A note on gamma difference distributions

Bernhard Klar

Institute of Stochastics,

Karlsruhe Institute of Technology (KIT), Germany,

bernhard.klar@kit.edu

December 18, 2014

It is the aim of this note to point out that the double gamma difference distribution recently introduced by Augustyniak and Doray (2012) is well-known in financial econometrics: it is the symmetric variance gamma family of distributions. We trace back to the various origins of this distribution. In addition, we consider in some detail the difference of two independent gamma distributed random variables with different shape parameters.

Keywords. Double gamma difference distribution; leptokurtic distr.; variance gamma distr.; generalized Laplace distr.; Bessel function distr.; McKay distr.

1 Introduction

In Augustyniak and Doray (2012), the authors considered the distribution of the difference of two independent and identically distributed (i.i.d.) random variables (r.v.'s) from a gamma distribution with parameters of shape $1/\lambda$ and scale $\sqrt{\lambda\theta}$, $(\lambda, \theta > 0)$. The characteristic function (ch.f.) of this distribution, named double gamma difference (DGD) distribution, was shown to be

$$\varphi(t) = (1 + t^2\lambda\theta)^{-1/\lambda} .$$

In Seneta (2004), the ch.f. φ_{vg} of a random variable following a variance gamma (VG) distribution with parameters $(c, \sigma, \vartheta, \lambda)$ is given by

$$\varphi_{vg}(t) = e^{ict}(1 - i\vartheta\lambda t + \sigma^2\lambda t^2/2)^{-1/\lambda} . \tag{1}$$

Obviously, setting the location parameter c and the parameter ϑ, which is related to skewness, to zero, and identifying θ and $\sigma^2/2$, both φ and φ_{vg} coincide. This is just the ch.f. of the original symmetric VG distribution proposed by Madan and Seneta (1990) in a slightly different parametrisation. Hence, all properties of the DGD distribution given in Augustyniak and Doray (2012) follow from the corresponding properties of the VG distribution.

In the light of this equivalence, it seems advisable to use a standard test for the hypothesis $H_0 : \vartheta = 0$ to check for symmetry instead of using a general test for symmetry as proposed in Augustyniak and Doray (2012). Contrary to the statement in Augustyniak and Doray (2012), the density function of a VG distribution has a closed-form expression in terms of a modified Bessel function of the third kind and elementary functions (Seneta, 2004); it is given by

$$f_{vg}(x) = \frac{2\exp(\vartheta(x-c)/\sigma^2)}{\sigma\sqrt{2\pi}\lambda^{1/\lambda}\Gamma(1/\lambda)} \left(\frac{|x-c|}{\sqrt{2\vartheta^2/\lambda+\sigma^2}}\right)^{\frac{1}{\lambda}-\frac{1}{2}} K_{1/\lambda-1/2}\left(\frac{|x-c|\sqrt{2\sigma^2/\lambda+\vartheta^2}}{\sigma^2}\right). \tag{2}$$

Hence, maximum likelihood estimation is possible. This and several other methods for parameter estimation are discussed in Finlay and Seneta (2008). Nevertheless, estimation based on the empirical ch.f. is particularly convenient for the VG family; details can be found in Yu (2004).

In Augustyniak and Doray (2012), the authors propose goodness-of-fit testing for the DGD family basically by applying the test of Koutrouvelis and Kellermeier (1981) based on the empirical ch.f. However, this test has several disadvantages: it is not consistent against all alternatives, the choice of the evaluation points is rather arbitrary, and the simulation study in Augustyniak and Doray (2012) reveals that this test does not maintain its theoretical level. An empirical ch.f. based test which does not have these deficiencies is proposed in Fragiadakis et al. (2013). Their test statistic is a weighted integral incorporating the empirical ch.f. of suitably standardized data, and is easily computable. Their simulations show that the test, performed by using a parametric bootstrap procedure, captures the nominal level of significance in the majority of cases quite closely. Alternatively, one may use classical tests like the Cramér-von Mises- or the χ^2-test (Fragiadakis et al., 2013). Clearly, these tests can also be used for the symmetric VG distribution by restricting the parameter space appropriately.

