Volume 13, pages 20-29, 2012

JOURNAL OF THE INDIAN SOCIETY FOR PROBABILITY AND STATISTICS

Estimation of Parameters in a Generalized Double Gamma Distribution

S. Rao Jammalamadaka
University of California, Santa Barbara
Kaushik Ghosh
University of Nevada, Las Vegas

ABSTRACT

In this article, we consider the problem of estimating the parameters of a generalized double gamma distribution using a method based on spacings, the gaps between successive ordered observations, and compare it with other methods that have been used in the past.

Keywords and phrases: Generalized double gamma distribution, maximum likelihood estimator, maximum product of spacings estimator, best linear unbiased estimator, method of moments estimator, location and scale parameters.

1.1 Introduction

A continuous random variable X is said to have a "generalized double gamma distribution" with parameters $(\mu, \sigma, \alpha, \beta)$ if its pdf is given by

$$f_x(x) = \frac{\beta}{3\Gamma(\alpha)\sigma^{\beta\alpha}} e^{-\left|\frac{x-\mu}{\sigma}\right|^{\beta}} |x-\mu|^{\beta\alpha-1}, -\infty < x < \infty. \tag{1}$$

We will henceforth denote this using $X \sim GDG(\mu, \sigma, \alpha, \beta)$. The parameter $\mu \in R$ is a location parameter, $\sigma > 0$ is a scale parameter and $\alpha > 0, \beta > 0$ are two shape parameters.

It can be easily checked that $X \sim GDG(\mu, \sigma, \alpha, \beta)$, if and only if

$$Y = \left|\frac{X-\mu}{\sigma}\right|^{\beta} \quad \text{Gamma}(\alpha, 1)$$

which is the standard gamma distribution with pdf

$$f_y(y) = \frac{1}{\Gamma(\alpha)} e^{-y} y^{\alpha-1}, y > 0$$

The GDG distribution is obtained by reflecting (i.e., doubling) a generalized gamma distribution (Johnson and Kotz, 1970, p. 197) about the location parameter μ, and is thus symmetric about μ.

The resulting family constitutes a very broad class of symmetric distributions, which can be either unimodal or bimodal, according as $\alpha\beta \geq 1$ or $\alpha\beta < 1$ respectively. Figure 1 gives the plots of the generalized double gamma density for $\mu = 0, \sigma = 1$ and selected values of α, β.

The GDG family (1) includes several well-known members as special cases. Setting $\alpha = \frac{1}{2}, \beta = 2$ gives the $N(\mu, \sigma^2/2$ distribution while setting $\alpha = 1, \beta = 1$ results in the standard double exponential (or Laplace) distribution DE (0, 1). Other special cases of this distribution have appeared elsewhere in the literature. Plucinska (1965), Plucinska (1966), Plucinska (1967) proposed a special subclass of (1) with $\mu = 0$, which is obtained by reflecting a generalized gamma distribution about the origin. The corresponding density is given by

$$f(x) = \frac{\beta}{2\Gamma(\alpha)\sigma^{\beta\alpha}} e^{-\left|\frac{x}{\sigma}\right|^{\beta}} |x|^{\beta\alpha - 1}, \quad -\infty < x < \infty.$$

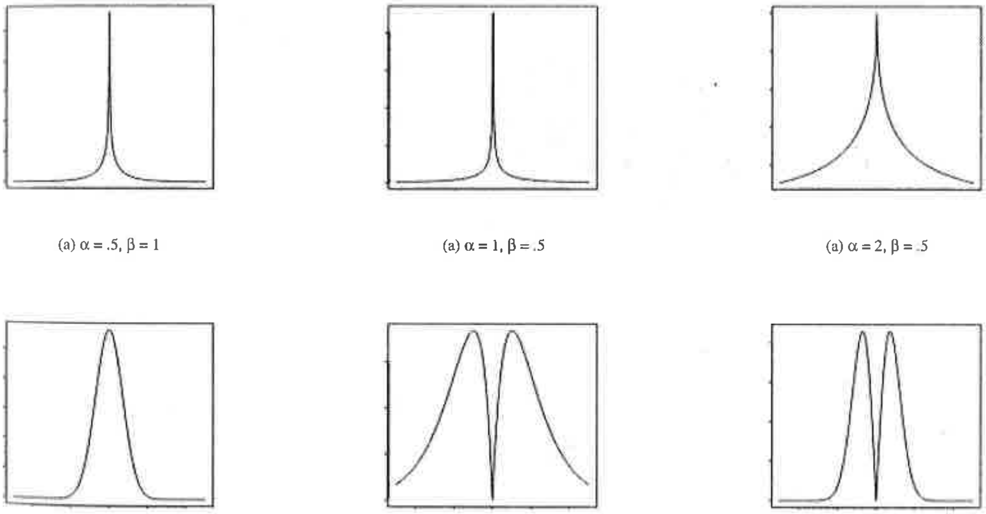
Borghi (1965) introduced the reflected gamma distribution, which is given by the pdf

