The Annals of Statistics
1999, Vol. 27, No. 1, 374–385

GENERALIZED VARIANCE AND EXPONENTIAL FAMILIES

BY ABDELHAMID HASSAIRI

Sfax University

Let μ be a positive measure on $\mathbb{R}^d$ and let $F(\mu) = \{P(\theta, \mu); \theta \in \Theta\}$ be the natural exponential family generated by μ. The aim of this paper is to show that if μ is infinitely divisible then the generalized variance of μ, i.e., the determinant of the covariance operator of $P(\theta, \mu)$, is the Laplace transform of some positive measure $\rho(\mu)$ on $\mathbb{R}^d$. We then investigate the effect of the transformation $\mu \to \rho(\mu)$ and its implications for the skewness vector and the conjugate prior distribution families of $F(\mu)$.

1. Introduction. Consider a distribution μ on $\mathbb{R}^d$. Assume that μ has a Laplace transform,

$$L_\mu(\theta) = \int \exp\langle \theta, x\rangle \mu(dx)$$

for θ in a neighborhood of 0. In Kokonendji and Seshadri (1996), the following result is proved. Let L''_μ denote the (d, d) Hessian matrix of L_μ. Then $\det L''_\mu$ is itself the Laplace transform of some positive measure on $\mathbb{R}^d$ [a far-reaching generalization of the result of Lindsay (1989)]. The aim of the present paper is to give a parallel result to this. Let $k_\mu = \operatorname{Log} L_\mu$ be the cumulant generating function of μ. Assume now that μ is infinitely divisible. Under this circumstance, we prove that there exists a positive measure $\rho(\mu)$ on $\mathbb{R}^d$ such that

$$\det k''_\mu(\theta) = L_{\rho(\mu)}(\theta). \tag{1.1}$$

Let us recall a few facts from the literature about this function $\det k''_\mu(\theta)$. Wilks (1932) calls $\det k''_\mu(0)$ the generalized variance of μ since $k''_\mu(0)$ is the covariance matrix of μ. Consider the natural exponential family (NEF) $F(\mu)$ generated by μ, that is, the set of distributions

$$P(\theta, \mu)(dx) = \exp\left[\langle \theta, x\rangle - k_\mu(\theta)\right] \mu(dx),$$

where θ belongs to $\Theta(\mu)$ defined as the interior of $\{\theta, L_\mu(\theta) < +\infty\}$. Then, for θ in $\Theta(\mu)$, $\det k''_\mu(\theta)$ is the generalized variance of $P(\theta, \mu)$. Also,

$$u(\theta) = \frac{d}{d\theta}\left[\log \det k''_\mu(\theta)\right] \tag{1.2}$$

is called the skewness vector of $F(\mu)$ [see Gutiérrez-Peña (1995)].

Received July 1995; revised October 1998.
AMS 1991 *subject classifications*. Primary 62E10; secondary 62H05.
Key words and phrases. Natural exponential family, variance function, generalized variance, skewness vector.

The third fact concerns the dimension 1. It is well known [see Letac (1992) or Kokonendji and Seshadri (1994)] that a distribution μ on $\mathbb{R}$ such that $\Theta(\mu) \neq \emptyset$ is infinitely divisible if and only if $k''_\mu(\theta)$ is the Laplace transform of some positive measure $\rho(\mu)$. Thus, our main result is the extension to $\mathbb{R}^d$ of the second part of this statement. Furthermore, still for dimension 1, and whether μ is infinitely divisible or not, Kokonendji and Seshadri (1994) have proved that $k''_\mu L^2_\mu$ is the Laplace transform of some positive measure $T(\mu)$. It is therefore interesting to remark that if, furthermore, μ is infinitely divisible, then $T(\mu) = \mu * \mu * \rho(\mu)$. One can also observe that if the variance function $V_{F(\mu)}$ of the NEF $F(\mu)$ is quadratic [see Morris (1982)], both $V_{F(T(\mu))}$ and $V_{F(\rho(\mu))}$ are quadratic. If $V_{F(\mu)}$ is cubic [see Letac and Mora (1990)], then both $V_{F(T(\mu))}$ and $V_{F(\rho(\mu))}$ are polynomials of degree 4 in $\sqrt{m}$.

Our paper has the following plan. In Section 2 we state our main result, whose long and technical proof is postponed to Section 3. In Section 2 we also study properties of the map $\mu \to \rho(\mu)$, which is defined on the set of infinitely divisible distributions in $\mathbb{R}^d$. In particular, if $F = F(\mu)$ and $\overline{F} = F(\rho(\mu))$, Theorem 2.3 links the variance functions $V_{\overline{F}}$ and V_F. This provides a way, in certain cases, to obtain $\rho(\mu)$ explicitly, or to discover new multivariate simple structures. Another surprising result is the fact that the skewness vector of $P(\theta, \mu)$ is nothing but the mean vector of $P(\theta, \rho(\mu))$. This leads Corollary 2.5 to a new characterization, in terms of $\rho(\mu)$, of the NEF such that the conjugate families obtained with the canonical parameter θ coincide with the conjugate families obtained when the parameter is the mean. [See Consonni and Veronese (1992) for results in the one-dimensional case and Gutiérrez-Peña (1995) for the d-dimensional case.]

