J. Statist. Comput. Simul., 1997, Vol. 59, pp. 375–384
Reprints available directly from the publisher
Photocopying permitted by license only

NORMAL SCALE MIXTURES AND DUAL PROBABILITY DENSITIES

TILMANN GNEITING

*Department of Mathematics, Universität Bayreuth,
D-95440 Bayreuth, Germany*

(Received 18 March 1997; In final form 25 June 1997)

Let p be a probability density function representable as a scale mixture of normal probability densities. Whenever p is bounded, there is a dual density $\hat{p}$ proportional to the characteristic function of p. We present numerous examples for pairs of dual probability densities representable as normal scale mixtures, among them generalized hyperbolic, Student's t and McKay's Bessel function, symmetric stable and exponential power, generalized logistic and generalized hyperbolic secant, α-Laplace or Linnik and generalized Cauchy densities.

Keywords: Dual density; normal scale mixture; random variate generation

Whenever a probability density function p has a representation of the form

$$p(x) = \int_0^\infty \frac{1}{(2\pi v)^{1/2}} \exp\left(\frac{-x^2}{2v}\right) dF(v), \tag{1}$$

where F is a distribution function on $(0, \infty)$, we talk of p as a *scale mixture* or *variance mixture* of normal probability densities. In many situations in statistics it is useful to identify a given density as a normal scale mixture, and to determine the *mixing distribution* F in the representation (1). The well-known Monte Carlo swindle of the Princeton robustness study (Andrews *et al.*, 1972), for example, relies on the normal scale mixture property of the sampling distributions. Knowledge of the mixing distribution allows to generate random variates with density p as the product $X = ZV^{1/2}$, where Z is standard normal, and V is an independent random variable with distribution function F. Equivalently, we might generate X as the ratio $X = Z/U$ where $U = 1/V^{1/2}$ is independent of Z. This representation accounts for the alternative expression of *normal/independent* distributions popular in robustness studies (Lange and Sinsheimer, 1993; Liu, 1996).

Various examples of normal scale mixtures and the corresponding mixing distributions have been given: Finite mixtures (Newcomb, 1886), contaminated normal and slash (Rogers and Tukey, 1972), the Student's t family and Laplace or double exponential (Andrews *et al.*, 1972), symmetric stable and logistic (Andrews and Mallows, 1974), hyperbolic (Barndorff-Nielsen, 1977), generalized hyperbolic and Fisher's z or generalized logistic (Barndorff-Nielsen *et al.*, 1982), and exponential power (West, 1987). In the present note we add some nonstandard densities to this list and discuss an interesting duality property for a broad class of normal scale mixtures.

Consider the characteristic function

$$\varphi(t) = \int_0^\infty \exp\left(\frac{-vt^2}{2}\right) dF(v) \tag{2}$$

of the normal scale mixture p in (1). In case

$$p(0) = \int_0^\infty \frac{1}{(2\pi v)^{1/2}} dF(v)$$

is finite, both p and φ are bounded, strictly positive, and integrable. We may then interchange the roles of the density and the characteristic function, and there is a *dual*, *reciprocal*, or *adjoint* density $\hat{p}$ which is proportional to the characteristic function φ. In turn, the dual density $\hat{p}$ has characteristic function $\hat{\varphi}$ proportional to p, such that

$$\varphi(t) = \frac{\hat{p}(t)}{\hat{p}(0)}, \quad \hat{\varphi}(t) = \frac{p(t)}{p(0)}, \quad 2\pi p(0)\hat{p}(0) = 1; \tag{3}$$

the last equality following from the Fourier inversion theorem (Rossberg, 1990). The dual probability density $\hat{p}$ is evidently a normal scale mixture of the form (1) and admits the representation

$$\hat{p}(x) = \int_0^\infty \frac{1}{(2\pi v)^{1/2}} \exp\left(\frac{-x^2}{2v}\right) d\hat{F}(v) \tag{4}$$

for some mixing distribution $\hat{F}$ on $(0, \infty)$. From (3) and (2) we see that

$$\begin{aligned}\hat{p}(x) &= \hat{p}(0)\ \varphi(x)\\ &= \frac{1}{2\pi p(0)}\int_0^\infty \exp\left(\frac{-vx^2}{2}\right)dF(v)\\ &= -\frac{1}{2\pi p(0)}\int_0^\infty \exp\left(\frac{-x^2}{2v}\right)dF\left(\frac{1}{v}\right).\end{aligned} \tag{5}$$