$$f(x) = \frac{\beta}{2\Gamma(\alpha)\sigma^{\alpha}} e^{-\left|\frac{x-\mu}{\sigma}\right|} |x - \mu|^{\alpha - 1}, \quad -\infty < x < \infty$$

and is obtained by setting $\beta = 1$ in (1). The double Weibull distribution with density function

$$f(x) = \frac{\beta}{2\sigma^{\beta}} e^{-\left|\frac{x-\mu}{\sigma}\right|^{\beta}} ||x - \mu||^{\beta - 1}, \quad -\infty < x < \infty. \qquad (2)$$

was introduced by Balakrishnan and Kocherlakota (1985), and is obtained by setting $\alpha = 1$ in (1).



(a) $\alpha = .5, \beta = 1$ (a) $\alpha = 1, \beta = .5$ (a) $\alpha = 2, \beta = .5$

(a) $\alpha = .5, \beta = 2$ (a) $\alpha = 2, \beta = 1$ (a) $\alpha = 1, \beta = .2$

Figure 1: Density of the generalized double gamma distribution for $\mu = 0$, $\sigma = 1$ and for various values of α, β.

Members of the GDG family are used in various modeling situations. The normal distribution is of course the most widely used distribution in statistics for modeling symmetric bell-shaped phenomenon. The Laplace distribution is commonly used as an alternative to the normal distribution in robustness studies (see Tiao and Lund, 1970, Andrews et al., 1972). It has also been used in astronomy, biological and environmental sciences, engineering, finance and quality control (see Nadarajah, 2004). Sornette et al. (2000), Malevergne et al. (2005), Malevergne et al. (2006), Malevergne and Sornette (2004), Malevergne and Sornette (2005) have used the modified Weibull distribution

$$f(x) = \frac{\beta}{2\sqrt{\pi}\sigma^{\beta/2}} e^{-\left|\frac{x}{\sigma}\right|^{\beta}} |x|^{\beta/2-1}, \quad -\infty < x < \infty$$

to model asset returns in finance. It is obtained by setting $\mu = 0, \alpha = \frac{1}{2}$ in (1). Other notable uses of members of the GDG family of distributions have been in engineering, such as modeling brittle and ductile strengths (Nadarajah and Kotz, 2008), modeling system reliability (Plucinska, 1967), modeling speech amplitudes and adjacent-sample differences in video signals (Jayant and Noll, 1984), (Eggerton and Srinath, 1986, Joshi and Fischer, 1995, Lam and Goodman, 2000, Chang et al., 2005).

In this article, we investigate a new method of estimation of the location and scale parameters of the GDG family. The rest of the article is organized as follows. The estimation procedure is proposed in Section 2. A small-scale simulation study is used to compare the proposed estimator to some previously known estimators in Section 3. Use of the proposed method is illustrated with an example in Section 4. The article ends with some concluding remarks in Section 5.

1.2 Parameter Estimation

Given a random sample $X_1, \ldots, X_n$ from GDG$(\mu, \sigma, \alpha, \beta)$, our goal is to estimate the parameters μ and σ in (1), where α, β are assumed to be known. The most popular method of parameter estimation in statistics, is of course, maximum likelihood (ML). However, when $\alpha\beta < 1$, this method breaks down when applied to (1), since the likelihood function becomes unbounded at $\mu = X_i$, for $i = 1, \ldots, n$, making estimation of σ inconsistent. Thus, we need to consider alternate estimation procedures.