Albeit a variance gamma distributed r.v. can be represented as the difference of two independent gamma variates with equal shape parameters (see below), this model has its origin in a specific stochastic process and is obtained formally from a normal variate by mixing on the variance parameter (Madan and Seneta, 1990). This genesis is a bit involved, and one may guess that the distribution of differences of gamma r.v.'s should have been investigated much earlier. This is indeed the case. We shortly review the appearance of such distributions in the mathematical and engineering literature in the following section. Subsequently, we discuss in more detail the case of two independent gamma variates with different shape parameters.

2 Historical remarks

A first line of research goes back to Pearson et al. (1929) in connection with the distribution of the sample covariance for a random sample drawn from a bivariate normal population. The pertaining density is also given in Press (1967, p. 356) having the same form as the density in (2). Press also derived an explicit expression for the distribution function of the sample covariance for even sample sizes. Furthermore, he showed that this distribution can be represented as the difference of two independent gamma variates with equal shape parameters (Press, 1967, Section 4).

The distribution of Pearson et al. (1929) has been also investigated by McKay (1932) who termed it *Bessel function distribution.* His aim was to develop an alternative to the well-known Pearson family of distributions. This is the reason why the distribution is also known under the name *McKay distribution*, mainly in the engineering literature. For example, Holm and Alouini (2004) rederive the moments of this distribution and show that it can be represented as the distribution of the difference of two independent gamma variates with equal shape parameters. However, this McKay distribution should not be confused with another distribution introduced by McKay (1934), termed McKays bivariate gamma distribution (see, e.g., Gupta and Nadarajah (2006)).

Moreover, the VG or Bessel function distribution appears repeatedly in engineering applications when considering several arithmetic combinations of Gaussian random variables. This leads, amongst others, to differences of chi-squared r.v.'s. An early example is the technical report of Omura and Kailath (1965) who treat differences of central and non-central chi-squared r.v.'s in Section IV. For the special case of equal and even degrees of freedom, distribution function and density are given on pages 25 and 73, respectively. There, one can also find the density in case of arbitrary (but still equal) degrees of freedom. On page 75, they give the density for different degrees of freedom in terms of elementary functions and Whittaker's W function (see Section 3 below). An updated version of these technical notes is prepared by Simon (2002).

A further line of research starts with the symmetric or asymmetric Laplace distribution. If $X_1, \ldots, X_n$ are i.i.d. Laplace r.v.s with mean zero and variance σ^2, then their sum S_n has the ch.f.

$$\psi_{S_n}(t) = (1 + \sigma^2 t^2/2)^{-n}, \quad -\infty < t < \infty,$$

which is a proper ch.f. even if n is not an integer. Introducing a skewness parameter in the same way as for the asymmetric Laplace distribution leads to the ch.f.

$$\psi(t) = (1 - i\mu t + \sigma^2 t^2/2)^{-\tau}, \quad -\infty < t < \infty, \tag{3}$$

where $\mu \in \mathbb{R}, \sigma^2, \tau > 0$. Adding a location parameter, we come once again back to the ch.f. in (1). A comprehensive account of this *generalized (asymmetric) Laplace distribution* is given in Section 4 of Kotz et al. (2001), where one can also find a number of different representations of this distribution. In particular, the aforementioned result of Press (1967) is stated on p. 229, taking the following form: let G_1, G_2 be i.i.d gamma distributed r.v.'s with shape parameter τ and scale parameter 1. Then, $Y = \sigma(G_1/\kappa - \kappa G_2)/\sqrt{2}$ has the ch.f. given in (3), where $\kappa = (\sqrt{2\sigma^2 + \mu^2} - \mu)/(\sqrt{2}\sigma)$.

Different generalizations of the (asymmetric) Laplace distribution can be found in Section 4.4 of Kotz et al. (2001), and in Yu and Zhang (2005).

3 Gamma difference distribution

In contrast to the VG or Bessel function distribution, the distribution of the difference of two independent gamma r.v.'s with different shape parameters has received very little attention. Therefore, it seems justified to treat this case in more detail. Besides the aforementioned results, we are only aware of the work of Mathai (1993) who derived a closed form expression for the density; there, the distribution is termed *gamma difference distribution*. It is also mentioned in Krishna and Jose (2011) under the name *generalized asymmetric Laplace distribution*. We adopt the first name since there already exist different distributions under the second one.