In view of the uniqueness theorem for Laplace transforms, the representations (4) and (5) for $\hat{p}$ imply

$$\frac{1}{(2\pi v)^{1/2}}d\hat{F}(v) = -\frac{1}{2\pi p(0)}dF\left(\frac{1}{v}\right), \quad v > 0.$$

Integration with respect to v and a change of variables lead to the formula

$$\hat{F}(u) = \frac{1}{(2\pi)^{1/2}p(0)}\int_{1/u}^\infty \frac{1}{v^{1/2}}dF(v), \quad u > 0, \tag{6}$$

for the cumulative distribution function of the mixing distribution $\hat{F}$. In particular, if F has density $f(v)$ for $v > 0$, then $\hat{F}$ has density

$$\hat{f}(v) = \frac{1}{(2\pi)^{1/2}p(0)v^{3/2}}\,f\left(\frac{1}{v}\right), \quad v > 0. \tag{7}$$

The present work has been motivated by a number of references, among them Kelker (1971), West (1987), and Good (1995). Kelker (1971, p. 807) mentions dual normal scale mixtures in a study of infinite divisibility. West (1987) makes use of the duality of the symmetric stable and exponential power family to derive a normal scale mixture representation for the exponential power distribution. His result is a special case of the identity (7). Good (1995) notes the duality of Student's t and the symmetric Bessel family, but does not refer to normal scale mixtures. Indeed, the notion of dual densities is not restricted to normal scale mixtures, but tied to positive definiteness (Rossberg, 1990). Yet it appears that almost all dual densities of statistical significance arise as normal scale mixtures. The following examples may therefore provide a reasonably complete response to Good's (1995) recent request for a list of dual pairs of density functions. Moreover, some of the normal scale mixture representations are new and possibly useful in Monte Carlo studies.

Example 1 Barndorff-Nielsen *et al.* (1978; 1981; 1982) consider a class of variance-mean mixtures of normal distributions which they term *generalized hyperbolic* distributions. The univariate and symmetric members of this family arise as normal scale mixtures whose mixing distribution is the generalized inverse Gaussian distribution with density

$$f(v) = \frac{(\kappa/\delta)^{\lambda}}{2\,K_{\lambda}(\delta\kappa)} v^{\lambda-1} \exp\left(-\frac{1}{2}\left(\kappa^2 v + \frac{\delta^2}{v}\right)\right), \quad v > 0; \tag{8}$$

here K is a modified Bessel function and the parameter $(\lambda, \delta, \kappa)$ satisfies

$$\begin{aligned} &\delta \geq 0, \quad \kappa > 0 \text{ for } \lambda > 0, \\ &\delta > 0, \quad \kappa > 0 \text{ for } \lambda = 0, \\ &\delta > 0, \quad \kappa \geq 0 \text{ for } \lambda < 0. \end{aligned}$$

The resulting normal scale mixture has probability density function

$$\frac{\kappa^{1/2}}{(2\pi)^{1/2}\delta^{\lambda}K_{\lambda}(\delta\kappa)} (\delta^2 + x^2)^{\frac{1}{2}(\lambda-\frac{1}{2})}\, K_{\lambda-\frac{1}{2}}(\kappa(\delta^2 + x^2)^{1/2}), \tag{9}$$

interpreted as the corresponding limit expression if $\delta = 0$ or $\kappa = 0$. In view of the mixing density (8), formula (7) shows that generalized hyperbolic densities with parameter $(\lambda, \delta, \kappa)$ and $(1/2-\lambda, \kappa, \delta)$ are dual. In particular, if $\lambda = 1/4$ and $\delta = \kappa > 0$, the generalized hyperbolic density is self-dual (Barndorff-Nielsen *et al.*, 1981; 1982).