Balakrishnan and Kocherlakota (1985) proposed estimation of μ and σ in (1) using linear combinations of the order statistics. They provided tables of coefficients of order statistics needed to construct best linear unbiased estimator (BLUE) of μ and σ for $\beta\{0.5, 0.75, 2, 3\}$ and $n = 2(1)(10)$. Later, Rao and Narasimham (1989) extended the tables of Balakrishnan and Kocherlakota (1985) to $n = 11(1)20$ for both complete and censored data. Kantam and Narasimham (1991) discussed best linear unbiased estimation of μ and σ in GDG$(\mu, \sigma, \alpha, 1)$ again for complete and censored data and provided tables for the coefficients when $n = 2(1)10$ for $\alpha\{2, 3\}$. The disadvantage of these methods is that the coefficients used in the linear combinations are cumbersome to obtain and the estimators will likely have poor behavior when compared to the MLE if the latter is available.

We propose to use an alternative procedure, called the Maximum Product of Spacings (MPS) method, which produces asymptotically equivalent estimators as the ML method but avoids problems such as non-uniqueness and unboundedness that is sometimes encountered by the ML method. See Ghosh and Jammalamadaka (2001) and the references contained therein for details about estimation based on spacings, and Ranneby (1984) for the special case of maximum product of spacings estimator.

The MPS method proceeds as follows. Let be $X_{(1)} < X_{(2)} < \ldots < X_{(n)}$ the order statistics based on a random sample from a distribution with cdf $F(.|\Theta)$, where Θ is the vector of unknown parameters. We then define the function

$$g(\Theta) = \prod_{i=1}^{n+1} \left[F(X)_{(i)}|\Theta) - F(X_{(i-1)}|\Theta)\right],$$

where $X_{(0)} \equiv -\infty$ and $X_{(n+1)} \equiv \infty$. The quantity $g(\Theta)$ is the product of the gaps (or spacings) between consecutive values of the transformed order statistics arising from the sample. The MPS estimator of Θ is given by the maximizer of $g(\Theta)$. For our case, we assume that the two shape parameters are completely known and thus $\Theta = (\mu, \sigma)$.

The estimator is motivated by the fact that when Θ_0 is the true value of the parameter vector, by the probability integral transform, the quantities $F(X_{(i)}|\Theta_0) - F(X_{(i-1)}|\Theta_0)$ are distributed as spacings from U(0,1), which, on average, partition the interval (0,1) into equal segments. Maximizing $g(\Theta)$ is an attempt to find the Θ that bring all the $F(X_{(i)}|\Theta) - F(X_{(i-1)}|\Theta)$ close $\frac{1}{n+1}$, since it is well known that subject to the restriction that $a_i > 0$ and $\sum_{i=1}^{n+1} a_i = 1$, the quantity $\prod_{i=1}^{n+1} a_i$ is maximized at $a_i = \frac{1}{n+1}$.

While the maximizer will typically not have an analytical closed-form solution, one can perform a numerical maximization to arrive at the desired answer for any given data set (note that this is also often the case with ML estimation). However, since the function $g(\Theta)$ is always bounded, this method will always give rise to a valid estimator and not suffer from the drawback faced by ML method where the likelihood function may become unbounded (such as when $\alpha\beta < 1$ in the GDG family). In addition, as discussed in Ghosh and Jammalamadaka (2001), the resulting estimator enjoys all the nice asymptotic properties of the ML estimator when the latter exists.

One may also use the method of moments (MOM) to estimate μ and σ. It is easily shown that if $X \sim GDG(\mu, \sigma, \alpha, \beta)$, then

$$E(X) = \mu, \quad \text{and} \quad \mathrm{Var}(X) = \sigma^2 \frac{\Gamma\left(\alpha + \frac{2}{\beta}\right)}{\Gamma(\alpha)}.$$

Hence, the method of moments estimators of μ and σ are

$$\mu_{\mathrm{MOM}} = \overline{X}, \quad \text{and} \quad \sigma_{\mathrm{MOM}} = S\sqrt{\frac{\Gamma(\alpha)}{\Gamma\left(\alpha + \frac{2}{\beta}\right)}},$$

where $\overline{X}$ and S are the sample mean and standard deviation respectively.