Let X_1 and X_2 be two independent random variables such that $X_1 \sim \Gamma(\alpha_1, \beta_1)$ and $X_2 \sim \Gamma(\alpha_2, \beta_2)$, where $\alpha_i > 0, \beta_i > 0, i = 1, 2$. Then $X = X_1 - X_2$ follows a gamma difference distribution (GDD) with parameters $\alpha_1, \beta_1, \alpha_2, \beta_2$. For practical applications, it will be convenient to introduce an additional location parameter $\theta \in \mathbb{R}$ and to consider $X_1 - X_2 + \theta$; however, in the following, we put $\theta = 0$. Then, the moment generating function (mgf) of X is

$$M_X(t) = (1 - t/\beta_1)^{-\alpha_1}(1 + t/\beta_2)^{-\alpha_2}$$

for $|t| < \min(\beta_1, \beta_2)$. Integer moments can easily be computed by means of the moments of X_1 and X_2 using the formula

$$E(X^n) = \sum_{k=0}^{n} \binom{n}{k} (-1)^k \, E(X_1^{n-k}) \, E(X_2^k).$$

In particular, mean μ, variance σ^2 and skewness $\gamma_1 = E[((X - \mu)/\sigma)^3]$ are given by

$$\mu = \frac{\alpha_1}{\beta_1} - \frac{\alpha_2}{\beta_2}, \qquad \sigma^2 = \frac{\alpha_1}{\beta_1^2} + \frac{\alpha_2}{\beta_2^2}, \qquad \gamma_1 = \frac{2(\alpha_1\beta_2^3 - \alpha_2\beta_1^3)}{(\alpha_1\beta_2^2 + \alpha_2\beta_1^2)^{3/2}}.$$

Hence, the skewness is zero for $\alpha_1\beta_2^3 = \alpha_2\beta_1^3$, but the distribution of X is asymmetric except for the case $\alpha_1 = \alpha_2, \beta_1 = \beta_2$.

Using the fact that X_1 and X_2 are independent random variables with positive support and denoting the density of X by $f(z)$, we obtain

$$f(z) = \begin{cases} c\, e^{\beta_2 z} \int_z^\infty x^{\alpha_1 - 1}(x - z)^{\alpha_2 - 1}\, e^{-(\beta_1+\beta_2)x}\, dx, & z > 0 \\ c\, e^{-\beta_1 z} \int_{-z}^\infty x^{\alpha_2 - 1}(x + z)^{\alpha_1 - 1}\, e^{-(\beta_1+\beta_2)x}\, dx, & z < 0, \end{cases} \tag{4}$$

with $c = \beta_1^{\alpha_1}\beta_2^{\alpha_2}/(\Gamma(\alpha_1)\Gamma(\alpha_2))$, where $\Gamma(\cdot)$ denotes the gamma function. A similar formula can be found in Mathai (1993). Using formula 3.383(4) in Gradshteyn and Ryzhik (1980), the integrals appearing in (4) can be expressed by means of Whittaker's W function, a confluent hypergeometric function (Olver et al., 2010, 13.14.3), as follows:

$$f(z) = \begin{cases} \frac{\tilde{c}}{\Gamma(\alpha_1)}\, z^{\frac{\alpha_1+\alpha_2}{2}-1} e^{\frac{\beta_2-\beta_1}{2} z}\, W_{\frac{\alpha_1-\alpha_2}{2}, \frac{1-\alpha_1-\alpha_2}{2}}\left((\beta_1+\beta_2)z\right), & z > 0 \\ \frac{\tilde{c}}{\Gamma(\alpha_2)}\, (-z)^{\frac{\alpha_1+\alpha_2}{2}-1} e^{\frac{\beta_1-\beta_2}{2}(-z)}\, W_{\frac{\alpha_2-\alpha_1}{2}, \frac{1-\alpha_1-\alpha_2}{2}}\left((\beta_1+\beta_2)(-z)\right), & z < 0, \end{cases} \tag{5}$$