2. The generalized variance transform. Let $\mathscr{M}_d$ be the set of positive measures μ on $\mathbb{R}^d$ not concentrated on a hyperplane and such that $\Theta(\mu) = \text{interior}\{\theta;\ L_\mu(\theta) < \infty\}$ is nonempty. In this section, we first state the main theorem related to the existence of $\rho(\mu)$ for μ in $\mathscr{M}_d$ infinitely divisible. We then introduce and study the generalized variance transform.

THEOREM 2.1. *Let μ be in $\mathscr{M}_d$. If μ is infinitely divisible then there exists a positive measure $\rho(\mu)$ on $\mathbb{R}^d$ such that*

$$\det k''_\mu(\theta) = L_{\rho(\mu)}(\theta) \quad \textit{for all } \theta \textit{ in } \Theta(\mu).$$

COMMENTS. (i) The measure $\rho(\mu)$ is not always in $\mathscr{M}_d$. This can be seen by taking the measure generating the Poisson family on $\mathbb{R}^d$,

$$\mu(dx) = \sum_{k \in \mathbb{N}^d} \frac{1}{k!} \delta_k(dx),$$

where $k! = k_1! k_2! \cdots k_d!$. Here $\Theta(\mu) = \mathbb{R}^d$, $k_\mu(\theta) = \sum_{i=1}^d e^{\theta_i}$ and $L_{\rho(\mu)}(\theta) = \det k''_\mu(\theta) = \exp(\sum_{i=1}^d \theta_i)$. Thus the measure $\rho(\mu)$ is concentrated on the point $\mathbf{1} = (1, 1, \dots, 1)$ of $\mathbb{R}^d$.

(ii) As mentioned above, if d is one, the infinite divisibility of μ is a necessary condition for the existence of a measure $\rho(\mu)$ satisfying (1.1). This need not be the case when $d > 1$. Consider, in fact, the Wishart family of distributions on the space S_d of symmetric real (d, d) matrices generated by the measure μ such that $-\Theta(\mu) = S_d^+$, the cone of positive definite elements of S_d, and

$$L_\mu(\theta) = (\det(-\theta))^{-1}.$$

It is known that the measure μ is not infinitely divisible [see Letac (1992)]. However,

$$\det k''_\mu(\theta) = (\det(-\theta))^{-(d+1)} = L_{\mu_{(d+1)}}(\theta),$$

where $\mu_{(d+1)} = \mu^{*(d+1)}$.

(iii) Let $(e_1, e_2, \dots, e_d)$ be the canonical basis of $\mathbb{R}^d$. The measure $\mu = \delta_0 + \delta_{e_1} + \cdots + \delta_{e_d}$ generating the multinomial family on $\mathbb{R}^d$ provides an example of a measure μ which is not infinitely divisible and such that $\det k''_\mu(\theta)$ is not a Laplace transform. Actually

$$k_\mu(\theta) = \operatorname{Log}\left(1 + \sum_{i=1}^{d} e^{\theta_i}\right),$$

and a simple calculation gives

$$\det k''_\mu(\theta) = \left(1 + \sum_{i=1}^{d} e^{\theta_i}\right)^{-(d+1)} \exp(\theta_1 + \theta_2 + \cdots + \theta_d).$$

It is easy to show that this is not the Laplace transform of a positive measure.

Definition 2.1. If $\mathscr{G}$ is the set of measures μ in $\mathscr{M}_d$ such that $\rho(\mu)$ exists, then the transform $\mu \to \rho(\mu)$, defined on $\mathscr{G}$, will be called the generalized variance transform.

We now examine the effect of the generalized variance transform $\mu \to \rho(\mu)$ when the transformed measure $\rho(\mu)$ is also in $\mathscr{M}_d$. Our aim is to obtain some information on the NEF generated by $\rho(\mu)$. We say that two NEFs, $F(\mu)$ and $F(\nu)$, on $\mathbb{R}^d$ are in the same G_0-orbit if the first can be obtained from the other by an affine transformation and a power of convolution. Given that a NEF is usually defined by one element of its basis and that the classifications of NEF are done up to G_0-orbits [see Morris (1982), Letac and Mora (1990) and Hassairi (1992)], we first prove the following result.

Proposition 2.2. *Let μ in $\mathscr{G}$ be such that $\rho(\mu)$ is in $\mathscr{M}_d$.*

(i) *If μ' is a basis of $F(\mu)$ then $F(\rho(\mu')) = F(\rho(\mu))$.*

(ii) *If μ' is such that $F(\mu')$ is in the G_0-orbit of $F(\mu)$, then $F(\rho(\mu))$ and $F(\rho(\mu'))$ are also in the same G_0-orbit.*

Proof. (i) If μ' is a basis of the NEF $F(\mu)$, then there exist a in $\mathbb{R}^d$ and b in $\mathbb{R}$ such that

$$k_{\mu'}(\theta) = k_\mu(\theta + a) + b \quad \text{for } \theta \text{ in } \Theta(\mu').$$

Hence

$$\det k''_{\mu'}(\theta) = \det\big(k''_{\mu}(\theta + a)\big).$$

Thus, we obtain

$$L_{\rho(\mu')}(\theta) = L_{\rho(\mu)}(\theta + a) \qquad \forall\, \theta \in \Theta(\rho(\mu')),$$

which implies that $F(\rho(\mu')) = F(\rho(\mu))$.