1.3 A Simulation Study on the Performance of Estimators

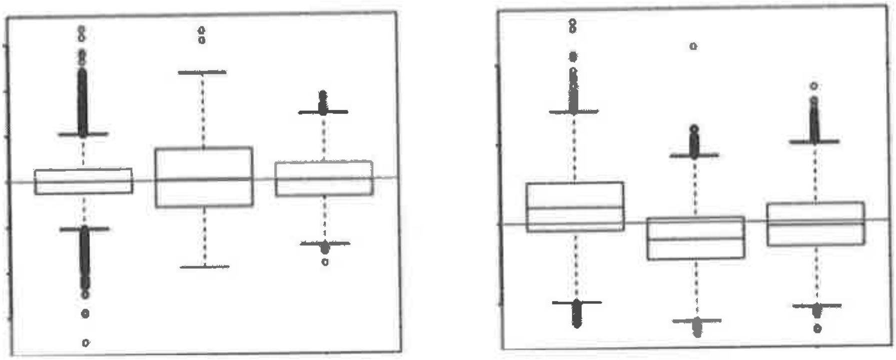
We compare properties of the MPSE, the MLE (when available), the BLUE and MOME of μ and σ using a brief simulation study. The study is conducted as follows: For a specific

$(\mu, \sigma, \alpha, \beta)$ combination, we first generate random samples of size $n = 10$ from GDG$(\mu, \sigma, \alpha, \beta)$ distribution (the choice of this sample size was somewhat dictated by the availability of BLUE coefficients). We then use the samples to estimate (μ, σ) using the four methods, assuming (α, β) are known. This procedure is repeated 5000 times, to obtain estimates of the respective root mean squared errors (RMSE), which are reported in Table 1.

Table 1: Estimated root mean squared error (RMSE) in estimating μ (corresponding figures for estimating σ are in parentheses) using the different methods when $n = 10$. The RMSE calculations are based on 5000 simulations. Lower RMSE indicates better performance. Note that MLE does not exist when $\alpha\beta < 1$. However, when MLE exists, it performs roughly the same as MPSE, which is always more efficient than BLUE. The calculations for BLUE are based on coefficients from Balakrishnan and Kocherlakota (1985) and Kantam and Narasimham (1991).

$(\mu, \sigma, \alpha, \beta)$	MPSE	MLE	BLUE	MOM
(0,1,1,.5)	0.30 (1.51)	NA	5.27 (2.17)	1.58 (0.69)
(0,1,1,.75)	0.37 (0.73)	NA	1.48 (1.21)	0.64 (0.46)
(0,1,1,2)	0.36 (0.25)	0.42 (0.18)	0.48 (1.07)	0.31 (0.18)
(0,1,1,3)	0.42 (0.28)	0.44 (0.19)	0.52 (1.11)	0.30 (0.14)
(0,1,2,1)	0.86 (0.32)	0.94 (0.24)	1.06 (1.08)	0.78 (0.24)
(0,1,3,1)	1.24 (0.26)	1.54 (0.22)	1.60 (1.07)	1.08 (0.20)
(0,1,4,1)	1.57 (0.23)	2.17 (0.21)	2.24 (1.07)	1.41 (0.18)
(0,1,5,1)	2.06 (0.22)	2.87 (0.22)	2.92 (1.07)	1.76 (0.16)

From the table, it is obvious that the BLUEs perform uniformly worse than the other two methods of estimation, both in estimating μ and σ. The MPSE has an advantage for estimating μ when $\alpha\beta < 1$ (note that in this case, the MLE does not even exist). For estimating μ when $\alpha\beta > 1$, MOM gives slightly better estimates than MPS, which already improves upon the ML method. For estimating σ, MOM seems to be performing better than ML, which is also better than MPS. An idea of the comparative sampling distributions of the resulting estimators can be obtained from the side-by-side boxplots of the estimators of μ and σ using three of the methods based on samples from GDG(0,1,3,1) given in Figure 2.



