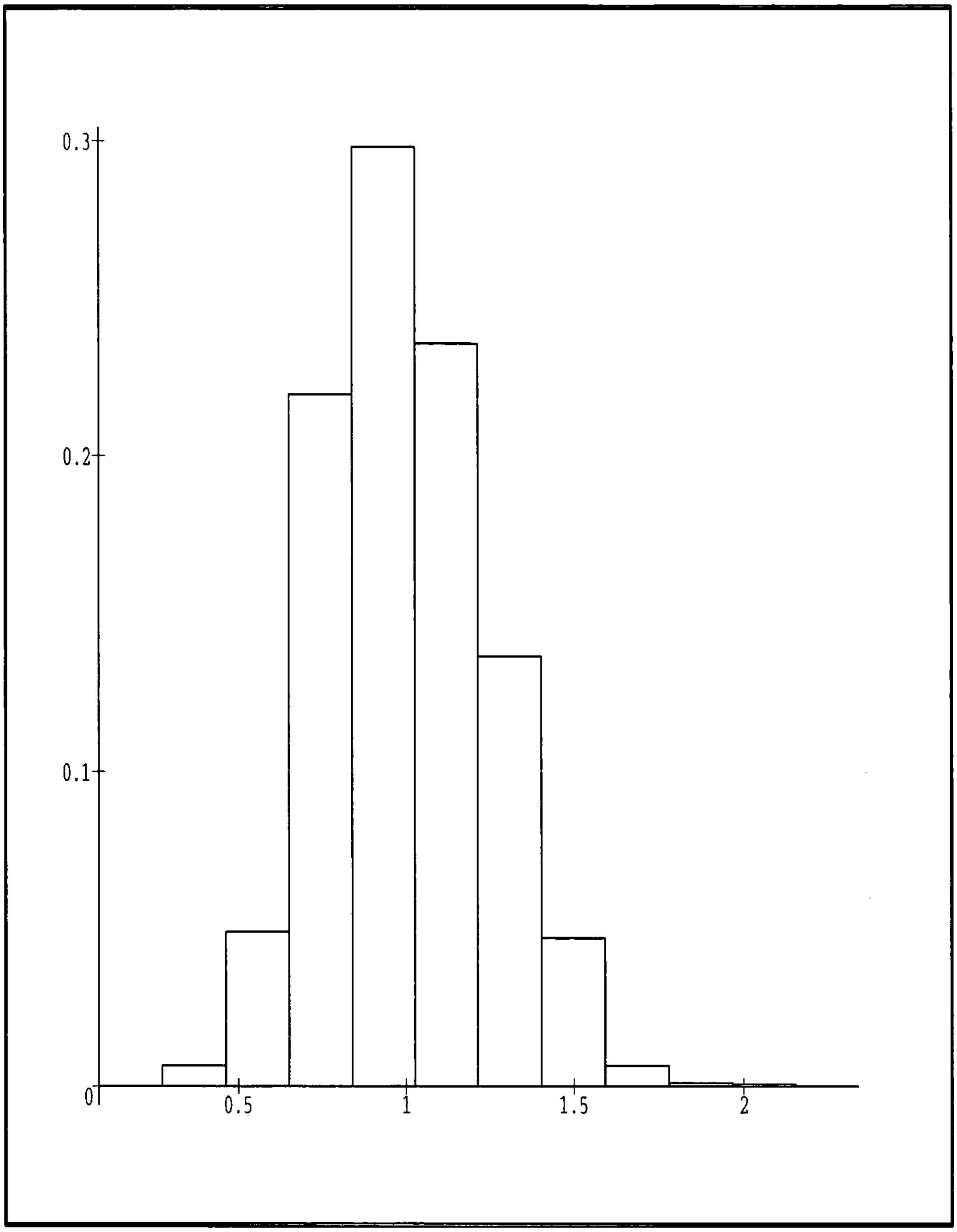

Figure 5.2: Histogram of the standardised G/σ for 3000 replications of the sample size $n = 10$ from $\mathrm{N}(0, \sigma^2)$.

5.3 Determination of control limits

When constructing EWMA control charts based on process data, m regularly spaced samples, each of size n, must be taken to estimate the process centre and spread; for example, $\overline{X_i}$ and G_i for the ith sample ($i = 1, ...m$). In general, theoretical control chart limits are often of the form

$$\begin{aligned} \text{LCL} &= \mu_T - K\sigma_T \\ \text{Centre line} &= \mu_T \\ \text{UCL} &= \mu_T + K\sigma_T \end{aligned} \tag{5.8}$$

Where T is a statistics that measures a quality characteristic, LCL and UCL the lower and upper control limits, μ_T and σ_T are the expectation and standard deviation of T, and usually $K = 3$; see, for example, Wetherill and Brown (1991). To set control limits we must estimate μ_T and σ_T based on m samples of size n obtained from the process when it is under control. In what follows, we will base our control chart limits on the assumption that when the process is under control all the observed data are independent and come from the same normal distribution, implying, in particular, that the process mean and variance are constant.

5.3.1 EWMAM limits

Each sample mean $\overline{X_i}$ is transformed into an EWMA value $Z_i = \lambda \overline{X}_i + (1 - \lambda) Z_{i-1}$ before it is plotted on the control charts: here, Z_0 is the average $\overline{\overline{X}}$ of the m sample means .

From the expressions for Var($\overline{X}_i$) and Var(Z_i), the control limits of the EWMA mean are

$$\begin{aligned} \text{LCL} &= \overline{\overline{X}} - h\overline{G} \\ \text{Centre line} &= \overline{\overline{X}} \\ \text{UCL} &= \overline{\overline{X}} + h\overline{G} \end{aligned} \tag{5.9}$$

where $h = 3\sqrt{\lambda/n(2 - \lambda)}$

5.3.2 EWMAG limits

As with EWMAM, each sample Gini's mean G_i is transformed into a EWMA value $Z_i = \lambda G_i + (1 - \lambda) Z_{i-1}$ before it is plotted on the control chart: here, Z_0 is the average

$\overline{G}$ of the m sample Gini means.

From the expression for Var(G_i) and Var(Z_i), the control limits of the EWMAG chart are

$$\begin{aligned} \text{LCL} &= (1-H)\overline{G} \\ \text{Centre line} &= \overline{G} \\ \text{UCL} &= (1+H)\overline{G} \end{aligned} \tag{5.10}$$

where $H = 3\sqrt{\lambda(0.51n+0.12)/\left(n(n-1)(2-\lambda)\right)}$

5.3.3 EWMATM and EWMATG limits

In this section the control limits of the EWMAM and EWMAG charts will be modified by replacing certain means with trimmed means. The trimmed mean $\overline{X}(\alpha)$ of m observations $X_1, ..., X_m$ is defined to be

$$\overline{X}(\alpha) = \sum_{i=r+1}^{m-r} X_{i:m}/\left(m-2r\right) \tag{5.11}$$

where $0 \leq \alpha < 0.5$ and $r = [m\alpha]$; see, for example, Hoaglin et al. (1983). Thus, the procedure involves trimming $100\alpha\%$ of the order observations from each tail and computing the mean of the remaining observations; for example, the 5% trimmed mean of 20 observations is calculated as the average of the 18 observations $X_{2:20}, ..., X_{19:20}$.

We suggest that in the control limits for the EWMAM and EWMAG, $\overline{\overline{X}}$is replaced by $\overline{\overline{X}}(\alpha)$ and $\overline{G}$ by $c\overline{G}(\alpha)$, where c (a function of m, n and α) is chosen to make $c\overline{G}(\alpha)$ an unbiased estimator of σ. We will refer to these modified EWMAM and EWMAG charts as EWMATM and EWMATG charts, respectively. Under the assumption that the process is in control and that the observations are independent and normally distributed with mean μ and variance σ^2 we find that

$$\text{E}\left(\overline{\overline{X}}\right) = \text{E}\left(\overline{\overline{X}}(\alpha)\right) = \mu \quad \text{and} \quad \text{E}\left(\overline{G}\right) = c\text{E}\left(\overline{G}(\alpha)\right) = \sigma$$

Table 5.3 contains values of c corresponding to values of m and n commonly used in the construction of the EWMA charts. Observing that we can write

$$c = \frac{\text{E}\left(\overline{G}\right)}{\text{E}\left(\overline{G}\left(\alpha\right)\right)} \tag{5.12}$$

we used simulation to obtain each value of c in Table 5.3 from 1000 replicates of mn standard normal random quantities by computing the average of the $1000\overline{G}$ values divided by the average of the $1000\overline{G}(\alpha)$ values for different values of m, n and α. We see from Table 5.3 that, for each n, the value of c depends more strongly on α than it does on m.

The steps required to compute the proposed control limits for the EWMATM and EWMATG scale charts are summarised as follows:

- Select a value of α
- Compute $r = [m\alpha]$
- Compute the $\overline{X_i}$ and G_i values, rank them and eliminate the r smallest and the r largest values
- Compute the overall trimmed L-mean and overall trimmed L-scale of μ and σ as follows:

$$\overline{\overline{X}}(\alpha) = \sum_{i=r+1}^{m-r} \overline{X}_{i:m} / (m-2r) \quad \text{and} \quad \overline{G}(\alpha) = \sum_{i=r+1}^{m-r} G_{i:m} / (m-2r)$$

- Obtain c from Table 5.3, using linear interpolation if necessary
- Compute the control limits for the process mean as

$$\begin{aligned} \text{LCL} &= \overline{\overline{X}}(\alpha) - hc\overline{G}(\alpha) \\ \text{Centre line} &= \overline{\overline{X}}(\alpha) \\ \text{UCL} &= \overline{\overline{X}}(\alpha) + hc\overline{G}(\alpha) \end{aligned} \tag{5.13}$$

  where h is as before.