(ii) Suppose that $F(\mu')$ and $F(\mu)$ are in the same G_0-orbits. Then there exist an affine transformation of $\mathbb{R}^d$, $\varphi(x) = \delta(x) + \gamma$ and $\alpha > 0$ such that

$$k_{\mu'}(\lambda) = \alpha k_{\mu}(\delta^*(\lambda)) + \langle \lambda, \gamma\rangle \qquad \forall\, \lambda \in \delta^{*-1}(\Theta(\mu)).$$

Hence

$$L_{\rho(\mu')}(\lambda) = \alpha^d (\det \delta)^2 L_{\rho(\mu)}(\delta^*(\lambda)),$$

which implies that $F(\rho(\mu))$ is the image of $F(\rho(\mu'))$ under the linear map δ. □

Recall now that $M_F = k'_{\mu}(\Theta(\mu))$ is called the domain of the means of the NEF $F = F(\mu)$. Since μ is in $\mathscr{M}_d$, k_μ is strictly convex and $\theta \to k'_\mu(\theta)$ is a bijection between $\Theta(\mu)$ and M_F. The map $m \to \psi_\mu(m)$ will denote its inverse function from M_F to $\Theta(\mu)$. The map from M_F to $L_S(\mathbb{R}^d)$, defined by $m \to V_F(m) = k''_\mu(\psi_\mu(m))$, is called the variance function of the NEF. It is easily proved that $V_F(m) = [\psi'_\mu(m)]^{-1}$ and an important feature of V_F is that it characterizes F in the following sense: if F_1 and F_2 are two NEFs whose variance function coincide on a nonempty open set of $M_{F_1} \cap M_{F_2}$, then $F_1 = F_2$.

The following theorem gives the link between the variance function of the NEF generated by the measure μ and the variance function of the transformed NEF generated by $\rho(\mu)$. If $V_{F(\mu)}(m)$ belongs to a certain class, then the generalized variance transform produces a new class of variance functions. For example, if $F(\mu)$ is in the Mora class of NEF in $\mathbb{R}^d$ with strictly cubic variance function [Hassairi (1993)], then $F(\rho(\mu))$ will be in the extension to $\mathbb{R}^d$ of the class of NEF's in $\mathbb{R}$ with polynomial variance function of degree 4 in $\sqrt{m}$. Note also that $V_{F(\rho(\mu))}$ represents an important tool in the determination of $F(\rho(\mu))$; it permits, in certain cases, explicit determination of $\rho(\mu)$.

THEOREM 2.3. *Let μ in $\mathscr{G}$ be such that $\rho(\mu)$ is in $\mathscr{M}_d$. If we define $F = F(\mu)$, $\overline{F} = F(\rho(\mu))$ and $\overline{m} = k'_{\rho(\mu)}(\psi_\mu(m))$ for m in M_F, then*:

(i) $\overline{m} = (V_F(m))(V'_F(m))^*(V_F(m))^{-1}$.

(ii) *If $(e_i)_{1 \le i \le d}$ is an orthonormal basis of $\mathbb{R}^d$, then*

$$\overline{m} = \sum_{i=1}^{d} (V'_F(m)e_i)e_i. \tag{2.1}$$

(iii) $\forall\ m \in M_F$,

$$V_{\overline{F}}(\overline{m}) = \frac{d\overline{m}}{dm} V_F(m). \tag{2.2}$$

To illustrate the last result concerning the variance function of $F(\rho(\mu))$, consider the inverse Gaussian NEF generated by the measure defined on a point (x_1, x') of $]0, +\infty[\times\mathbb{R}^{d-1}$ by

$$\mu(dx_1, dx') = \frac{1}{(2\pi)^{d/2} x_1^{(d+2)/2}} \exp\left[-\frac{1}{2x_1}(1 + \|x'\|^2)\right] \mathbf{1}_{]0,+\infty[}(x_1)\, dx_1\, dx'.$$

The variance function of $F(\mu)$ is

$$V_{F(\mu)}(m) = \langle e_1, m\rangle[m \otimes m + I_d - e_1 \otimes e_1]$$

[see Hassaïri (1992)]. A direct calculation, using (2.1) and (2.2), shows that $\overline{F} = F(\rho(\mu))$ has the simple quadratic variance function

$$V_{\overline{F}}(\overline{m}) = \frac{2}{d+2}\overline{m} \otimes \overline{m} + \langle e_1, \overline{m}\rangle(I_d - e_1 \otimes e_1).$$

Therefore $\overline{F}$ is one of the quadratic NEF's described by Casalis (1996).