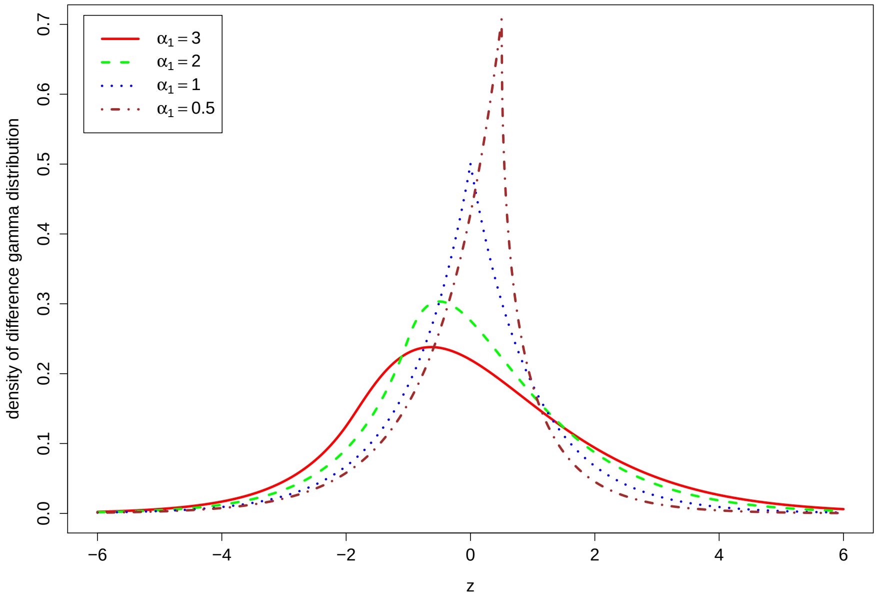

Figure 1: Densities of centered gamma difference distributions with $\alpha_2 = \beta_1 = \beta_2 = 1$.

where

$$\tilde{c} = \frac{\beta_1^{\alpha_1} \beta_2^{\alpha_2}}{(\beta_1 + \beta_2)^{(\alpha_1+\alpha_2)/2}} \ .$$

This corresponds to formula (2.5) in Mathai (1993), using a slightly different parametrization. However, even if Whittaker's W function is available in many statistical or mathematical programming environments, one has to take care using (5) since $W_{\kappa,\mu}$ may not be defined for all μ. Hence, a direct approach using (4) together with numerical quadrature may be more convenient. As mentioned above, a corresponding expression for chi-squared r.v.'s can be found in Omura and Kailath (1965). For even degrees of freedom, or, equivalently, integer shape parameters of the gamma distributions, the integral in (4) can be expressed as a finite sum of elementary functions, see Omura and Kailath (1965, p. 73) or Simon (2002, p. 28).

Figure 1 shows GDD densities for $\alpha_2 = \beta_1 = \beta_2 = 1$ and several values of α_1; the distributions are centered, i.e. $EX = 0$. Note that the distributions are skewed to the right for $\alpha_1 = 3$ and $\alpha_1 = 2$, symmetric for $\alpha_1 = 1$ (i.e. the Laplace distribution), and left-skewed for $\alpha_1 = 0.5$.

By using the formula

$$P(X_1 - X_2 \leq t) = \int_0^\infty P(X_1 \leq x + t)\, f_{X_2}(x) dx,$$

where $f_{X_2}(x)$ denotes the density function of X_2, it is also possible to give a practically