- Compute the control limits for the process standard deviation σ as

$$\begin{aligned} \text{LCL} &= (1-H)\, c\overline{G}(\alpha) \\ \text{Centre line} &= c\overline{G}(\alpha) \\ \text{UCL} &= (1+H)\, c\overline{G}(\alpha) \end{aligned} \tag{5.14}$$

  where H is as before.

		α			
n	m	0.05	0.10	0.15	0.20
	20	1.006	1.013	1.020	1.024
	25	1.006	1.012	1.020	1.024
4	30	1.005	1.012	1.020	1.023
	35	1.005	1.012	1.020	1.022
	40	1.005	1.012	1.019	1.021
	20	1.006	1.011	1.019	1.021
	25	1.006	1.010	1.018	1.018
5	30	1.005	1.010	1.015	1.017
	35	1.005	1.011	1.014	1.017
	40	1.005	1.009	1.013	1.016
	20	1.006	1.009	1.012	1.015
	25	1.005	1.009	1.011	1.015
6	30	1.005	1.009	1.011	1.015
	35	1.005	1.008	1.011	1.014
	40	1.004	1.008	1.011	1.013
	20	1.003	1.009	1.012	1.014
	25	1.003	1.008	1.012	1.013
7	30	1.003	1.008	1.011	1.013
	35	1.002	1.008	1.011	1.013
	40	1.002	1.007	1.011	1.012
	20	1.002	1.006	1.010	1.013
	25	1.002	1.006	1.011	1.013
8	30	1.002	1.006	1.010	1.013
	35	1.002	1.006	1.010	1.012
	40	1.002	1.006	1.010	1.011
	20	1.002	1.005	1.011	1.013
	25	1.002	1.005	1.012	1.012
9	30	1.002	1.005	1.011	1.012
	35	1.002	1.004	1.012	1.012
	40	1.002	1.004	1.011	1.012
	20	1.002	1.004	1.012	1.011
	25	1.002	1.005	1.011	1.010
10	30	1.002	1.004	1.011	1.010
	35	1.002	1.003	1.011	1.010
	40	1.002	1.003	1.010	1.0

Table 5.3: Values of c which satisfy $\mathrm{E}\left(\overline{G}\right) = c\mathrm{E}\left(\overline{G}\left(\alpha\right)\right)$ for different values of m and n and for various choices of the trimming percentage α.

5.4 Evaluation of proposed control charts

The control limits for the proposed method are based upon the assumption that the process outcomes are normally and independently distributed. Actually, the observation generated by the process need not be normally distributed. Processes which are generated by distributions with heavier tails than the normal distribution would tend to have more than the expected number of points falling outside the control limits (EPO). Moderate departures from normality are common. We study the consequences of using the proposed method for non-normal observations.

In particular, we consider random observations generated from a $t-$distribution: a symmetric family of heavy tailed distributions which approach the normal as the degrees-of-freedom increase. Tables 5.4 and Table 5.5 contain results for independent observations generated from t-distributions with 1, 3, 8, 30 and ∞ degrees- of- freedom. The control limits for the non-trimmed and trimmed control charts in Tables 5.4 and 5.5 were computed by averaging the results of 100 replications of $m = 20$ random samples of $n = 5$ observations from each of the t-distributions. The trimmed chart limits were based on trimming 2 out of 20 sample values of $\overline{X}_i$ and $\overline{G}_i$. An additional 1000 samples of size 5 were generated from each of the $t-$distributions and the number of points falling outside the control limits was noted. For normal data, we would expect about 3 for EWMAM charts.

Tables 5.4 and 5.4 reveal that the simulated control limits for the non-trimmed and trimmed control charts are identical when the process is generated by a normal distribution. It can further be observed that the control limits for non-trimmed and trimmed control charts are approximately the same for the normal distribution and $t-$distributions with 30, 8, and 3 degrees-of-freedom. On the other hand, results for the Cauchy distribution (one degree-of-freedom) differ greatly: the trimmed charts are far tighter than the corresponding limits for the non-trimmed charts. The effect of these tighter limits is that more points fall outside for both EWMATM and EWMATG, signalling the need for appropriate corrective measures when extreme departures from normality are encountered.

5.5 Applications

In this section we give two applications to show how to use the proposed control charts.

Distribution	% trimming	$\bar{X}$ UCL	$\bar{X}$ LCL	$\bar{X}$ EPO	G UCL	G LCL	G EPO
Cauchy	0% trimmed	11.54	−9.87	43	16.30	0	65
	10% trimmed	4.72	−4.54	119	7.28	0	165
t_3	0% trimmed	1.96	−1.99	18	3.09	0	31
	10% trimmed	1.82	−1.86	23	2.88	0	42
t_8	0% trimmed	1.52	−1.53	6	2.40	0	10
	10% trimmed	1.48	−1.50	7	2.34	0	13
t_{30}	0% trimmed	1.36	−1.36	4	2.14	0	5
	10% trimmed	1.35	−1.35	4	2.13	0	6
Normal	0% trimmed	1.34	−1.34	3	2.08	0	4
	10% trimmed	1.33	−1.33	4	2.07	0	5

Table 5.4: The expected number of points falling outside the control limits (out of 1000) and UCL and LCL for samples of size 5 from various t-distribution and an EWMA weighting factor of $\lambda = 1$.

Distribution	% trimming	$\bar{X}$ UCL	$\bar{X}$ LCL	$\bar{X}$ EPO	G UCL	G LCL	G EPO
Cauchy	0% trimmed	5.07	−2.44	125	11.47	5.31	543
	10% trimmed	1.57	−1.62	285	4.95	2.33	666
t_3	0% trimmed	0.72	−0.63	30	2.07	0.96	42
	10% trimmed	0.64	−0.58	43	1.86	0.85	91
t_8	0% trimmed	0.51	−0.51	6	1.55	0.72	7
	10% trimmed	0.50	−0.49	7	1.52	0.71	9
t_{30}	0% trimmed	0.45	−0.46	4	1.40	0.65	5
	10% trimmed	0.44	−0.44	4	1.38	0.64	5
Normal	0% trimmed	0.44	−0.44	3	1.36	0.63	4
	10% trimmed	0.44	−0.43	4	1.35	0.62	5

Table 5.5: The expected number of points falling outside the control limits (out of 1000) and UCL and LCL for samples of size 5 from various $t-$distribution and an EWMA weighting factor of $\lambda = 0.20$.

Date	Aug. 1			Aug. 2			Aug. 3			Aug. 7
Shift	3	1	4	3	1	4	2	1	4	2
	218	228	280	210	243	225	240	244	238	228
	224	236	228	249	240	250	238	248	233	238
	220	247	228	241	230	258	240	265	252	220
	231	234	221	246	230	244	243	234	243	230
Date			Aug. 8			Aug. 9			Aug. 10	
Shift	4	3	1	4	3	1	4	3	1	4
	218	226	224	230	224	232	243	247	224	236
	232	231	221	220	228	240	250	238	228	230
	230	236	230	227	226	241	248	244	228	230
	226	242	222	226	240	232	250	230	246	232

Table 5.6: Subgroups of melt index measurements

5.5.1 Example 1

The data for our first example, which comes from page 207 of Wadsworth et al. (1986), is the melt index of an extrusion grade polyethylene compound measured over $m = 20$ consecutive shifts with $n = 4$ measurements per shift: see also Rocke (1989) for various robust charts for these data. Figure 5.3 shows two sets of three EWMAM $\overline{X}$ control charts for the process mean: one set, panels (a), (b) and (c) have $\lambda = 1$ (no smoothing) and the other set, panels (d), (e) and (f) have $\lambda = 0.20$ (to detect small shifts). The limits in (a) and (d) are determined from the mean and the range, (b) and (e) use the Gini mean difference, and (c) and (d) use the trimmed mean and the trimmed Gini mean difference. Similarly, Figure 5.4 gives two sets of three EWMA control charts for the process standard deviation for the same choices of λ. The control limits for (a) and (d) are determined from the range, panels (b) and (e) are determined from the Gini mean difference, and panels (c) and (f) are determined from the trimmed Gini mean difference. Also plotted on these charts are corresponding EWMA values for the same data. Panel (a) in Figure 5.3 and panel (a) in Figure 5.4 are the same as those in Figure 7-2 in Wadsworth et al. (1986).

For the three charts with $\lambda = 1$ in Figure 5.3, no points are out of limits on charts (a) and (b), while two are out on (c); one of them was identified as problematical by Wadsworth et al. (1986) using auxiliary rules. Thus, the robust control limits have detected possible problems without the use of these auxiliary rules. When $\lambda = 0.20$, out-of-control behaviour is readily apparent in all three charts (d), (e) and (f), none of them requiring much sensitivity.