(a) Estimating μ (b) Estimating σ

Figure 2 Boxplots of estimates of μ and σ using MPS, ML and MOM. Estimates were obtained from random samples of size n = 10 from GDG(0, 1, 3, 1). Results are based on 5000 simulations.

1.4 A Practical Example

Video compression is widely used in many applications such as transmission of high-definition TV signals and storage of surveillance footage. Most video compression is currently based on so-called hybrid coding, which is motion compensation followed by a two-dimensional spatial transform. An alternative method is to view a video signal as 3-dimensional — two spatial dimensions and one time-dimension — and apply the 3-dimensional discrete cosine transform (DCT). This method was first proposed by Roese et al. (1977), but due to high computational cost and memory requirements compared to hybrid coding, has not been in much use. However, with availability of cheap computational power and memory, there has been a renewed interest in 3D DCT. More details and further references can be found in Bhaskaranand (2008).

Knowledge of the distribution of the 3D DCT coefficients is essential in designing optimal quantizers for the 3D DCT method. To this effect, Bhaskaranand (2008) undertook a study of the distributional properties of the 3D DCT coefficients arising from the luminance components of videos with low or structured motion. Eight different test sequences were considered, from which samples of 3D DCT coefficients were obtained. Table 2 provides a random sample of 100 such coefficients.

Table 2: A random sample of 100 coefficients from discrete cosine transform (DCT) of a video signal

59.27	−43.01	1.22	−241.91	28.32
8.27	3.94	−18.54	−29.93	53.08
22.59	−16.01	9.88	−2.16	−53.75
20.42	46.90	−29.53	−10.49	−4.45
6.05	50.26	−1.55	−27.32	0.50
113.17	1.89	10.68	18.72	6.99
−417.62	101.99	−72.72	−8.10	56.00
10.09	−38.88	−55.94	7.56	−3.96
18.64	−14.00	19.45	29.62	−118.44
−2.44	37.13	195.62	−2.13	54.09
12.38	0.00	29.26	−11.96	16.62
76.59	−1.29	−8.94	−32.33	17.56
0.18	−8.77	3.70	10.44	−7.45
33.56	43.73	7.74	10.29	10.74
−0.04	2.94	19.47	12.54	36.93
6.32	74.41	7.30	51.26	80.76
−74.40	0.46	4.76	0.31	1.41
−543.44	7.01	−9.01	99.73	19.74
−4.44	−22.65	−1.65	8.41	−91.95
−24.75	35.84	−6.79	−32.93	-70.03

The Gaussian, Laplace, double-gamma and Rayleigh distributions are commonly used for modeling DCT coefficients. The four possible contenders were

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right), -\infty < x < \infty, \tag{3}$$

$$f_X(x) = \frac{\lambda}{12} \exp\left(-\lambda|x-\mu|\right), -\infty < x < \infty, \tag{4}$$

$$f_X(x) = \frac{\sqrt[4]{3}}{\sqrt{8\pi\sigma|x-m|}} \exp\left(\frac{-\sqrt{3}|x-\mu|}{2\sigma}\right), -\infty < x < \infty, \tag{5}$$

and

$$f_X(x) = \frac{x}{\sigma^2} \exp\left(\frac{-x^2}{2\sigma^2}\right), x > 0. \tag{6}$$

In engineering literature, the family (5) has been found to be a good fit for such data, a fact re-confirmed through various goodness of fit tests in Bhaskaranand (2008). This is the density of $\mathrm{GDG}\left(\mu, \frac{2\sigma}{\sqrt{3}}, \frac{1}{2}, 1\right)$. Note that $\alpha\beta = \frac{1}{2} < 1$, and hence the ML method does not work for estimation of μ, σ. We thus need to resort to the MPS method, which gives $\mu_{MPS} = 3.74, \sigma_{MPS} = 74.14$.

Note also that for this distribution, since $E(X) = \mu$ and $\mathrm{Var}(X) = \sigma^2$, as an alternative, one can use the method of moments to get $\mu_{\mathrm{MOM}} = \overline{X} = -4.31$ and $\sigma_{MOM} = S = 84.38$. The corresponding density plots as well as the histogram of the data are given in Figure 3. Close examination of Figure 3 indicates that the estimated density from MPS provides a better fit to the observed data. This can also be confirmed with a formal test of goodness-of-fit.