Now consider the skewness vector $u(\theta)$ of a NEF $F = F(\mu)$ defined by (1.2). It may be expressed in terms of the mean parameter m as

$$\tilde{u}(m) = u(\psi_\mu(m)) = \frac{d}{dm}[\operatorname{Log}\det V_F(m)]V_F(m).$$

Theorem 2.3 provides two alternative expressions for $\tilde{u}(m)$.

COROLLARY 2.4. *Let μ be in $\mathscr{M}_d$. Then*

$$\tilde{u}(m) = V_{F(\mu)}(m)(V_F'(m))^*(V_F(m))^{-1} = \sum_{i=1}^{d}(V_F'(m)e_i)e_i,$$

where (e_i) is any orthonormal basis of $\mathbb{R}^d$.

Furthermore, if μ is in $\mathscr{G}$, in particular if μ is infinitely divisible, then the skewness vector of the distribution $P(\theta, \mu)$ is equal to the mean vector of the distribution $P(\theta, \rho(\mu))$.

Another corollary of Theorem 2.3 is related to Bayesian theory and concerns the equality of two conjugate prior distribution families of a NEF $F = F(\mu)$. Let Π be the family of prior distributions on $\Theta(\mu)$ introduced by Diaconis and Ylvisaker (1979),

$$\pi_{t,m_0}(d\theta) = C_{t,m_0}\exp t[\langle\theta, m_0\rangle - k_\mu(\theta)]\mathbf{1}_{\Theta(\mu)}(\theta)\, d\theta,$$

where $m_0 \in M_F$, $t > 0$ and C_{t,m_0} is a normalizing constant. The image of Π by k_μ' is a family of prior distributions on the domain of the means M_F of the NEF F. Besides $k_\mu'(\Pi)$, Consonni and Veronese (1992) consider another family Π^* of prior distributions on M_F,

$$\pi_{t,m_0}^*(dm) = C_{t,m_0}^*\exp t[\langle\psi_\mu(m), m_0\rangle - k_m(\psi_\mu(m))]\mathbf{1}_{M_F}(m)\, dm.$$

Casalis (1996) states that $k_\mu'(\Pi) = \Pi^*$ if and only if there exists B in $\mathbb{R}^d$ and a in $\mathbb{R}$ such that, for any basis $(e_i)_{i=1}^d$ of $\mathbb{R}^d$,

$$\sum_{i=1}^{d}(V_F'(m)e_i)e_i = am + B, \qquad m \in M_F.$$

Provided that the generalized variance transform of μ exists, we have the following criteria for the equality of $k'_\mu(\Pi)$ and Π^*.

COROLLARY 2.5. *Let μ in $\mathscr{G}$ be such that $\rho(\mu)$ is also in $\mathscr{M}_d$. Then $k'_\mu(\Pi) = \Pi^*$ if and only if $F(\mu)$ leads to $F(\rho(\mu))$ by a translation and a power of convolution.*

PROOF OF THEOREM 2.3. (i) For $\theta \in \Theta(\mu) = \Theta(\rho(\mu))$, we have $L_{\rho(\mu)}(\theta) = \det k''_\mu(\theta)$. Hence

$$k_{\rho(\mu)}(\theta) = \operatorname{Log}\det k''_\mu(\theta).$$

If $L_S(\mathbb{R}^d)$ is equipped with the inner product $\langle a, b\rangle_{L_S(\mathbb{R}^d)} = \operatorname{trace}(a, b)$, then by differentation, one gets

$$\begin{aligned}\left\langle k'_{\rho(\mu)}(\theta), \cdot\right\rangle_{\mathbb{R}^d} &= \left\langle \left(k''_\mu(\theta)\right)^{-1}, k'''_\mu(\theta)(\cdot)\right\rangle_{L_S(\mathbb{R}^d)} \\ &= \left\langle \left(k'''_\mu(\theta)\right)^* \left(k''_\mu(\theta)\right)^{-1}, \cdot\right\rangle_{\mathbb{R}^d},\end{aligned} \tag{2.3}$$

which implies $k'_{\rho(\mu)}(\theta) = (k'''_\mu(\theta))^*(k''_\mu(\theta))^{-1}$. Let $\theta = \psi_\mu(m)$. Then

$$\overline{m} = k'_{\rho(\mu)}\big(\psi_\mu(m)\big) = \big(k'''_\mu\big(\psi_\mu(m)\big)^* \big(k''_\mu\big(\psi_\mu(m)\big)\big)^{-1}. \tag{2.4}$$

Since $V_F(m) = k''_\mu(\psi_\mu(m)) = (\psi'_\mu(m))^{-1}$, we have $V'_F(m) = k'''_\mu(\psi_\mu(m))\psi'_\mu(m)$. Inserting this in (2.4) yields

$$\overline{m} = V_F(m)\big(V'_F(m)\big)^*\big(V_F(m)\big)^{-1}.$$

(ii) From (2.3) and (2.4), we have

$$\langle \overline{m}, \alpha\rangle = \operatorname{trace}\Big((V(m))^{-1}V'(m)V(m)\alpha\Big).$$