Out-of-control behaviour is readily apparent in all six charts in Figure 5.4. However,

Method	$\lambda = 1$				$\lambda = 0.20$		
	LCL	CL	UCL		LCL	CL	UCL
				$\overline{X}$			
EWMAMR	221.4	235	248.7		230.5	235	239.6
EWMAMG	221.8	235	248.3		230.6	235	239.4
EWMATM	222.8	235	248.8		230.8	235	238.8
				R, G			
EWMAR	0	18.7	42.8		10.7	18.7	26.7
EWMAG	0	8.8	20		5.1	8.8	12.5
EWMATG	0	8.8	18.2		4.6	8.8	11.4

Table 5.7: Control limits for melt index data

Position	8	10	12	14	16	18	20	24
Jan. 15	31.5	31.2	31.0	31.5	31.0	30.5	31.7	30.5
	31.4	31.5	31.5	31.0	31.0	32.5	30.0	31.8
	30.8	31.0	31.0	31.0	30.0	31.5	34.0	31.0
	33.0	32.5	32.5	31.0	32.0	32.0	31.5	32.5
	7	9	11	13	17	19	21	23
Jan. 16	32.0	30.5	30.5	31.0	30.5	33.5	32.0	33.5
	31.5	32.0	32.0	29.0	30.0	32.0	30.0	30.8
	32.0	31.0	31.0	31.5	32.0	30.0	30.0	30.0
	31.8	31.0	31.0	31.0	32.5	34.5	30.5	30.5
	8	10	12	14	16	20	22	24
Jan. 17	32.5	32.5	32.5	31.5	30.8	31.5	31.0	32.5
	32.8	31.0	31.0	30.7	30.5	30.4	30.5	30.4
	32.0	30.5	30.5	29.0	30.5	31.5	31.0	30.8
	32.5	34.0	34.0	31.0	29.8	30.0	31.0	31.5

Table 5.8: Subgroups of cotton yarn data

four -points are outside the control limits in (f), twice as many as found by the standard methods in (c) or (d).

The analysis of this example supports the claim that robust methods give tighter limits and thus, in general, provide greater sensitivity than the standard control charts.

5.5.2 Example 2

The second example concerns the production of cotton yarn. Samples are taken from the spinning frames at eight positions daily. Four measurements of yarn count are obtained for each position to form subgroups of size 4.

Three day's results were used, see Table 5.8 and Wadsworth et al. (1986) page (229).

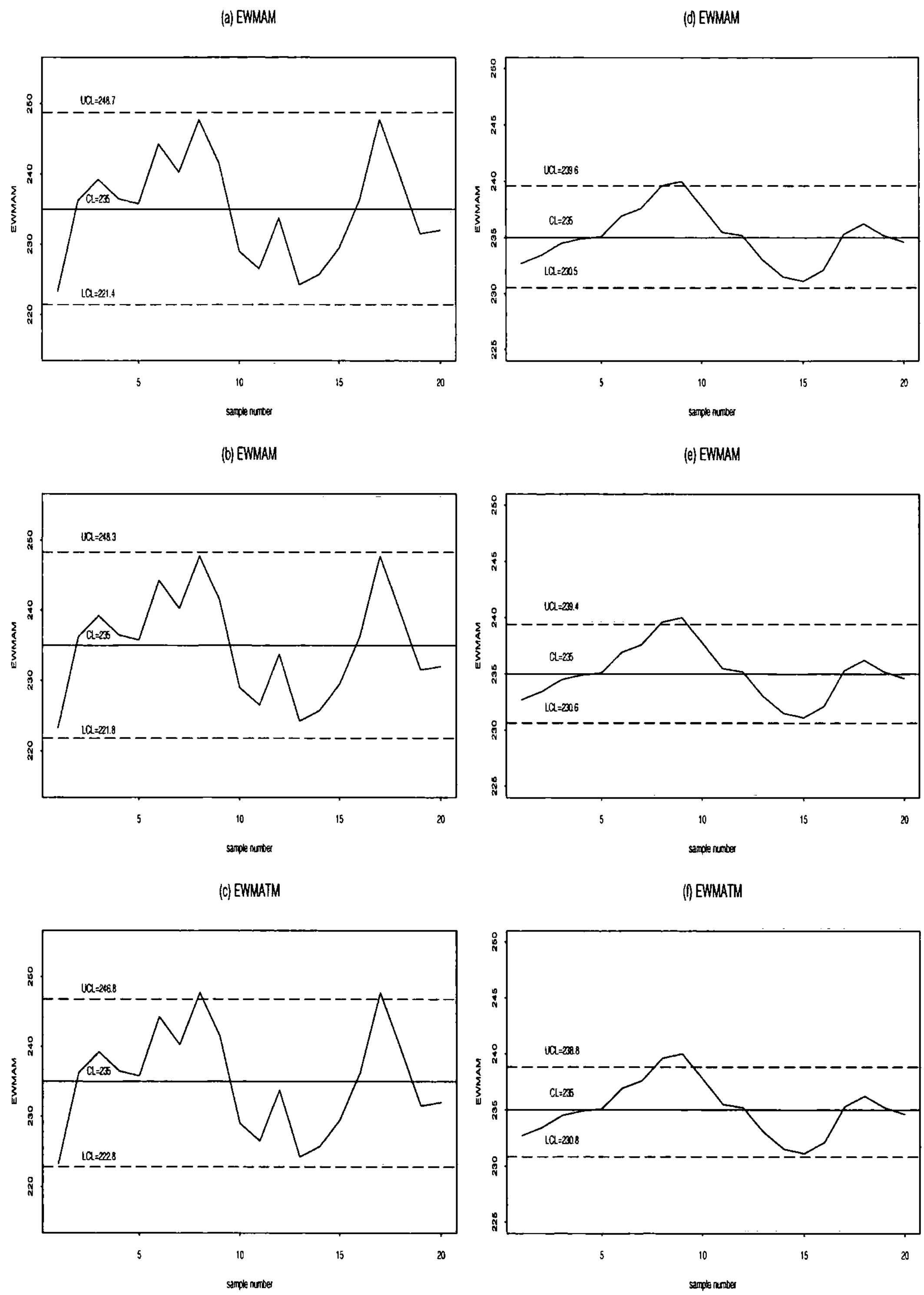

Figure 5.3: EWMA control charts for the melt index process mean μ. (a) mean, $\lambda = 1$, range estimate of σ; (b) mean, $\lambda = 1$, Gini estimate of σ; (c) trimmed mean, $\lambda = 1$, trimmed $(\alpha = 0.10)$ Gini estimate of σ; (d) mean, $\lambda = 0.20$, range estimate of σ; (e) mean, $\lambda = 0.20$, Gini estimate of σ; (f) trimmed mean, $\lambda = 0.20$, trimmed $(\alpha = 0.10)$ Gini estimate of σ;

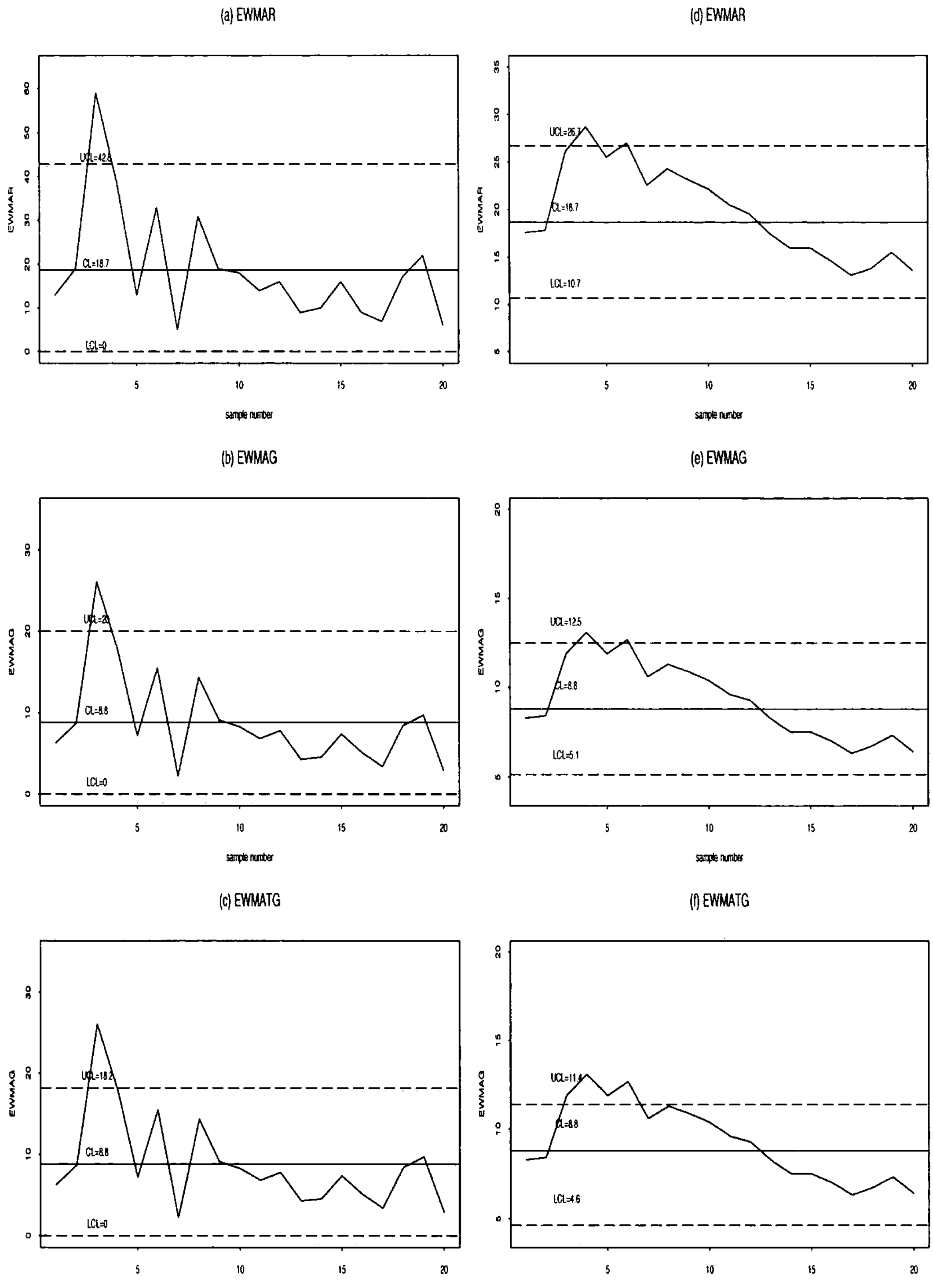

Figure 5.4: EWMA control charts for the melt index process standard deviation σ. (a) range estimate of σ; (b) Gini estimate of σ, $\lambda = 1$; (c) trimmed ($\alpha = 0.10$) Gini estimate of σ, $\lambda = 1$; (d) range estimate of σ, $\lambda = 0.20$; (e) Gini estimate of σ, $\lambda = 0.20$; (f) trimmed ($\alpha = 0.10$) Gini estimate of σ, $\lambda = 0.20$;