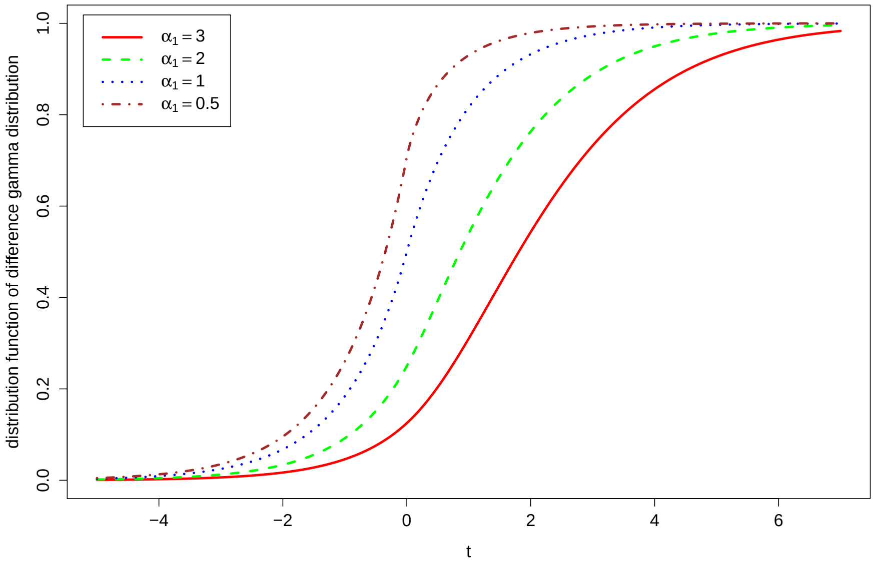

Figure 2: Distribution functions of the GDD with $\alpha_2 = \beta_1 = \beta_2 = 1$.

applicable form of the distribution function $F(t)$ of X. Let

$$\gamma(\alpha, y) = \int_0^y t^{\alpha-1} e^{-t} dt$$

denote the lower incomplete gamma function, then

$$F(t) = d \int_{\max\{0,-t\}}^{\infty} x^{\alpha_2 - 1} e^{-\beta_2 x} \, \gamma\left(\alpha_1, \beta_1(x+t)\right) dx \quad (t \in \mathbb{R}),$$

where $d = \beta_2^{\alpha_2}/(\Gamma(\alpha_1)\Gamma(\alpha_2))$. Again, for integer shape parameters of the gamma distributions, this can be expressed as a finite sum of elementary functions, see Omura and Kailath (1965, p. 25) or Simon (2002, p. 28). Distribution functions of (non-centered) gamma difference distributions with $\alpha_2 = \beta_1 = \beta_2 = 1$ and the same values of α_1 as above are shown in Figure 2.

In order to give an impression about statistical inference for the GDD, we employ characteristic function based parameter estimation. The ch.f. of the GDD and the empirical ch.f. of a sample $X_1, \ldots, X_n$ from the GDD are given by

$$\psi_X(t;\theta) = (1 - it/\beta_1)^{-\alpha_1}(1 + it/\beta_2)^{-\alpha_2} \qquad \text{and} \qquad \psi_n(t) = \frac{1}{n}\sum_{j=1}^{n} e^{itX_j},$$

respectively, where $\theta = (\alpha_1, \alpha_2, \beta_1, \beta_2)$ denotes the parameter vector. Defining

$$T_n = \int_{-\infty}^{\infty} |\psi_n(t) - \psi_X(t;\theta)|^2 \, e^{-t^2} \, dt, \tag{6}$$

		$\alpha_1 = 2$	$\alpha_2 = 1$	$\beta_1 = 1$	$\beta_2 = 0.5$
	bias	1.42	0.50	0.37	0.10
$n = 100$	st.dev.	3.65	1.61	0.96	0.34
	MAD	1.94	0.77	0.57	0.22
	bias	0.20	0.090	0.055	0.025
$n = 400$	st.dev.	0.89	0.39	0.27	0.12
	MAD	0.56	0.26	0.19	0.088
		$\alpha_1 = 4$	$\alpha_2 = 0.5$	$\beta_1 = 1$	$\beta_2 = 0.5$
	bias	2.02	1.17	0.26	0.30
$n = 100$	st.dev.	4.92	2.90	0.62	0.80
	MAD	2.76	1.41	0.41	0.49
	bias	0.45	0.25	0.068	0.082
$n = 400$	st.dev.	1.34	0.81	0.21	0.31
	MAD	0.90	0.41	0.16	0.20

Table 1: Bias, standard deviation and MAD for estimators of the GDD parameters.

we can estimate the parameter vector θ by minimizing T_n. We have performed a small scale simulation study for sample size $n = 100$ and 400 for two different choices of θ, namely $\theta = (2, 1, 1, 0.5)$ and $(4, 0.5, 1, 0.5)$. In both cases, the variance of the underlying GDD distribution is 6; skewness is -0.82 in the first and 0 in the second case.