Therefore,

$$\langle \overline{m}, \alpha\rangle = \sum_{i=1}^d \Big\langle\Big((V(m))^{-1}V'(m)V(m)\alpha\Big)e_i, e_i\Big\rangle.$$

We now use the following property of symmetry of a variance function:

$$(V'(m)V(m)\alpha)\beta = (V'(m)V(m)\beta)\alpha) \quad \text{for all } \alpha \text{ and } \beta.$$

Thus

$$\begin{aligned}\langle \overline{m}, \alpha\rangle &= \sum_{i=1}^d \Big\langle\Big((V(m))^{-1}V'(m)V(m)e_i\Big)\alpha, e_i\Big\rangle \\ &= \sum_{i=1}^d \Big\langle (V'(m)V(m)e_i)\alpha, (V(m))^{-1}e_i\Big\rangle \\ &= \sum_{i=1}^d \Big\langle (V'(m)V(m)e_i)(V(m))^{-1}e_i, \alpha\Big\rangle \\ &= \sum_{i=1}^d \Big\langle (V'(m)e_i)e_i, \alpha\Big\rangle.\end{aligned}$$

Hence

$$\overline{m} = \sum_{i=1}^{d} (V'(m)e_i)e_i.$$

(iii) Since $\overline{m} = k'_{\rho(\mu)}(\psi_\mu(m))$, then $\psi_{\rho(\mu)}(\overline{m}) = \psi_\mu(m)$. Differentiating with respect to m yields

$$\psi'_{\rho(\mu)}(\overline{m})\frac{d\overline{m}}{dm} = \psi'_\mu(m).$$

This is equivalent to the desired result. □

3. Proof of Theorem 2.1. This section is entirely devoted to the proof of Theorem 2.1. In order to do so, we need to introduce some other notations and to establish several results on determinant calculations. If $A = (a_{ij})_{1 \le {i \atop j} \le d}$ is a (d, d)-matrix and T is a subset of $\{1, 2, \dots, d\}$, we denote by A_T the matrix $(a_{ij})_{(i,j) \in T \times T}$ and by $\det A_T$ the determinant of A_T with $\det A_\phi = 1$.

The following proposition is classical and stated without proof.

PROPOSITION 3.1. *Let A and B be two (d, d)-matrices. If A is diagonal, then*

$$\det(A + B) = \sum_{T \subset \{1, 2, \dots, d\}} \det A_{T'} \det B_T,$$

where $T' = \{1, 2, \dots, d\} \setminus T$.

We now prove a theorem concerning the expectation of a determinant. This theorem is stated without proof in the paper of Kokonendji and Seshadri (1996) and appears as a problem due to Pólya and Szegö (1972).

THEOREM 3.2. *Let E be an oriented Euclidean space and η a positive measure on $E \times E$ such that*

$$\int_{E \times E} \|X\| \, \|Y\| \eta(dX, dY) < +\infty.$$

If $\dim E = d$, *then*

$$\det\left[\int_{E \times E} X \otimes Y \eta(dX, dY)\right] \tag{3.1}$$
$$= \frac{1}{d!}\int_{(E \times E)^d} \det[X^{(1)}, \dots, X^{(d)}] \det[Y^{(1)}, \dots, Y^{(d)}] \times \eta(dX^{(1)}, dY^{(1)}) \cdots \eta(dX^{(d)}, dY^{(d)}).$$

Recall that, in an oriented Euclidean space E, $\det[X^{(1)}, \dots, X^{(d)}]$ is defined as the determinant of the matrix $([X^{(1)}]^{(e)}, \dots, [X^{(d)}]^{(e)})$ where (e) is any direct orthonormal basis of E; this number is independent of the choice of (e).

We need now the following lemma.

LEMMA 3.3. *Let $F = F_1 \times F_2 \times ,\dots, F_d$ be the product of d finite-dimensional linear spaces and let B: $F \to \mathbb{R}$ be a multilinear form on F. If ν is a positive measure on F and $f = (f_1, f_2, \dots, f_d)$ is a mapping from F to F such that $\int_F f\nu(df)$ exists, then*

$$B\left(\int_F f\nu(df)\right) = B\left(\int_F f_1\nu(df),\dots,\int_F f_d\nu(df)\right)$$
$$= \int_{F^n} B\big(f_1^{(1)},\dots,f_d^{(d)}\big)\nu\big(df^{(1)}\big)\cdots\ \nu\big(df^{(d)}\big).$$

PROOF. We use induction on d. It is obvious for $d = 1$. Suppose the result true for $d - 1$ and denote $m_j = \int_F f_j\nu(df)$. Then

$$B(m_1, m_2,\dots, m_d) = B\left(m_1,\dots,m_{d-1},\int_F f_d\nu(df)\right).$$

Using the linearity of B,

$$= \int_F \nu(df)\left[\int_{F^{n-1}} B\big(f_1^{(1)}, f_2^{(1)},\dots, f_{d-1}^{(d-1)}, f_d\big)\nu\big(df^{(1)}\big)\cdots\ \nu\big(df^{(d-1)}\big)\right]$$
$$= \int_{F^n} B\big(f_1^{(1)}, f_2^{(1)},\dots, f_{d-1}^{(d-1)}f_d\big)\nu\big(df^{(1)}\big)\cdots\ \nu\big(df^{(d-1)}\big)\nu(df). \qquad \square$$