Method	$\lambda = 1$				$\lambda = 0.20$		
	LCL	CL	UCL		LCL	CL	UCL
				$\overline{X}$			
EWMAMR	29.8	31.3	32.8		30.8	31.3	31.81
EWMAMG	29.86	31.3	248.3		30.83	31.3	31.80
EWMATM	29.87	31.28	32.69		30.81	31.28	31.75
				R, G			
EWMAR	0	2.04	4.65		1.17	2.04	2.91
EWMAG	0	0.97	2.20		0.55	0.97	1.37
EWMATG	0	0.93	2.14		0.54	0.93	1.34

Table 5.9: Control limits for cotton yarn data

Figures 5.5 and 5.6 show the results. In both the location and spread charts, the out-of-control behaviour is readily apparent only when the trimmed control limits are used. In this example the best results, with regard to detecting out-of-control observations, were obtained using EWMATG control limits, but in the previous example the best results were obtained using the EWMAG and EWMATG control limits. Thus in some cases it may be preferable to use both methods to produce control limits.

5.6 Conclusions

In this chapter we have developed exponentially weighted moving average control charts for a process mean and standard deviation which incorporate an L-scale estimate of the process standard deviation, and we also describe trimmed versions of these charts.

We have investigated the expected number of points falling outside the control limits of these charts by simulation and conclude that the trimmed control charts are simple to use and give essentially the same limits as those computed from the non-trimmed control charts for processes generated from normal observations, and to tighter limits otherwise. Also, while the limits based on the two methods are quite similar when the process observations deviate only moderately from the normal distribution, the trimmed control charts are more likely to signal a problem when the distribution is far from normal, or when the process is out-of-control or because assignable causes have not yet been identified where we are assuming the process follows the normal ditribution. Further support for these claims are borne out in the examples.

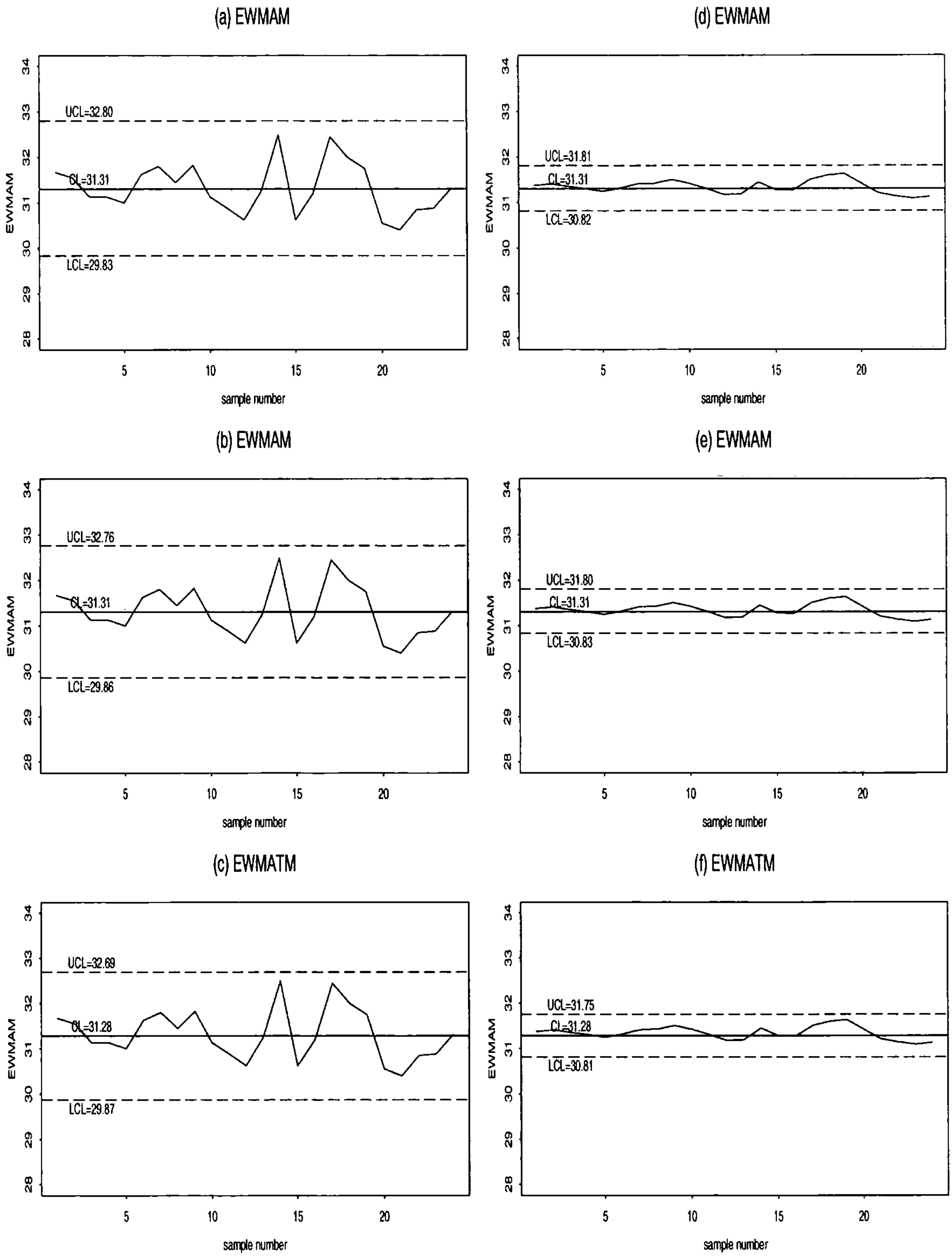

Figure 5.5: EWMA control charts for the cotton yarn process mean μ. (a) mean, $\lambda = 1$, range estimate of σ; (b) mean, $\lambda = 1$, Gini estimate of σ; (c) trimmed mean, $\lambda = 1$, trimmed ($\alpha = 0.10$) Gini estimate of σ; (d) mean, $\lambda = 0.20$, range estimate of σ; (e) mean, $\lambda = 0.20$, Gini estimate of σ; (f) trimmed mean, $\lambda = 0.20$, trimmed ($\alpha = 0.10$) Gini estimate of σ;