1.5 Concluding Remarks

We have proposed an alternative method of estimating μ and σ in the generalized double gamma distribution when α and β are known. This method performs very well in estimating μ when compared with other competing methods, and provides better estimates of the underlying distribution. From the simulation studies, it appears that MOM estimate of σ is better but this could be an artifact of the restricted sample size that we chose for these simulations. It would be interesting to perform more extensive simulations using different sample sizes, but one would then have to exclude the BLUE from consideration because of the limited availability of the coefficients needed.

Even when all the four parameters of the distribution in (1) are unknown, the proposed MPS method can be easily adapted to estimate all of them simultaneously. It is not possible to do so with the other approaches like the method of moments, the maximum likelihood and the best linear unbiased estimation methods.

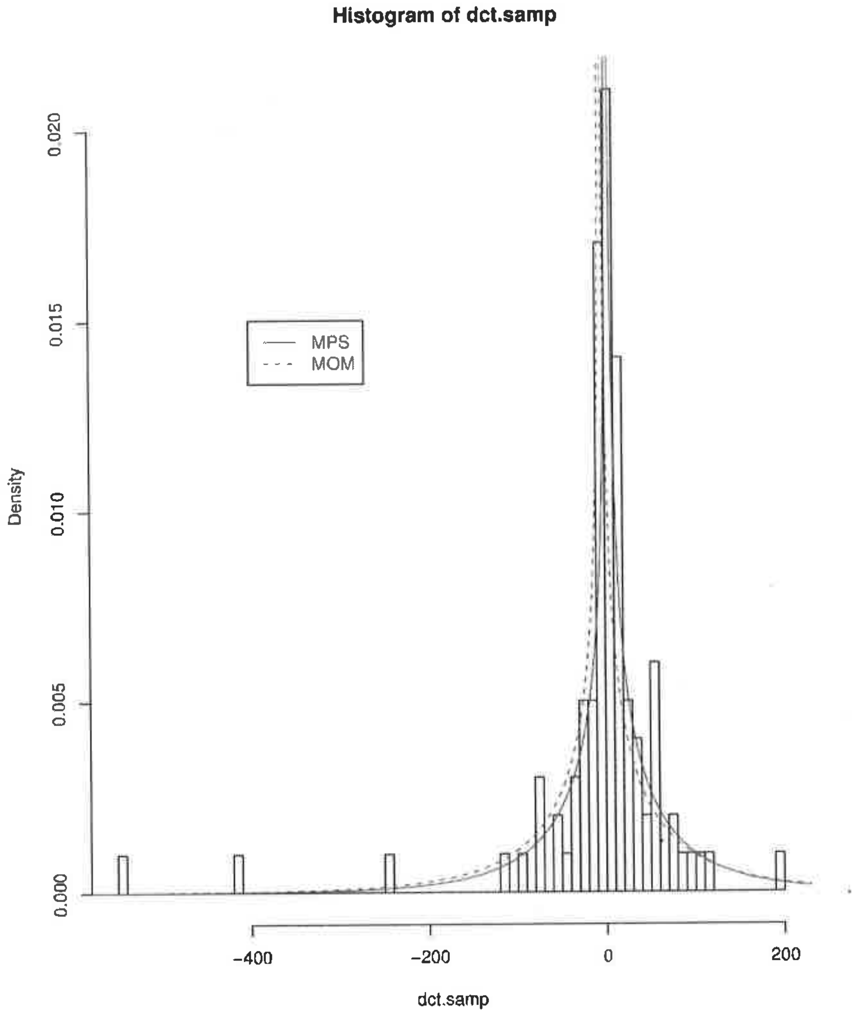


Figure 3: Histogram of the DCT coefficients with the superimposed estimated densities using MOM and MPS.

REFERENCES

Andrews, D.F., Bickel, P.J., Hampel, F.R., Huber, P.J., Rogers, W.H., and Tukey, J.W. (1972), *Robust Estimates of Location: Survey and Advances*, New Jersey: Princeton University Press.

Balakrishnan, N. and Kocherlakota, S. (1985), "On the double Weibull distribution: Order statistics and estimation," *Sankhya: The Indian Journal of Statistics, Series B*, 47, 161–178.