PROOF OF THEOREM 3.2. Let $(e) = (e_1, e_2,\dots, e_d)$ be an orthonormal basis in E. If the elements of $\mathbb{R}^d$ are written as column matrices, then the space E is identified with $\mathbb{R}^d$ by means of the map $\sum_{i=1}^d X_i e_i \to X = {}^t(X_1, X_2,\dots, X_d)$. Hence, in the basis (e), the equality (3.1) becomes

$$\begin{aligned}&\det\int_{\mathbb{R}^d\times\mathbb{R}^d} X\otimes Y\eta(dX, dY)\\ &\quad= \frac{1}{d!}\int_{(\mathbb{R}^d\times\mathbb{R}^d)^d}\det[X^{(1)},\dots,X^{(d)}]\det[Y^{(1)},\dots,Y^{(d)}]\\ &\qquad\times\eta(dX^{(1)}, dY^{(1)}),\dots,\eta(dX^{(d)}, dY^{(d)}).\end{aligned} \tag{3.2}$$

We apply Lemma 3.3 to $F_j = \mathbb{R}^d$, F being the space of the (d, d)-matrices and $B(f) = \det f$. We have

$$\det\left(\int_{\mathbb{R}^d\times\mathbb{R}^d} X\otimes Y\eta(dX, dY)\right) = \int_{(\mathbb{R}^d\times\mathbb{R}^d)^d}\det\big[X^{(1)}Y_1^{(1)},\dots,X^{(d)}Y_d^{(d)}\big]$$
$$\times\ \eta(dX^{(1)}, dY^{(1)})\cdots\ \eta(dX^{(d)}, dY^{(d)})$$

because the ith column of the matrix $X^{(i)}\otimes Y^{(i)}$ is $X^{(i)}Y_i^{(i)}$. Hence the left-hand side of (3.2) is equal to $\int_{(\mathbb{R}^d\times\mathbb{R}^d)^d} Y_1^{(1)}Y_2^{(2)}\cdots\ Y_d^{(d)}\det[X^{(1)},\dots, X^{(d)}]\eta(dX^{(1)}, dY^{(1)})\cdots\ \eta(dX^{(d)}, dY^{(d)})$.

Let us now examine the right-hand side of (3.2). If we denote it by S, we have

$$S = \frac{1}{d!}\int_{(\mathbb{R}^d\times\mathbb{R}^d)^d} \det\left[\langle X^{(i)}, Y^{(i)}\rangle_{i\le {}^{i}_{j}\le d}\right]\eta(dX^{(1)}, dY^{(1)})\cdots\ \eta(dX^{(d)}, dY^{(d)})$$

$$= \frac{1}{d!}\int_{(\mathbb{R}^d\times\mathbb{R}^d)^d} \det\left[\left(\sum_{k=1}^{d} X_k^{(i)}Y_k^{(j)}\right)\right]\eta(dX^{(1)}, dY^{(1)})\cdots\ \eta(dX^{(d)}, dY^{(d)}).$$

Using the multilinearity of det and denoting by $X_k^{(i)}$ the vector ${}^t(X_{k_1}^{(i)}, X_{k_2}^{(i)}, \ldots, X_{k_d}^{(i)})$, we obtain

$$S = \frac{1}{d!}\sum_{k_1, k_2, \ldots, k_d = 1}^{d}\int_{(\mathbb{R}^d\times\mathbb{R}^d)^d}\prod_{j=1}^{d} Y_{k_j}^{(j)} \det\left[X_k^{(1)}, X_k^{(2)}, \ldots, X_k^{(d)}\right]$$

$$\times\,\eta(dX^{(1)}, dY^{(1)})\cdots\ \eta(dX^{(d)}, dY^{(d)}).$$

Now $\det[X_k^{(1)}, X_k^{(2)}, \ldots, X_k^{(d)}]$ is equal to $\det[X_\sigma^{(1)}, X_\sigma^{(2)}, \ldots, X_\sigma^{(d)}]$ if $(k_1, \ldots, k_d)$ is a permutation of σ and is equal to zero otherwise. Hence

$$S = \frac{1}{d!}\sum_{\sigma}\int_{(\mathbb{R}^d\times\mathbb{R}^d)^d}\prod_{j=1}^{d} Y_{\sigma(j)}^{(j)} \det\left[X_\sigma^{(1)}, X_\sigma^{(2)}, \ldots, X_\sigma^{(d)}\right] \tag{3.3}$$

$$\times\ \eta(dX^{(1)}, dY^{(1)})\cdots\ \eta(dX^{(d)}, dY^{(d)}).$$

Introducing the permutation $\tau = \sigma^{-1}$, we have

$$\prod_{j=1}^{d} Y_{\sigma(j)}^{(j)} \det\left[X_\sigma^{(1)}, X_\sigma^{(2)}, \ldots, X_\sigma^{(d)}\right] = \prod_{j=1}^{d} Y_j^{\tau(j)} \det\left[X_1^{\tau(1)}, X_2^{\tau(2)}, \ldots, X_d^{\tau(d)}\right].$$