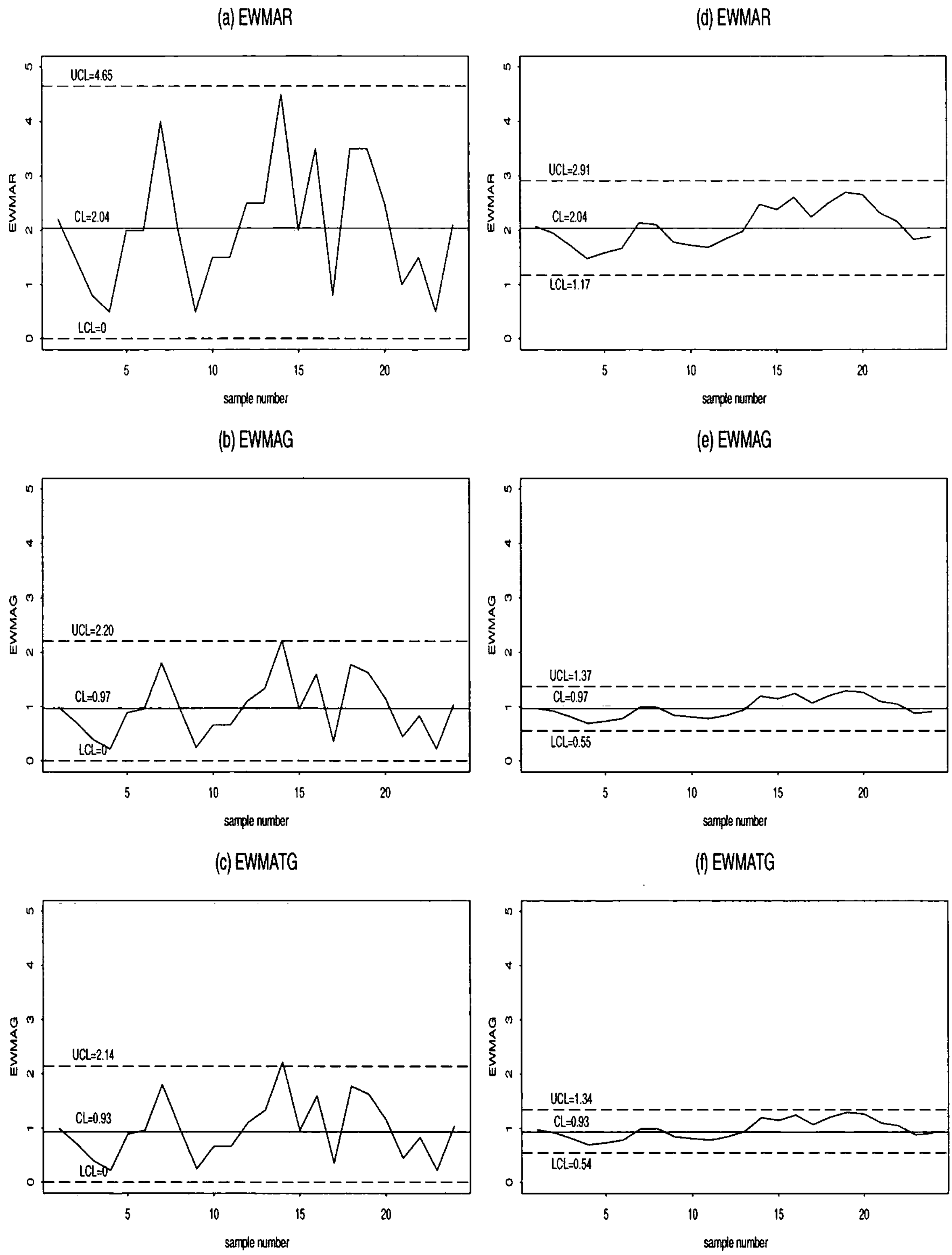

Figure 5.6: EWMA control charts for the cotton yarn process standard deviation σ. (a) range estimate of σ; (b) Gini estimate of σ, $\lambda = 1$; (c) trimmed $(\alpha = 0.10)$ Gini estimate of σ, $\lambda = 1$; (d) range estimate of σ, $\lambda = 0.20$; (e) Gini estimate of σ, $\lambda = 0.20$; (f) trimmed $(\alpha = 0.10)$ Gini estimate of σ, $\lambda = 0.20$;

Chapter 6

Conclusions and suggestions for further work

Throughout this thesis we have used L-moments and their generalisations (TL-moments and LQ-moments) to develop methods of parameter estimation and to obtain control limits in quality control for process parameters. In Chapter 2, we saw how to obtain exact variances and covariances of concepts of sample L-moments in terms of a few first and second-order moments of order statistics and how we can characterise the normal distribution in terms of their covariances.

Having reached the end of Chapter 2, one is faced with the question of "where do we go from here?". We have extended L-moments to trimmed L-moments. We have defined the sample TL-mean, TL-scale, TL-skewness and TL-kurtosis and have obtained the variances and covariances of sample TL-moments in closed form. We have also investigated the properties of TL-moments for some symmetric distributions. Also, we have shown that the TL-mean is a robust measure of location which protects against outliers. We have described the trimmed probability weighted method (TPWM) and its relation to the TL-moment method.

In Chapters 4 and 5, we apply the methods of maximum likelihood, L-moments, LQ-moments and TL-moments to estimate the parameters of the symmetric lambda distribution. Also, we have developed exponentially weighted moving average control charts for a process mean and standard deviation which incorporate an L-scale estimate of the process standard deviation, and we also describe trimmed versions of these charts.

In the light of these conclusions we see there are still many loose ends which we would like to tie up. Some of these are

- The method of maximum likelihood yields estimators which are consistent and asymptotically efficient, and for this reason is favoured by statisticians: The method of maximum likelihood should be compared with those of TL-moments and LQ-moments.

- Departures from normality. Royston (1992) argues that sample L-skewness and L-kurtosis should be used in place of the corresponding estimators of the product moment measures of skewness and kurtosis. In fact, Vogel and Fennessey (1993) argue for the replacement of the traditional moment diagram, based on $\sqrt{\beta_1}$ and β_2, by the Pearson system with a corresponding diagram based upon L-moments. Their arguments are based principally upon ease of interpretation, robustness to outliers, approximate normality and ability to indicate the type of departure from normality and ease of use in the case of censored data. We suggest that the sample TL-moments may share the same advantages.

- We can use sample TL-moments to determine control limits for a process mean and standard deviation, for example, we may consider trimming before we calculate grouped sample mean and standard deviation.

- Develop a test of normality based on the characterisation theorem given in Chapter 2.

- The method of TL-moments may also be compared with other methods which use linear combinations of order statistics to estimate the parameters of probability distributions; such as least squares, simplified linear estimates, asymptotically best linear estimates and Blom's estimates; see, David (1981, Chapter 6).

Appendix A

Covariances of the first four sample L-moments

A.1 Variances and Covariances of sample L-moments

In this appendix we give exact variances and covariances of the first four sample L-moments and all of them come from (2.40).

A.1.1 Variance of l_1

$$\mathrm{Var}\{l_1\} = \theta_{00} = \frac{\mathrm{E}\{Y_{1:1}^2\} - \mathrm{E}^2\{Y_{1:1}\}}{n}$$

A.1.2 Variance of l_2

$$\mathrm{Var}\{l_2\} = 4\theta_{11} - 4\theta_{01} + \theta_{00}$$

$$\begin{aligned}\mathrm{Var}\{l_2\} \;=\; & \left\{\frac{4}{3}(n-2)\left(\mathrm{E}\{Y_{3:3}^2\} + \mathrm{E}\{Y_{1:3}Y_{2:3}\} + \mathrm{E}\{Y_{2:3}Y_{3:3}\}\right) - 2(n-3)\mathrm{E}\{Y_{1:2}Y_{2:2}\}\right. \\ & -2(n-2)\mathrm{E}\{Y_{2:2}^2\} + (n-1)\mathrm{E}\{Y_{1:1}^2\} - 2(2n-3)\mathrm{E}^2\{Y_{2:2}\} \\ & \left.+\mathrm{E}\{Y_{1:1}\}\left(4(2n-3)\mathrm{E}\{Y_{2:2}\} - 5(n-1)\mathrm{E}\{Y_{1:1}\}\right)\right\} / \left(n\,(n-1)\right)\end{aligned}$$

A.1.3 Covariance of l_1 and l_2

$$\mathrm{Cov}\{l_1, l_2\} = 2\theta_{01} - \theta_{00}$$

$$\mathrm{Cov}\{l_1, l_2\} = \frac{(\mathrm{E}\{Y_{2:2}^2\} - \mathrm{E}\{Y_{1:2}Y_{2:2}\} - \mathrm{E}\{Y_{1:1}^2\}) + \mathrm{E}\{Y_{1:1}\}(3\mathrm{E}\{Y_{1:1}\} - 2\mathrm{E}\{Y_{2:2}\})}{n}$$

A.1.4 Variance of l_3

$$\mathrm{Var}\{l_3\} = 36\theta_{22} - 72\theta_{12} + 12\theta_{02} + 36\theta_{11} - 12\theta_{01} + \theta_{00}$$

$$\begin{aligned}
\mathrm{Var}\{l_3\} \;=\;& \Big[\frac{1}{5}\left(-18n^2 + 126n - 216\right)\left(\mathrm{E}\{Y_{3:5}Y_{4:5}\} + \mathrm{E}\{Y_{2:5}Y_{3:5}\} - 2\mathrm{E}\{Y_{5:5}^2\}\right. \\
& \left. + 5\mathrm{E}\{Y_{4:4}^2\}\right) + \left(6n^2 - 66n + 144\right)\mathrm{E}\{Y_{2:4}Y_{3:4}\} \\
+\;& (-18n + 54)(\mathrm{E}\{Y_{3:4}Y_{4:4}\} + \mathrm{E}\{Y_{1:4}Y_{2:4}\}) \\
+\;& \left(16n^2 - 108n + 176\right)\mathrm{E}\{Y_{3:3}^2\} + \left(20n^2 - 24n - 104\right)\mathrm{E}\{Y_{2:3}Y_{3:3}\} \\
+\;& \left(12n^2 - 120\right)\mathrm{E}\{Y_{1:3}Y_{2:3}\} + \left(-6n^2 + 36n - 48\right)\mathrm{E}\{Y_{2:2}^2\} \\
+\;& \left(-30n^2 + 90n + 48\right)\mathrm{E}\{Y_{1:2}Y_{2:2}\} + \left(n^2 - 3n + 2\right)\mathrm{E}\{Y_{1:1}^2\} \\
+\;& \left(-36n^2 + 180n - 240\right)\mathrm{E}^2\{Y_{3:3}\} + \left(-108n^2 + 486n - 540\right)\mathrm{E}^2\{Y_{2:2}\} \\
+\;& \left(108n^2 - 540n + 720\right)\mathrm{E}\{Y_{2:2}\}\mathrm{E}\{Y_{3:3}\} + \mathrm{E}\{Y_{1:1}\} \\
& \Big[\left(-12n^2 + 108n - 240\right)\mathrm{E}\{Y_{3:3}\} + \left(60n^2 - 288n + 336\right)\mathrm{E}\{Y_{2:2}\} \\
& + \left(-13n^2 + 39n - 26\right)\mathrm{E}\{Y_{1:1}\}\Big]\Big] / n(n-1)(n-2)
\end{aligned}$$

A.1.5 Covariance between l_1 and l_3

$$\mathrm{Cov}\{l_1, l_3\} = 6\theta_{02} - 6\theta_{01} + \theta_{00}$$

$$\begin{aligned}
\mathrm{Cov}\{l_1, l_3\} \;=\;& \Big\{\left(2\mathrm{E}\{Y_{3:3}^2\} - 2\mathrm{E}\{Y_{2:3}Y_{3:3}\} - 3\mathrm{E}\{Y_{2:2}^2\} + 3\mathrm{E}\{Y_{1:2}Y_{2:2}\} + \mathrm{E}\{Y_{1:1}^2\}\right) \\
+\;& \mathrm{E}\{Y_{1:1}\}(12\mathrm{E}\{Y_{2:2}\} - 6\mathrm{E}\{Y_{3:3}\} - 7\mathrm{E}\{Y_{1:1}\})\Big\} / n
\end{aligned}$$