Bhaskaranand, M. (2008), "Video compression using three-dimensional discrete cosine transform," Ph.D. dissertation proposal, dept. of electrical and computer engineering, University of California Santa Barbara.

Borghi, O. (1965), "Sobre Una Distribucion De Frequencias," *Trabajos De Estadstica Y De Investigacon Operativa*, 16, 171–192.

Chang, J.-H., Shin, J.W., Kim, N.S., and Mitra, S. K. (2005), "Image probability distribution based on generalized gamma function," *IEEE Signal Processing Letters*, 12, 325–328.

Eggerton, J.D. and Srinath, M.D. (1986), "Statistical distributions of image DCT coefficients," *Computer and Electrical Engineering*, 12, 137–145.

Ghosh, K. and Jammalamadaka, S.R. (2001), "A general estimation method using spacings," *Journal of Statistical Planning and Inference*, 93, 71–82.

Jayant, N.S. and Noll, P. (1984), *Digital Coding of Waveforms*, New Jersey: Prentice-Hall.

Johnson, N.L. and Kotz, S. (1970), *Distributions in Statistics: Continuous Univariate Distributions-1*, Wiley.

Joshi, R.L. and Fischer, T.R. (1995), "Comparison of generalized Gaussian and Laplacian modeling in DCT image coding," *IEEE Signal Processing Letters*, 2, 81–82.

Kantam, R.R.L. and Narasimham, V.L. (1991), "Estimation in reflected Gamma distribution," *Sankhya: The Indian Journal of Statistics, Series B*, 53, 25–47.

Lam, E.Y. and Goodman, J.W. (2000), "A mathematical analysis of the DCT coefficient distributions for Images," *IEEE Transactions on Image Processing*, 9, 1661–1666.

Malevergne, Y., Pisarenko, V., and Sornette, D. (2005), "Empirical distributions of stock returns: between the stretched exponential and the power law?" *Quantitative Finance*, 5, 379–401.

—(2006), "The modified weibull distribution for asset returns: reply," *Quantitative Finance*, 6, 451–451.

Malevergne, Y. and Sornette, D. (2004), "Value-at-risk-efficient portfolios for a class of super and sub-exponentially decaying assets returning distributions," *Quantitative Finance*, 4, 17–36.

—(2005), "Multivariate Weibull distributions for Asset Returns: I," *Finance Letters*, 2, 16–32.

Nadarajah, S. (2004), "Reliability for Laplace Distributions," *Mathematical Problems in Engineering*, 2, 169–183.

Nadarajah, S. and Kotz, S. (2008), "Strength modeling using Weibull distributions," *Journal of Mechanical Science and Technology*, 22, 1247–1254.

Plucinska, A. (1965), "On Certain Problems Connected with a division of a normal population into parts," *Zastosowania Matematyki*, 8, 117–125, in Polish.

—(1966), "On a general form of the probability density function and its application to the investigation of the distribution of rheostat Resistance," *Zastosowania Matematyki*, 9, 9–19, in Polish.

—(1967), "The reliability of a compound system under consideration of the system elements prices," *Zastosowania Matematyki*, 9, 123–134, in Polish.

Ranneby, B. (1984), "The maximum spacing method: An estimation method related to the maximum likelihood method," *The Scandinavian Journal of Statistics*, 11, 93–112.

Rao, A. V. D. and Narasimham, V. L. (1989), "Linear estimation in double Weibull distribution," *Sankhya: The Indian Journal of Statistics, Series B*, 51, 24–64.

Roese, J., Pratt, W., and Robinson, G. (1977), "Interframe cosine transform image coding," *IEEE Transactions on Communications*, 25, 1329–1339.

Sornette, D., Simonetti, P., and Andersen, J. V. (2000), "q-field Theory for Portfolio Optimization: 'Fat tails' and nonlinear correlations," *Physics Reports Review Section of Physics Letters*, 335, 20–92.

Tiao, G. C. and Lund, D. R. (1970), "The use of OLUMV estimators in inference robustness studies of the location parameters of a class of symmetric distributions," *Journal of the American Statistical Association*, 65, 371–386.