Let us consider some special cases. If $\lambda = 1$, the probability density function (9) reduces to

$$\frac{1}{2\delta\, K_1(\delta\kappa)} \exp(-\kappa(\delta^2 + x^2)^{1/2}),$$

that is, the symmetric *hyperbolic* density. The dual density is generalized hyperbolic with parameter $(-1/2, \kappa, \delta)$. It has recently been studied and applied under the name *normal inverse Gaussian* distribution (Barndorff-Nielsen, 1996; Rydberg, 1996).

If $\lambda = -\nu/2 < 0$, $\delta^2 = \nu$, and $\kappa = 0$, the density (9) reduces to

$$\frac{\Gamma((\nu+1)/2)}{(\pi\nu)^{1/2}\,\Gamma(\nu/2)}\left(1+\frac{x^2}{\nu}\right)^{-(\nu+1)/2},$$

the probability density function of *Student's t* with parameter $\nu > 0$. The mixing density (8) is then such that $1/v^2 f(1/v)$ is a Gamma$(\nu/2, \nu/2)$ density. This corresponds to the well-known result that a Student's t variable may be generated as $Z/V^{1/2}$, where Z is standard normal, and V is independent of Z and distributed as Gamma$(\nu/2, \nu/2)$, or, for $\nu = n$ an integer, as χ_n^2/n. The dual probability density is generalized hyperbolic with parameter $\lambda = (\nu+1)/2$, $\delta = 0$, and $\kappa^2 = \nu$. Then (8) is a gamma density, and formula (9) gives

$$\frac{\nu^{1/2}}{\pi^{1/2}\,\Gamma((\nu+1)/2)}\left(\frac{\nu^{1/2}|x|}{2}\right)^{\nu/2} K_{\nu/2}(\nu^{1/2}|x|), \tag{10}$$

that is, the symmetric members of McKay's (1932) *Bessel function* density. For $\nu = 1$, formula (10) recovers the *Laplace* or *double exponential* density.

Example 2 The *symmetric stable* distribution with exponent $\alpha \in (0, 2]$ is defined by its characteristic function

$$\varphi(t) = \exp(-|t|^\alpha).$$

It is well-known that symmetric stable distributions are normal scale mixtures. If $\alpha \in (0, 2)$, the mixing density in the representation (1) is given by

$$f(v) = \frac{1}{2} s_{\alpha/2}\left(\frac{v}{2}\right), \quad v > 0, \tag{11}$$

where $s_{\alpha/2}$ is the density of the positive stable distribution with index $\alpha/2$ (Feller, 1971, p. 596).

The dual family is the class of *exponential power* distributions with index $\alpha\in(0,2]$ and probability density function

$$[2\,\Gamma(1+1/\alpha)]^{-1}\exp(-|x|^{\alpha}) \tag{12}$$

(Box and Tiao, 1992, pp. 156–159; Devroye, 1986, pp. 174–175). By equations (7) and (3), we identify the density (12) as a normal scale mixture with mixing density

$$\hat{f}(v)=\frac{\pi^{1/2}}{2^{3/2}\,\Gamma(1+1/\alpha)v^{3/2}}\,s_{\alpha/2}\left(\frac{1}{2v}\right),\quad v>0. \tag{13}$$

For $\alpha\in[1,2)$, this corresponds to a result of West (1987). Lange and Sinsheimer (1993) have a related series expansion.