Over the entire $B = 1000$ simulation runs, comparison between estimation results is made on the basis of bias, standard deviation and mean absolute deviation (MAD) of estimated parameter values from the true values. If ϑ is the true parameter value, and $\hat{\vartheta}_i$ is the estimated parameter value from the i-th simulation run, the bias is given by $\bar{\vartheta} - \vartheta$, where $\bar{\vartheta} = \frac{1}{B}\sum_{i=1}^{B} \hat{\vartheta}_i$. Standard deviation and MAD are defined by

$$\text{st.dev.} = \sqrt{\frac{1}{B}\sum_{i=1}^{B}(\hat{\vartheta}_i - \bar{\vartheta})^2}, \qquad \text{MAD} = \frac{1}{B}\sum_{i=1}^{B}|\hat{\vartheta}_i - \vartheta|.$$

The results of the simulations are reported in Table 1. The values of bias, standard deviation and MAD appear very high for sample size $n = 100$, and strongly decrease for $n = 400$. The reason is that, for $n = 100$, we see in a significant proportion of cases estimated parameter values far away from the true values. However, this does not indicate problems with the estimation procedure or a bad fit in general. As a typical example, consider the case $n = 100$ and $\alpha_1 = 4, \alpha_2 = 0.5, \beta_1 = 1, \beta_2 = 0.5$. In one of the simulation runs, we obtained the estimates $\alpha_1 = 12.73, \alpha_2 = 6.81, \beta_1 = 1.93, \beta_2 = 2.10$. Figure 3 shows the fitted and true density functions of the GDD together with a nonparametric density estimate. One can observe that all three curves are in good agreement despite the large discrepancy between true and estimated parameter values.

Based on the statistic T_n in (6) one could also develop goodness-of-fit tests similar to the test given in Fragiadakis et al. (2013).

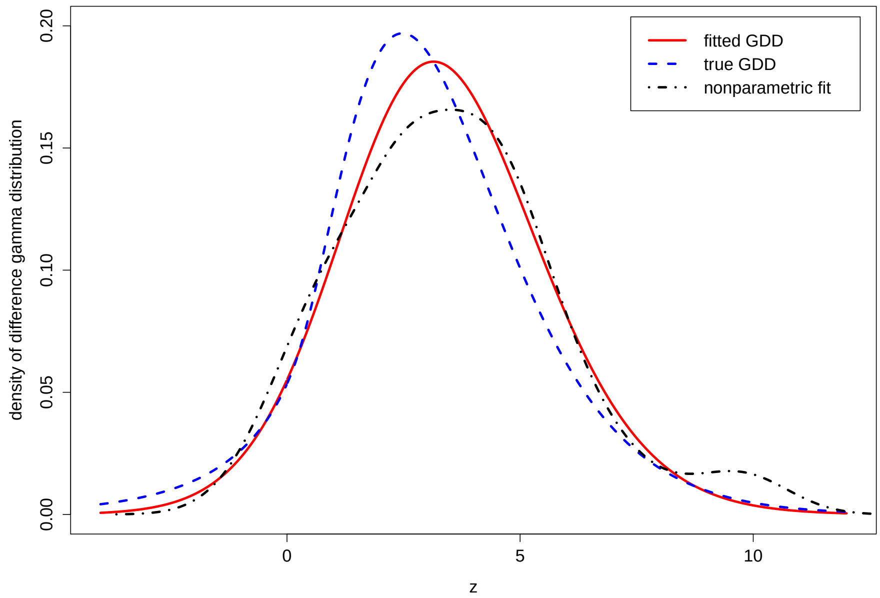

Figure 3: True and fitted GDD density function together with kernel density estimate.