A.1.6 Covariance between l_2 and l_3

$$\mathrm{Cov}\{l_2, l_3\} = 12\theta_{12} - 12\theta_{11} + 8\theta_{01} - 6\theta_{02} - \theta_{00}$$

$$\begin{aligned}
\mathrm{Cov}\{l_2, l_3\} &= \Big\{(n-3)\left(3\mathrm{E}\{Y_{4:4}^2\} + 3\mathrm{E}\{Y_{3:4}Y_{4:4}\} + 2\mathrm{E}\{Y_{2:4}Y_{3:4}\} - 6\mathrm{E}\{Y_{3:3}^2\}\right. \\
&\quad \left. -10\mathrm{E}\{Y_{2:3}Y_{3:3}\}\right) - 4\mathrm{E}\{Y_{1:3}Y_{2:3}\} + (4n-10)\mathrm{E}\{Y_{2:2}^2\} \\
&+ (8n-32)\mathrm{E}\{Y_{1:2}Y_{2:2}\} + (1-n)\mathrm{E}\{Y_{1:1}^2\} \\
&+ (24n-54)\mathrm{E}^2\{Y_{2:2}\} + (24-12n)\mathrm{E}\{Y_{2:2}\}\mathrm{E}\{Y_{3:3}\} \\
&+ \mathrm{E}\{Y_{1:1}\}\left((9n-9)\mathrm{E}\{Y_{1:1}\} + (50-26n)\mathrm{E}\{Y_{2:2}\}\right. \\
&\quad \left. +(6n-6)\mathrm{E}\{Y_{3:3}\}\right)\Big\} / \left(n(n-1)\right)
\end{aligned}$$

A.1.7 Variance of l_4

$$\begin{aligned}
\mathrm{Var}\{l_4\} &= 400\theta_{33} - 1200\theta_{23} + 480\theta_{13} - 40\theta_{03} + 900\theta_{22} - 720\theta_{12} + 60\theta_{02} \\
&\quad +144\theta_{11} - 24\theta_{01} + \theta_{00}
\end{aligned}$$

$$\begin{aligned}
\mathrm{Var}\{l_4\} &= \Big\{\frac{1}{7}\left(80n^3 - 1200n^2 + 5920n - 9600\right) \\
&\quad \times \left(5\mathrm{E}\{Y_{7:7}^2\} + \mathrm{E}\{Y_{4:7}Y_{5:7}\} + \mathrm{E}\{Y_{3:7}Y_{4:7}\}\right) \\
&+ \left(-200n^3 + 3000n^2 - 14800n + 24000\right)\mathrm{E}\{Y_{6:6}^2\} \\
&+ \left(120n^2 - 1080n + 2400\right)\left(\{Y_{4:6}Y_{5:6}\} + \mathrm{E}\{Y_{2:6}Y_{3:6}\}\right) \\
&+ \left(-20n^3 + 480n^2 - 3100n + 6000\right)\mathrm{E}\{Y_{3:6}Y_{4:6}\} \\
&+ \left(276n^3 - 4104n^2 + 19956n - 31824\right)\mathrm{E}\{Y_{5:5}^2\} \\
&+ \left(-144n^3 + 1296n^2 - 3264n + 1536\right)\mathrm{E}\{Y_{4:5}Y_{5:5}\} \\
&+ \left(-162n^3 + 1188n^2 - 882n - 5112\right)\mathrm{E}\{Y_{3:5}Y_{4:5}\} \\
&+ \left(-90n^3 + 540n^2 + 990n - 6840\right)\mathrm{E}\{Y_{2:5}Y_{3:5}\} \\
&+ (480n - 1920)\,\mathrm{E}\{Y_{1:5}Y_{2:5}\} \\
&+ \left(-190n^3 + 2760n^2 - 12890n + 19560\right)\mathrm{E}\{Y_{4:4}^2\} \\
&+ \left(330n^3 - 3420n^2 + 8130n + 1080\right)\mathrm{E}\{Y_{2:4}Y_{3:4}\} \\
&+ \left(540n^3 - 5310n^2 + 14940n - 9210\right)\mathrm{E}\{Y_{3:4}Y_{4:4}\}
\end{aligned}$$

$$
\begin{aligned}
+ & \left(-450n^2 + 900n + 3750\right) \mathrm{E}\{Y_{1:4}Y_{2:4}\} \\
+ & \left(68n^3 - 936n^2 + 3988n - 5376\right) \mathrm{E}\{Y_{3:3}^2\} \\
+ & \left(-392n^3 + 5664n^2 - 20272n + 15624\right) \mathrm{E}\{Y_{2:3}Y_{3:3}\} \\
+ & \left(48n^3 + 1224n^2 - 6432n + 2184\right) \mathrm{E}\{Y_{1:3}Y_{2:3}\} \\
+ & \left(-12n^3 + 144n^2 - 492n + 504\right) \mathrm{E}\{Y_{2:2}^2\} \\
+ & \left(-132n^3 + 1044n^2 + 9228n - 9924\right) \mathrm{E}\{Y_{1:2}Y_{2:2}\} \\
+ & \left(n^3 - 6n^2 + 11n - 6\right) \mathrm{E}\{Y_{1:1}^2\} \\
+ & \left(1600n^3 - 16800n^2 + 63200n - 84000\right) \mathrm{E}\{Y_{3:3}\}\mathrm{E}\{Y_{4:4}\} \\
+ & \left(-480n^3 + 5400n^2 - 25080n + 43200\right) \mathrm{E}\{Y_{2:2}\}\mathrm{E}\{Y_{4:4}\} \\
+ & \left(2100n^3 - 21960n^2 + 78900n - 95760\right) \mathrm{E}\{Y_{2:2}\}\mathrm{E}\{Y_{3:3}\} \\
+ & \left(-400n^3 + 4200n^2 - 15800n + 21000\right) \mathrm{E}^2\{Y_{4:4}\} \\
+ & \left(-2100n^3 + 21600n^2 - 75900n + 90000\right) \mathrm{E}^2\{Y_{3:3}\} \\
+ & \left(-864n^3 + 7416n^2 - 20664n + 18576\right) \mathrm{E}^2\{Y_{2:2}\} \\
+ & \mathrm{E}\{Y_{1:1}\} \left(\left(-25n^3 + 150n^2 - 275n + 150\right) \mathrm{E}\{Y_{1:1}\}\right. \\
& + \left(228n^3 - 2232n^2 + 6828n - 6552\right) \mathrm{E}\{Y_{2:2}\} \\
& + \left(-100n^3 + 2400n^2 - 14300n + 24000\right) \mathrm{E}\{Y_{3:3}\} \\
& \left. + \left(40n^3 - 240n^2 + 2840n - 9840\right) \mathrm{E}\{Y_{4:4}\}\right)\Big\} \\
/ & \; n\,(n-1)\,(n-2)\,(n-3)
\end{aligned}
$$

A.1.8 Covariance between l_1 and l_4

$$
\mathrm{Cov}\{l_1, l_4\} = 20\theta_{03} - 30\theta_{02} + 12\theta_{01} - \theta_{00}
$$

$$
\begin{aligned}
\mathrm{Cov}\{l_1, l_4\} = & \Big\{5\left(\mathrm{E}\{Y_{4:4}^2\} - \mathrm{E}\{Y_{3:4}Y_{4:4}\} - 2\mathrm{E}\{Y_{3:3}^2\} + 2\mathrm{E}\{Y_{2:3}Y_{3:3}\}\right) \\
+ & \; 6\left(\mathrm{E}\{Y_{2:2}^2\} - \mathrm{E}\{Y_{1:2}Y_{2:2}\} - \frac{1}{6}\mathrm{E}\{Y_{1:1}^2\}\right) \\
+ & \; \mathrm{E}\{Y_{1:1}\}\left(13\mathrm{E}\{Y_{1:1}\} - 42\mathrm{E}\{Y_{2:2}\} + 50\mathrm{E}\{Y_{3:3}\} - 20\mathrm{E}\{Y_{4:4}\}\right)\Big\} \\
/ & \; n
\end{aligned}
$$

A.1.9 Covariance between l_2 and l_4

$$\text{Cov}\{l_2, l_4\} = 40\theta_{13} - 60\theta_{12} + 24\theta_{11} - 14\theta_{01} - 20\theta_{03} + 30\theta_{02} + \theta_{00}$$

$$\begin{aligned}
\text{Cov}\{l_2, l_4\} &= \Big\{(4n-16)\left(2\text{E}\{Y_{5:5}^2\} + 2\text{E}\{Y_{4:5}Y_{5:5}\} + \text{E}\{Y_{3:5}Y_{4:5}\}\right) \\
&+ (-10n+40)\left(\text{E}\{Y_{2:4}Y_{3:4}\} + 2\text{E}\{Y_{4:4}^2\} + 3\text{E}\{Y_{3:4}Y_{4:4}\}\right) \\
&+ (18n-66)\text{E}\{Y_{3:3}^2\} + (8n-36)\text{E}\{Y_{1:3}Y_{2:3}\} + (38n-166)\text{E}\{Y_{2:3}Y_{3:3}\} \\
&+ (-7n+19)\text{E}\{Y_{2:2}^2\} + (-17n+101)\text{E}\{Y_{1:2}Y_{2:2}\} \\
&+ (n-1)\text{E}\{Y_{1:1}^2\} + (-84n+276)\text{E}\{Y_{2:2}\}^2 + (100n-280)\text{E}\{Y_{2:2}\}\text{E}\{Y_{3:3}\} \\
&+ (-40n+100)\text{E}\{Y_{2:2}\}\text{E}\{Y_{4:4}\} + \text{E}\{Y_{1:1}\}\left((-15n+15)\text{E}\{Y_{1:1}\}\right. \\
&\quad \left.+(68n-152)\text{E}\{Y_{2:2}\} + (-50n+50)\text{E}\{Y_{3:3}\} + (20n-20)\text{E}\{Y_{4:4}\}\right)\Big\} \\
&/ \quad (n(n-1))
\end{aligned}$$