To complete the proof, we observe that the measure $\eta(dX^{(1)}, dY^{(1)})\cdots \eta(dX^{(d)}, dY^{(d)})$ is invariant by the permutation τ defined on $(\mathbb{R}^d\times\mathbb{R}^d)^d$ by

$$\tau\left[(X^{(1)}, Y^{(1)}), (X^{(2)}, Y^{(2)}), \ldots, (X^{(d)}, Y^{(d)})\right]$$

$$= \left[(X^{\tau(1)}, Y^{\tau(1)}), (X^{\tau(2)}, Y^{\tau(2)}), \ldots, (X^{\tau(d)}, Y^{\tau(d)})\right].$$

Therefore, the measure

$$\prod_{j=1}^{d} Y_j^{\tau(j)} \det\left[X_1^{\tau(1)}\ldots X_d^{\tau(d)}\right]$$

$$\times\ \eta(dX^{(1)}, dY^{(1)})\eta(dX^{(2)}, dY^{(2)}), \ldots, \eta(dX^{(d)}, dY^{(d)})$$

does not depend on τ, and using (3.3), we obtain

$$S = \int_{(\mathbb{R}^d\times\mathbb{R}^d)}(Y_1^{(1)}\ \cdots\ Y_d^{(d)})\det[X^{(1)}, X^{(2)}, \ldots, X^{(d)}]$$

$$\times\ \eta(dX^{(1)}, dY^{(1)})\cdots\ \eta(dX^{(d)}, dY^{(d)}). \qquad \square$$

We are now in position to prove Theorem 2.1. Since μ is infinitely divisible, there exist a positive Σ in $L_S(\mathbb{R}^d)$ and a positive measure ν such that

$$k''_\mu(\theta) = \Sigma + \int_{\mathbb{R}^d} X\otimes X\exp\langle\theta, X\rangle\nu(dX).$$

[For a proof, see Gikhman and Skorohod (1973), page 342.] Since Σ is symmetric and positive, there exists in $\mathbb{R}^d$ an orthonormal basis $(e) = (e_1, \dots, e_d)$ in which Σ is represented by a diagonal matrix A. Also,

$$\det k''_\mu(\theta) = \det\left[A + \int_{\mathbb{R}^d} X \otimes X \exp\langle \theta, X\rangle \nu(dX)\right].$$

We next apply Proposition 3.1 with $B = \int_{\mathbb{R}^d} X \otimes X \exp\langle \theta, X\rangle \nu(dX)$. Let $T = \{i_1, i_2, \dots, i_k\}$, with $1 \le i_1 < i_2 < \cdots < i_k \le d$, a nonempty subset of $\{1, 2, \dots, d\}$ and $\tau_T\colon \mathbb{R}^d \to \mathbb{R}^k$ the map defined by

$$\tau_T(X_1, X_2, \dots, X_d) = (X_{i_1}, X_{i_2}, \dots, X_{i_k}).$$

Then, for X and Y in $\mathbb{R}^d$, one has

$$(X \otimes Y)_T = (\tau_T(X)) \otimes (\tau_T(Y)). \tag{3.4}$$

Now, if we consider the measure defined on $\mathbb{R}^d \times \mathbb{R}^d$ by

$$\eta(dX, dY) = \exp\langle \theta, X\rangle \nu(dX)\, \delta_X(dY), \tag{3.5}$$

then

$$B = \int_{\mathbb{R}^d} X \otimes X \exp\langle \theta, X\rangle \nu(dX) = \int_{\mathbb{R}^d \times \mathbb{R}^d} X \otimes Y \eta(dX, dY).$$

Using (3.4), we have

$$B_T = \int_{\mathbb{R}^d \times \mathbb{R}^d} \tau_T(X) \otimes \tau_T(Y) \eta(dX, dY).$$

In order to use Theorem 3.2, we introduce the measure $\tilde{\eta}$ as the image of η under the function from $\mathbb{R}^d \times \mathbb{R}^d$ into $\mathbb{R}^k \times \mathbb{R}^k$ given by $(X, Y) \to (\tau_T(X), \tau_T(Y))$. Then

$$B_T = \int_{\mathbb{R}^k \times \mathbb{R}^k} X \otimes Y \tilde{\eta}(dX, dY),$$

and Theorem 3.2 implies that

$$\det B_T = \frac{1}{k!} \int_{(\mathbb{R}^k \times \mathbb{R}^k)^k} \det[X^{(1)}, X^{(2)}, \dots, X^{(k)}] \det[Y^{(1)}, Y^{(2)}, \dots, Y^{(k)}] \times \tilde{\eta}(dX^{(1)}, dY^{(1)}) \cdots \tilde{\eta}(dX^{(k)}, dY^{(k)}).$$