Example 3 The symmetric *Fisher's z* or symmetric *generalized logistic* distribution with index $\alpha>0$ (Johnson *et al.*, 1994b, p. 141; Barndorff-Nielsen *et al.*, 1982) has density function

$$[B(\alpha,\alpha)]^{-1}\frac{e^{-\alpha x}}{(1+e^{-x})^{2\alpha}}=[4^{\alpha}B(\alpha,\alpha)]^{-1}\operatorname{sech}^{2\alpha}\left(\frac{x}{2}\right), \tag{14}$$

where B is the beta function. If $\alpha=1$ we have the standard *logistic* distribution, and $\alpha=1/2$ recovers the standard *hyperbolic cosine* or *hyperbolic secant* distribution. The characteristic function of the density (14) is

$$\varphi(t)=\frac{\Gamma(\alpha+it)\,\Gamma(\alpha-it)}{(\Gamma(\alpha))^{2}}. \tag{15}$$

Whenever 2α is an integer, (15) can be expressed in terms of hyperbolic functions (Gradshteyn and Ryzhik, 1994, Section 3.985). In particular, the logistic distribution has characteristic function $\varphi(t)=\pi t/\sinh(\pi t)$, and the characteristic function of the hyperbolic secant distribution is $\varphi(t)=\operatorname{sech}(\pi t)$. Barndorff-Nielsen *et al.* (1982) show that the symmetric z densities are normal scale mixtures with mixing density

$$h_{\alpha}(v)=\sum_{k=0}^{\infty}\binom{-2\alpha}{k}\frac{(\alpha+k)}{B(\alpha,\alpha)}\exp\left(-\frac{1}{2}(\alpha+k)^{2}v\right),\quad v>0,$$

in the canonical representation (1).

The dual family is the class of *generalized hyperbolic secant* distributions (Harkness and Harkness, 1968; Johnson *et al.*, 1994b, p. 148) defined by the characteristic function

$$\hat{\varphi}(t) = \operatorname{sech}^{2\alpha}\left(\frac{t}{2}\right)$$

which is proportional to the density (14). Formula (7) shows that generalized hyperbolic secant distributions are normal scale mixtures with mixing density

$$\frac{4^{\alpha}B(\alpha,\alpha)}{(2\pi)^{1/2}v^{3/2}}\, h_{\alpha}\left(\frac{1}{v}\right), \quad v > 0.$$

Example 4 The *Linnik* or *α-Laplace* distribution with index $\alpha\in(0,2]$ (Johnson *et al.*, 1994b, pp. 199–200) has characteristic function

$$\varphi(t) = (1 + |t|^{\alpha})^{-1}.$$

A result of Devroye (1990) leads to a simple proof that Linnik distributions are normal scale mixtures. Devroye in fact shows that a random variable with characteristic function

$$\varphi(t) = (1 + |t|^{\alpha})^{-\beta}, \tag{16}$$

where $\alpha\in(0,2]$ and $\beta>0$, can be generated as the product $X = S_{\alpha}V_{\beta}^{1/(\alpha\beta)}$. Here, S_{α} is a symmetric stable random variable of exponent α, and V_{β} is a positive random variable independent of S_{α} and with density $[\Gamma(1+\beta)]^{-1}\exp(-v^{1/\beta})$, $v>0$. Recalling the normal scale mixture representation for the symmetric stable distribution in Example 2, we see that the distribution with characteristic function (16) is a normal scale mixture that can be generated as $Z(2UV_{\beta}^{(2/\alpha\beta)})^{1/2}$, where Z, U, V_{β} are independent, Z is standard normal, U is positive stable with index $\alpha/2$, and V_{β} is as above.

Expression (16) gives an integrable characteristic function whenever $\alpha\in(0,2]$ and $\beta>1/\alpha$. For these values of α and β, we have the dual class of *generalized Cauchy* distributions (Johnson *et al.*, 1994a, p. 327) with probability density function

$$\frac{\alpha\,\Gamma(\beta)}{2\,\Gamma(1/\alpha)\,\Gamma(\beta-1/\alpha)}(1+|x|^{\alpha})^{-\beta}.$$