A.1.10 Covariance between l_3 and l_4

$$\text{Cov}\{l_3, l_4\} = 120\theta_{23} - 180\theta_{22} + 252\theta_{12} - 36\theta_{02} - 120\theta_{13} - 72\theta_{11} + 18\theta_{01} + 20\theta_{03} - \theta_{00}$$

$$\begin{aligned}
\text{Cov}\{l_3, l_4\} &= \Big\{\left(20n^2 - 180n + 400\right)\text{E}\{Y_{6:6}^2\} + \left(-8n^2 + 72n - 160\right)\text{E}\{Y_{4:6}Y_{5:6}\} \\
&+ \left(-6n^2 + 54n - 120\right)\text{E}\{Y_{3:6}Y_{4:6}\} + \left(68n^2 - 600n + 1288\right)\text{E}\{Y_{4:4}^2\} \\
&+ \left(-60n^2 + 540n - 1200\right)\text{E}\{Y_{5:5}^2\} + \left(42n^2 - 378n + 840\right)\text{E}\{Y_{3:5}Y_{4:5}\} \\
&+ \left(24n^2 - 216n + 480\right)\text{E}\{Y_{4:5}Y_{5:5}\} + \left(18n^2 - 162n + 360\right)\text{E}\{Y_{2:5}Y_{3:5}\} \\
&+ \left(-62n^2 + 690n - 1762\right)\text{E}\{Y_{3:4}Y_{4:4}\} \\
&+ \left(-48n^2 + 570n - 1488\right)\text{E}\{Y_{2:4}Y_{3:4}\} \\
&+ (90n - 330)\text{E}\{Y_{1:4}Y_{2:4}\} + \left(-24n^2 - 180n + 1056\right)\text{E}\{Y_{1:3}Y_{2:3}\} \\
&+ \left(-36n^2 + 300n - 576\right)\text{E}\{Y_{3:3}^2\} + (-600n + 2280)\text{E}\{Y_{2:3}Y_{3:3}\} \\
&+ \left(9n^2 - 63n + 90\right)\text{E}\{Y_{2:2}^2\} + \left(300n^2 - 1980n + 3600\right)\text{E}^2\{Y_{3:3}\} \\
&+ \left(63n^2 + 27n - 1206\right)\text{E}\{Y_{1:2}Y_{2:2}\} + \left(-n^2 + 3n - 2\right)\text{E}\{Y_{1:1}^2\} \\
&+ \left(324n^2 - 1692n + 2088\right)\text{E}^2\{Y_{2:2}\}
\end{aligned}$$

$$
\begin{aligned}
+ & \left(-552n^2 + 3348n - 5568\right) \mathrm{E}\{Y_{2:2}\}\mathrm{E}\{Y_{3:3}\} \\
+ & \left(120n^2 - 540n + 600\right) \mathrm{E}\{Y_{2:2}\}\mathrm{E}\{Y_{4:4}\} \\
+ & \left(-120n^2 + 720n - 1200\right) \mathrm{E}\{Y_{3:3}\}\mathrm{E}\{Y_{4:4}\} \\
+ & \mathrm{E}\{Y_{1:1}\} \left[\mathrm{E}\{Y_{1:1}\} \left(19n^2 - 57n + 38\right) + \mathrm{E}\{Y_{2:2}\} \left(-126n^2 + 702n - 900\right)\right. \\
& \left. + \mathrm{E}\{Y_{3:3}\} \left(56n^2 - 528n + 1432\right) + \mathrm{E}\{Y_{4:4}\} \left(-20n^2 + 60n - 40\right)\right]\Big\} \\
/ & \left(n\left(n-1\right)\left(n-2\right)\right)
\end{aligned}
$$

Bibliography

Ali, S. and A. Luceno (1997). Maximum trimmed likelihood estimators: A unified approach. examples and algorithms. *Computational Statistics and Data Analysis 25*, 251–272.

Amin, R. W. and A. J. Searcy (1991). A nonparametric exponentially weighted moving average control scheme. *Communicatin in Statistics-Simulation and Computation 20*, 1049–1072.

Amin, R. W. and H. Wollf (1995). The behaviour of EWMA-type quality control schemes in the case of mixture alternatives. *Sequential Analysis 14*, 157–177.

Barnett, F. C., K. Mullen, and J. G. Saw (1967). Linear estimates of a population scale parameter. *Biometrika 54*, 551–554.

Barnett, V. (1966). Order statistics estimators of the location of the Cauchy distribution. *Journal of the American Statistical Association 61*, 1205–1218.

Barnett, V. and T. Lewis (1994). *Outliers in Statistical Data* (3rd ed.). New York: John Wiley and Sons.

Blom, G. (1958). *Statistical Estimates and Transformed Beta-variables* (1st ed.). New York: John Wiley and sons Inc.

Blom, G. (1989). Extrapolation of linear estimates to larger sample sizes. *Journal of the American Statistical Association 4*, 1139–1158.

Bowley, A. (1937). *Elements of Statistics* (6th ed.). Staples Press: London.

Box, G. and A. Luceno (1997). *Statistical Control by Monitoring and Feedback Adjustment* (1st ed.). New York: Wiley.

Chan, L. and L. Rhodin (1980). Robust estimation of location using optimally chosen sample quantile. *Technometrics 22*, 225–237.

Cheng, R. and N. Amin (1983). Estimating parameters in continuous univariate distributions with a shifted origin. *Journal of the Royal Statistical Society, B 45*, 394–403.

Chernoff, H., J. L. Gastwith, and M. V. Johns (1967). Asymptotic distribution of linear combinations of functions of order statistics with applications to estimation. *Annals of Mathematical Statistics 38*, 52–72.

Chernoff, H. and G. Lieberman (1954). Use of normal probability paper. *Journal of the American Statistical Association 49*, 778–785.

Chernoff, H. and G. J. Lieberman (1956). The use of generalised probability paper for continuous distributions. *Annals of Mathematical Statistics 27*, 806–818.

Collins, J. (2000). Robustness comparisons of some classes of location parameter estimators. *Annals of the Institute of Statistical Mathematics 52*, 351–366.

Crow, E. and M. Siddiqui (1967). Robust estimation of location. *Journal of the American Statistical Association 62*, 353–389.

Crowder, S. (1989). Design of exponentially weighted moving average schemes. *Journal of Quality Technology 21*, 155–162.

David, H. (1981). *Order Statistics* (2nd ed.). Wiley: New York.

Doodson, A. (1917). Relation of the mode, median and mean in frequency curves. *Biometrika 11*, 425–429.

Downton, F. (1966). Linear estimates with polynomial coefficients. *Biometrika 53*, 129–141.

Elamir, E. and A. Seheult (2001a). Control charts based on linear combinations of order statistics. *Journal of Applied Statistics 28*, 457–468.

Elamir, E. and A. Seheult (2001b). Exact covariances of sample L-moments. *Submitted paper.*

Elamir, E. and A. Seheult (2001c). Trimmed L-moments. *Submitted paper.*

Filliben, J. J. (1969). *Simple and Robust Linear Estimation of The Location Parameters of a Symmetric Distribution.* Ph.D thesis, Princeton University.

Gail, M. and J. L. Gastwirth (1978). A scale-free goodness-of-fit test for the exponential distribution based on the gini statstics. *Jouranl of Royal Statistics Society,B 40*, 350–357.

Gan, F. (1995). Joint monitoring of process mean and variance using exponentially weighted moving average control charts. *Technometrics 37*, 446–453.

Gastwirth, J. (1966). On robust procedures. *Journal of the American Statistical Association 61*, 929–948.

Gastwirth, J. and M. Cohen (1970). Small sample behaviour of some robust linear estimators of location. *Journal of the American Statistical Association 65*, 964–973.

Gastwirth, J. and H. Rubin (1969). On robust linear estimators. *The Annals of Mathematical Statistics 40*, 24–39.

Gingras, D. and K. Adamowski (1994). Performance of L-moments and nonparametric flood frequency analysis. *Canadian Journal of Civil Engineering 21*, 856–862.

Govindarajulu, Z. (1966a). Best linear estimates under censoring of the parameters of a double exponential population. *Journal of the American Statistical Association 61*, 248–258.

Govindarajulu, Z. (1966b). Characterisation of normal and generalised truncated normal distributions using order statistics. *Annals of Mathematical Statistics 35*, 1011–1015.

Greenwood, J. A., J. M. Landweher, and N. C. Matales (1979). Probability weighted moments: Definition and relation to parameters of several distributions expressable in inverse form. *Water Resources Research 15*, 1049–1054.

Groeneveld, R. A. (1998). A class of quantile measures for kurtosis. *The American Statistician 51*, 325–329.