By a similar argument, this becomes

$$\det B_T = \frac{1}{k!} \int_{(\mathbb{R}^d \times \mathbb{R}^d)^k} \det\left[\tau_T(X^{(1)}), \dots, \tau_T(X^{(k)})\right] \det\left[\tau_T(Y^{(1)}), \dots, \tau_T(Y^{(k)})\right] \times \eta(dX^{(1)}, dY^{(1)}) \cdots \eta(dX^{(k)}, dY^{(k)}).$$

Using (3.5) again, we obtain

$$\det B_T = \frac{1}{k!} \int_{(\mathbb{R}^d)^k} \left(\det\left[\tau_T(X^{(1)}), \dots, \tau_T(X^{(k)})\right]\right)^2 \times \exp\left\langle \theta, X^{(1)} + X^{(2)} + \cdots + X^{(k)}\right\rangle \nu(dX^{(1)}) \cdots \nu(dX^{(k)}).$$

Denoting by $T(\nu)$ the image of the measure

$$\frac{1}{k!}\left(\det\left[\tau_T(X^{(1)}),\dots,\tau_T(X^{(k)})\right]^2\right)\nu(dX^{(1)})\cdots\nu(dX^{(k)})$$

under the mapping from $(\mathbb{R}^d)^k$ to $\mathbb{R}^d$ defined by $(X^{(1)}, X^{(2)},\dots, X^{(k)}) \to X^{(1)} + X^{(2)} + \cdots + X^{(k)}$, we have

$$\det B_T = \int_{\mathbb{R}^d} \exp\langle\theta, X\rangle T(\nu)(dX).$$

Finally,

$$\det k''_\mu(\theta) = \det A + \int_{\mathbb{R}^d} \exp\langle\theta, X\rangle\left(\sum_{\substack{T\subset\{1,2,\dots,d\}\\ T\neq\varnothing}} \det A_{T'}T(\nu)\right)(dX).$$

Thus the measure

$$\rho(\mu) = (\det A)\delta_0 + \sum_{\substack{T\subset\{1,\dots,d\}\\ T\neq\varnothing}} \det A_{T'}T(\nu)$$

is such that $\det k''_\mu(\theta) = L_{\rho(\mu)}(\theta)$ and the proof of Theorem 2.1 is complete. □

Acknowledgments. I am very grateful to Gérard Letac for his constructive comments. I sincerely thank the Editor, Associate Editor and a referee for their suggestions and comments that led to significant improvements in the paper.

REFERENCES

BARNDORFF-NIELSEN, O. (1978). *Information and Exponential Families in Statistical Theory*. Wiley, New York.

CASALIS, M. (1996). The $(2d + 4)$ simple quadratic natural exponential families on $\mathbb{R}^d$. *Ann. Statist.* **24** 1828–1854.

CONSONNI, G. and VERONESE, P. (1992). Conjugate priors for exponential families having quadratic variance function. *J. Amer. Statist. Assoc.* **87** 1123–1127.

DIACONIS, P. and YLVISAKER, D. (1979). Conjugate priors for exponential families. *Ann. Statist.* **7** 269–281.

GIKHMAN, I. I. and SKOROHOD, A. V. (1973). *The Theory of Stochastic Processes* **2**. Springer, New York.

GUTIÉRREZ-PEÑA, E. (1995). Bayesian topics relating to the exponential family. Ph.D. thesis, Imperial College, London.

HASSAIRI, A. (1992). La classification des familles exponentielles naturelles sur $\mathbb{R}^n$ par l'action du groupe linéaire de $\mathbb{R}^{n+1}$. *C.R. Acad. Sci. Paris Sér. I* **315** 207–210.

HASSAIRI, A. (1993). Les $(d + 3)$ G-orbites de la classe de Morris-Mora des familles exponentielles de $\mathbb{R}^d$. *C.R. Acad. Sci. Paris Sér. I* **317** 887–890.

KOKONENDJI, C. and SESHADRI, V. (1994). The Lindsay transform of natural exponential families. *Canad. J. Statist.* **22** 259–272.

KOKONENDJI, C. and SESHADRI, V. (1996). On the determinant of the second derivative of the Laplace transform. *Ann. Statist.* **24** 1813–1827.

LETAC, G. (1992). *Lectures on Natural Exponential Families and Their Variance-Functions*. IMPA, Rio de Janeiro.

LETAC, G. and MORA, M. (1990). Natural real exponential families with cubic variance functions. *Ann. Statist.* **18** 1–37.
LINDSAY, B. G. (1989). On the determinant of moment matrices. *Ann. Statist.* **17** 711–721.
MORRIS, C. N. (1982). Natural exponential families with quadratic variance functions. *Ann. Statist.* **10** 65–80.
PÓLYA, G. and SZEGÖ, G. (1972). *Problems and Theorems in Analysis* **1**. Springer, Berlin.
WILKS, S. S. (1932). Certain generalizations in the analysis of variance. *Biometrika* **24** 471–794.

FACULTÉ DES SCIENCES
UNIVERSITÉ DE SFAX
ROUTE DE SOUKRA
3038 SFAX
TUNISIA