The special case $\alpha=2$ and $\beta>1/2$ corresponds to Student's t. A normal scale mixture representation for the generalized Cauchy density with parameters $\alpha\in(0,2]$ and $\beta>1/\alpha$ can be obtained as follows. Let E_α have exponential power distribution with index α, and suppose $W_{\alpha,\beta}$ is a positive random variate independent of E_α and with probability density proportional to $v^{-1/(\alpha\beta)}\exp(-v^{1/\beta})$, $v>0$. Then $X=E_\alpha/W_{\alpha,\beta}^{1/(\alpha\beta)}$ has probability density function proportional to

$$\int_0^\infty \exp(-\mid x\mid^{\alpha} v^{1/\beta}-v^{1/\beta})\,dv=\Gamma(\beta+1)(1+\mid x\mid^{\alpha})^{-\beta}$$

The normal scale mixture representation for the exponential power distribution in Example 2 gives a normal scale mixture representation for the generalized Cauchy density with parameters $\alpha\in(0,2]$ and $\beta>1/\alpha$.

Example 5 Romanowski (1979) studies two classes of normal scale mixtures which he terms *modulated normal* distributions. The *type I modulated normal* distributions are obtained when choosing a Beta $(a, 1)$ distribution, $a>0$, as the mixing distribution F in representation (1). Thus the mixing density is $f(v)=av^{a-1}$ for $v\in(0,1)$, and the resulting normal scale mixture has density function

$$\frac{a}{(2\pi)^{1/2}}\left(\frac{x^2}{2}\right)^{a-\frac{1}{2}}\Gamma\left(-\left(a-\frac{1}{2}\right),\frac{x^2}{2}\right),$$

where Γ denotes the complementary incomplete gamma function.

Romanowski's *modulated normal* distributions of *type II* arise for F a Pareto distribution on $[1,\infty)$ with parameter $b>0$. The Pareto distribution has density b/v^{b+1} for $v>1$, and the resulting normal scale mixture has probability density function

$$\frac{b}{(2\pi)^{1/2}}\left(\frac{x^2}{2}\right)^{-(b+\frac{1}{2})}\gamma\left(b+\frac{1}{2},\frac{x^2}{2}\right); \qquad (17)$$

here γ denotes the incomplete gamma function. For b an integer plus 1/2, the density (17) can be expressed in terms of elementary functions (Gradshteyn and Ryzhik, 1994, formula 8.352.1). If $b=1/2$, for example, the density of the type II modulated normal distribution is

$$(2\pi)^{-1/2}x^{-2}(1-\exp(-x^2/2)).$$

From equation (7) and the densities of the beta and Pareto distribution it is immediate that type I and type II modulated normal densities with parameter $a > 1/2$ and $b = a - 1/2$, respectively, are dual.

The modulated normal distribution of type II is in fact far better known as the *slash* distribution (Rogers and Tukey, 1972; Lange and Sinsheimer, 1993). The slash distribution is generated as X/U, where X is standard normal, U has a Beta$(c, 1)$ distribution, and X, U are independent. A simple calculation shows that the slash distribution and the type II modulated normal distribution coincide when $c = 2b > 0$.

As I. J. Good (1995) noted, the value of knowing that two densities are dual is that knowledge about the one provides knowledge about the other. A striking instance is an uncertainty principle for dual densities similar to the Heisenberg relation. More explicitly, the products of the variances of dual densities p and $\hat{p}$ have a greatest lower bound Λ, and it is known that $0.5276\ldots < \Lambda < 0.857070$ (Dreier, 1996; Gneiting, 1997). If p and $\hat{p}$ are normal scale mixtures, then $\mathrm{Var}(p)\mathrm{Var}(\hat{p}) \geq 1$ with equality if and only if p and $\hat{p}$ are normal probability densities (Gneiting, 1997).

Acknowledgement

The paper is based on a section in the author's PhD thesis (Universität Bayreuth, 1997). The author is grateful to his supervisor Professor Peter J. Huber for continuous support, discussions, and encouragement.

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*. Princeton University Press, Princeton.

Andrews, D. F. and Mallows, C. L. (1974) Scale mixtures of normal distributions. *J. Roy. Statist. Soc. Ser. B*, **36**, 99–102.