4 Discussion

At first glance, one may get the impression that the difference of two independent gamma r.v.'s has rarely been examined in the literature. For example, this topic is not mentioned in the chapter on the gamma distribution in Johnson et al. (1994), contrary to sums and products of such random variables. However, as we have seen, this impression is definitely wrong. A large bulk of results on this subject exists, but scattered in several directions and under different names. Only the case of independent gamma variates with different shape parameters has not received much attention. The reason seems to be that starting with the symmetric distribution of the difference of two i.i.d. gamma r.v.'s, skewness can be introduced by simply allowing for different scale parameters, leading to a versatile family of distributions with nice mathematical properties. However, even for the general gamma difference distribution, main quantities can be derived without major difficulties. Likewise, statistical inference is readily feasible as the results on parameter estimation show.

References

Augustyniak, M., and Doray, L.G., 2012. *Inference for a leptokurtic symmetric family of distributions represented by the difference of two gamma variables.* Journal of Statistical Computation and Simulation 82, 1621-1634.

Finlay, R., and Seneta, E., 2008. *Stationary-increment variance-gamma and t models: Simulation and parameter estimation.* Int. Statist. Rev. 76, 167-186.

Fragiadakis, K., Karlis, D., and Meintanis, S.G., 2013. *Inference procedures for the variance gamma model and applications.* J. Stat. Comput. Simul. 83, 555-567.

Gradshteyn, I.S., and Ryzhik, I.M., 1980. *Table of Integrals, Series, and Products.* Academic Press.

Gupta, A., Nadarajah, S., 2006. Sums, products and ratios for McKay's bivariate gamma distribution. Math. Comput. Modelling 43, 185193.

Holm, H., and Alouini, M.S., 2004. *Sum and difference of two squared correlated Nakagami variates in connection with the McKay distribution.* IEEE Trans. Commun. 52, 1367-1376.

Johnson, N.L., Kotz, S., and Balakrishnan, N., 1994. *Continuous Univariate Distributions*, Vol. 1. Wiley Series in Probability and Statistics.

Kotz, S., Kozubowski, T., Podgorski, K., 2001. *The Laplace Distribution and Generalizations. A Revisit with Applications to Communications, Economics, Engineering, and Finance.* Birkhäuser.

Koutrouvelis, I.A., and Kellermeier, J., 1981. *A goodness-of-fit test based on the empirical characteristic function when parameters must be estimated.* J. R. Stat. Soc. Ser. B Stat. Methodol. 43, 173-176.

Krishna, E., and Jose, K.K., 2011. *Marshall-Olkin Generalized Asymmetric Laplace distributions and processes.* Statistica 71, 453-467.

Madan D.B., and Seneta E., 1990. *Variance Gamma (V.G.) Model for Share Market Returns.* The Journal of Business 63, 511-524.

Mathai, A.M., 1993. *On non-central generalized Laplacianness of quadratic forms in normal variables.* Journal of Multivariate Analysis 45, 239246.

McKay, A.T., (1932). *A Bessel function distribution.* Biometrika XXIV, 39-44.

McKay, A. T., (1934). *Sampling from batches.* Journal of the Royal Statistical Society, 1 (Suppl.), 207216.

Olver, F.W.J., Lozier, D.W., Boisvert, R.F., and Clark, C.W., 2010. *NIST Handbook of Mathematical Functions.* Cambridge. Online resource: *NIST Digital Library of Mathematical Functions*, `http://dlmf.nist.gov/`

Omura, J.K., and Kailath, T. (1965). *Some Useful Probability Functions.* Tech. Rep. 7050-6, Stanford University.

Pearson, K., Jefferey, G.B. and Elderton, E.M. (1929). *On the distribution of the first product moment-coefficient, in samples drawn from an indefinitely large normal population.* Biometrika 21, 164-193.

Press, S.J., 1967. *On the sample covariance from a bivariate normal distribution.* Annals of the Institute of Statistical Mathematics 19, 355-361.

Seneta, E., 2004. *Fitting the variance-gamma model to financial data.* J. Appl. Prob. 41, 177-187.

Simon, M.K.. 2002. *Probability Distributions Involving Gaussian Random Variables. A Handbook for Engineers and Scientists.* Springer.

Yu, J., 2004. *Empirical characteristic function estimation and its applications.* Econometric Rev. 23, 93123.

Yu, K., and Zhang, J., 2005. *A three-parameter asymmetric Laplace distribution and its extension.* Communications in Statistics - Theory and Methods 34, 1867-1879.