Gumbel, E. J. (1941). The return period of flood flows. *Annals of Mathematical Statistics 12*, 163–190.

Gupta, A. (1967). Best linear unbiased estimators of the parameters of the logistic distribution using order statistics. *Technometrics 9*, 43–56.

Guttman, N. B., J. Hosking, and Wallis (1993). Regional precipitation of quantile values for the continental united states computed from l-moments. *Journal Climate 6*, 2326–2340.

Hampel, F., E. Ronchetti, P. Rousseeuw, and W. Stahel (1986). *Robust Statistics: The Approach Based on Influence Functions* (1st ed.). Wiley, New York, NY.

Harter, H. L. (1984). Another look at plotting positions. *Communications in Statistics-Theory and Methods 13*, 1613–1633.

Hawkins, D. (1980). *Identification of Outliers* (1st ed.). London: Chapman and Hall.

Hazen, A. (1914). Storage to be provided in impounding reservoirs for municipal water supply. *Transaction of the American Society of Civil Engineers 77*, 1539–1640.

Hinkley, D. V. (1975). On power transformation to symmetry. *Biometrika 62*, 101–111.

Hoaglin, D. C., F. Mosteller, and J. W. Tukey (1983). *Understanding Robust and Exploratory Data Analysis.* (1st ed.). John Wiley and Sons, New York, NY.

Hogg, R. (1974). Adaptive robust procedures: A partial review and some suggestions for future applications and theory. *Journal of the American Statistical Association 69*, 909–923.

Hogg, R., D. Fisher, and N. Randles (1975). A two-sample adaptive distribution-free test. *Journal of the American Statistical Association 70*, 656–661.

Hosking, J. (1986). The theory of probability weighted moments. *Research Report RC12210. IBM Research, Yorktown Heights.*

Hosking, J. R. M. (1990). L-moments: Analysis and estimation of distributions using linear combinations of statistics. *Journal of the Royal Statistical Society B 52*, 105–124.

Hosking, J. R. M. and J. Wallis (1987). Parameter and quantile estimation for the generalised pareto distribution. *Technometrics 29*, 339–349.

Hosking, J. R. M. and J. Wallis (1995). A comparison of unbiased and plotting-position estimators of L-moments. *Water Resources Research 31*, 2019–2025.

Hosking, J. R. M., J. Wallis, and E. Wood (1985). Estimation of the generalised extreme-value distribution by the method of probability-weighted moments. *Technometrics 27*, 251–261.

Huber, P. (1981). *Robust Statistics* (1st ed.). New York: Wiley.

Hunter, J. (1986). The exponentially weighted moving average. *Journal of Quality Technology 18*, 203–210.

Jaeckel, L. (1971). Robust estimates of location: Symmetry and asymmetric contamination. *Annals of Mathematical Statistics 42*, 1020–1034.

Joiner, B. and J. Rosenblatt (1971). Some properties of the range in samples from Tukey's symmetric lambda distribution. *Journal of the American Statistical Association 66*, 394–399.

Kaigh, W. D. and M. Driscoll (1987). Numerical and graphical data summary using O-statistics. *The American Statistician 41*, 25–32.

Karian, Z. and E. Dudewicz (2000). *Fitting Statistical Distributions: The Generalised Lambda Distribution and Generalised Bootstrap Methods* (1st ed.). New York: CRC.

Kendall, M. and A. Stuart (1987). *The Advanced Theory of Statistics* (5th ed.). London: Griffin.

Kimball, B. (1960). On the choice of the plotting positions on probability paper. *Journal of the American Statistical Association 55*, 546–560.

Kirby, W. (1974). Algebraic boundedness of sample statistics. *Water Resources Research 10*, 220–222.

Landweher, J. A., J. M. Matalas, and J. Wallis (1979). Probability weighted moments compared with some traditional techniques in estimating gumbel parameters and quantiles. *Water Resources Research 15*, 1055–1064.

Langenberg, P. and B. Iglewicz (1986). Trimmed mean $\bar{X}$ and R charts. *Journal of Quality Technology 18*, 152–161.

Lin, C. and G. Mudholker (1980). A simple test for normality against asymmetric alternatives. *Biometrika 67*, 455–461.

Locke, C. and J. Spurrier (1976). The use of U-statistics for testing normality against non-symmetric alternatives. *Biometrika 63*, 141–176.

Lomoniki (1951). The standard error of Gini's mean difference. *Biometrika 34*, 151–155.

Lucas, J. and M. Saccucci (1990). Exponentially weighted moving average control schemes: Properties and enhancements. *Technometrics 32*, 1–13.

Macgregor, J. F. and T. J. Harris (1993). The exponentially weighted moving variance. *Journal of Quality Technology 25*, 106–118.

Montgomery, D. C. (1996). *Introduction to Statistical Quality Control* (3rd ed.). New York, Wiley.

Mood, A., F. Graybill, and D. Boes (1974). *Introduction to the Theory of Statistics* (3rd ed.). McGraw-Hill.

Moor, D. (1968). An elementary proof of asymptotic normality of linear functions of order statistics. *Annal of Mathematical Statistics 39*, 263–265.

Mosteller, F. and J. Tukey (1977). *Data Analysis and Regression* (1st ed.). Massachusetts: Addison-Wesley, Reading.

Mudholkar, G. and A. Hutson (1998). LQ-moments: Analogs of L-moments. *Journal of Statistical Planning and Inference 71*, 191–208.

Mykytka, E. and J. Ramberg (1979). Fitting a distribution to data using an alternative to moments. *Winter Simulation Conference 1*, 361–375.

Nair, U. (1936). The standard error of Gini's mean difference. *Biometrika 23*, 428–436.

Ng, G. and E. Case (1989). Development and evaluation of control charts using exponentially weighted moving averages. *Journal of Quality Technology 21*, 242–250.

Oja, H. (1981). On location, scale, skewness and kurtosis of univariate distributions. *Scandinavian Journal of Statistics 8*, 154–168.

Page, E. (1954). Continuous inspection schemes. *Biometrika 41*, 100–114.

Pearson, C. (1993). Application of L-moments to maximum river flows. *The New Zealand Statistics 28*, 02–10.

Pearson, K. (1895). Contribution to the mathematical theory of evaluation II: Skew variation in homogeneous material. *Philosophy Transaction Royal Society London*, 343.

Pilon, P. and K. Adamowski (1992). The value of regional information to flood frequency analysis using the method of L-moments. *Canadian Journal Civil Engineering 19*, 137–147.

Ramberg, J. S., E. J. Dudewicz, P. Tadikamalla, and E. Mykytka (1979). A probability distribution and its uses in fitting data. *Technometrics 21*, 201–214.

Ramberg, J. S. and B. W. Schmeiser (1972). An approximate method for generating symmetric random variables. *Communication ACM 15*, 987–990.

Ramberg, J. S. and B. W. Schmeiser (1974). An approximate method for generating asymmetric random variables. *Communication ACM 17*, 78–82.

Rice, J. (1995). *Mathematical Statistics and Data Analysis* (2nd ed.). Duxbury Press.

Roberts, S. W. (1959). Control charts based on geometric moving averages. *Technometrics 1*, 239–250.

Rocke, D. M. (1989). Robust control charts. *Technometrics 31*, 173–184.

Rousseuw, P. and A. Leroy (1987). *Robust Regression and Outlier Detection* (1st ed.). Wiley, New York.

Royston, P. (1992). Which measures of skewness and kurtosis are best? *Statistical Medicine 11*, 333–343.

Sankarasubramanian, A. and K. Srinivasan (1999). Investigation and comparison of sampling properties of L-moments and conventional moments. *Journal of Hydrology 218*, 13–34.

Sarhan, A. and B. Greenberg (1962). *Contribution to Order Statistics* (1st ed.). New York: Wiley.

Siddiqui, M. and K. Raghunandanan (1967). Asymptotically robust estimators of location. *Journal of the American Statistical Association 62*, 950–953.

Sillito, G. (1951). Interrelations between certain linear systematic statistics of samples from any continuous distribution. *Biometrika 38*, 377–382.

Sillito, G. (1969). Derivation of approximations to the inverse distribution function of a continuous univariate population from the order statistics of a sample. *Biometrika 56*, 641–650.

Stigler, S. (1974). Linear functions of order statistics with smooth weight function. *Annals of Statistics 2*, 676–693.

Tatum, L. G. (1997). Robust estimation of the process standard deviation for control charts. *Technometrics 39*, 127–141.

Teichroew, D. (1954). Tables of expected values of order statistics and products of order statistics for samples of size twenty and less from the normal distribution. *Annals of Mathematical Statistics 26*, 410–426.

Tukey, J. (1960). A survey of sampling from contaminated distributions. *In Olkin et al. (Ed.), Contributions to probability and statistics, Stanford University Press 1*, 448–485.

Tukey, J. (1962). The future of data analysis. *Annals of Mathematical Statistics 33*, 1–67.

Tukey, J. (1970). *Exploratory Data Analysis* (1st ed.). Massachusetts: Addison-Wesley, Reading.

Vogel, R. and N. Fennessey (1993). L-moment diagram should replace product moment diagram. *Water Resource Research 29*, 1745–1752.

Wadsworth, H. M., K. Stephens, and A. B. Godfrey (1986). *Modern Methods for Quality Control and Improvement* (1st ed.). Singapore: Wiley.

Wetherill, G. B. and D. Brown (1991). *Statistical Process Control: Theory and Practice* (1st ed.). London: Chapman and Hall.

Yule, G. (1944). *Introduction to the Theory of Statistics* (13th ed.). London: Griffin.