Barndorff-Nielsen, O. (1977) Exponentially decreasing distributions for the logarithm of particle size. *Proc. Roy. Soc. London Ser. A*, **353**, 401–419.

Barndorff-Nielsen, O. (1978) Hyperbolic distributions and distributions on hyperbolae. *Scand. J. Statist.*, **5**, 151–157.

Barndorff-Nielsen, O. (1996) Normal inverse Gaussian distributions and stochastic volatility modelling. *Scand. J. Statist.*, **24**, 1–13.

Barndorff-Nielsen, O. and Blaesild, P. (1981) Hyperbolic distributions and ramifications: contributions to theory and applications. In Taillie, C., Patil, G. P. and Baldessari, B. A. (Eds.): *Statistical Distributions in Scientific Work*, **4**, 19–44. Reidel, Dordrecht.

Barndorff-Nielsen, O., Kent, J. and Sørensen, M. (1982) Normal variance-mean mixtures and z distributions. *Internat. Statist. Rev.*, **50**, 145–159.

Box, G. E. P. and Tiao, G. C. (1992) *Bayesian Inference in Statistical Analysis*. Wiley Classics Library edition. Wiley, New York.

Devroye, L. (1986) *Non-Uniform Random Variate Generation*. Springer, New York.

Devroye, L. (1990) A note on Linnik's distribution. *Statist. Probab. Lett.*, **9**, 305–306.

Dreier, I. (1996) On the uncertainty principle for positive definite densities. *Z. Anal. Anwendungen*, **15**, 1015–1023.

Feller, W. (1971) *An Introduction to Probability Theory and Its Applications, Vol.* **2**. Second edition. Wiley, New York.

Gneiting, T. (1997) On the uncertainty relation for positive definite probability densities. To appear in *Statistics*.

Good, I. J. (1995) Dual density functions. *J. Statist. Comput. Simulation*, **52**, 193–194.

Gradshteyn, I. S. and Ryzhik, I. M. (1994) *Table of Integrals, Series, and Products*. Fifth edition. Academic Press, Boston.

Harkness, W. L. and Harkness, M. L. (1968) Generalized hyperbolic secant distribution. *J. Amer. Statist. Assoc.*, **63**, 329–337.

Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994a) *Continuous Univariate Distributions Vol.* **1**. Second edition. Wiley, New York.

Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994b) *Continuous Univariate Distributions, Vol.* **2**. Second edition. Wiley, New York.

Kelker, D. (1971) Infinite divisibility and variance mixtures of the normal distribution. *Ann. Math. Statist.*, **42**, 802–808.

Lange, R. and Sinsheimer, J. S. (1993) Normal/independent distributions and their applications in robust regression. *J. Comput. Graph. Statist.*, **2**, 175–198.

Liu, C. (1996) Bayesian robust multivariate linear regression with incomplete data. *J. Amer. Statist. Assoc.*, **91**, 1219–1227.

McKay, A. T. (1932) A Bessel function distribution. *Biometrika*, **24**, 39–44.

Newcomb, S. (1886) A generalized theory of the combination of observations so as to obtain the best result. *Amer. J. Math.*, **8**, 343–366.

Rogers, W. H. and Tukey, J. W. (1972) Understanding some long-tailed distributions. *Statist. Neerlandica*, **26**, 211–226.

Romanowski, M. (1979) *Random Errors in Observation and the Influence of Modulation on Their Distribution*. Verlag Konrad Wittwer, Stuttgart.

Rossberg, H.-J. (1990) Positive definite probability densities. *Theory Probab. Appl.*, **35**, 169–174.

Rydberg, T. (1996) The normal inverse Gaussian Lévy process: simulation and approximation. Resarch Report 344, Department of Theoretical Statistics, Aarhus University. Revised version to appear in *Comm. Statist. Stochastic Models*.

West, M. (1987) On scale mixtures of normal distributions. *Biometrika*, **74**, 646–648.