A Characterization of Poisson–Gaussian Families by Convolution-Stability

A. E. Koudou

Institut Elie Cartan, Vandœuvre-lès-Nancy Cedex, France
E-mail: koudou@iecn.u-nancy.fr

and

D. Pommeret

CREST (L.S.M.), ENSAI, Bruz, France
E-mail: pommeret@ensai.fr

Received December 2, 1999; published online December 12, 2001

If the convolution of natural exponential families on $\mathbb{R}^d$ is still a natural exponential family, then the families are all Poisson–Gaussian, up to affinity. This statement is a generalization of the one-dimensional versions proved by G. Letac (1992, “Lectures on Natural Exponential Functions and Their Variance Functions,” Instituto de Matemática pura e aplicada: Monografias de matemática, 50, Rio de Janeiro) in the case of two families, and by D. Pommeret (1999, C. R. Acad. Sci. Ser. I 328, 929–933) for more than two families.

© 2001 Elsevier Science (USA)


Key words and phrases: convolution; marginal law; natural exponential families; Poisson–Gaussian families; variance function.

INTRODUCTION

A number of papers have been devoted to the so-called natural exponential families (NEF) in the past few years. Some of these works concern the classification of NEF (e.g. Morris (1982), Letac (1989), Letac and Mora (1990), Mora (1990), Hassairi (1992), Kokonendji (1994), Casalis (1996)); other papers deal with the projection of NEF (Bar-Lev et al., 1994, Barndorff-Nielsen and Koudou, 1995). The work presented in this paper originates in the remark that some results taken from that literature about classification and projection of NEF could lead to a generalization of the following theorem proved by Letac (1992): if the convolution product of two NEF $F_1$, $F_2$ on $\mathbb{R}$ is a NEF, then $F_1$ and $F_2$ are both Gaussian families or both Poisson families, up to affinity. (Note that the converse is clearly
true, as mentioned in Cohen and Sackrowitz (1991), since the convolution of two Gaussian (resp. Poisson) distributions is a Gaussian (resp. Poisson) distribution.

Pommeret (1999) provided another proof of that result, extending the property to the convolution of more than two families on $\mathbb{R}$. The proof is based on a characterization established by Bryc (see Bryc, 1987; see also Laha and Lukacs, 1960) of Poisson and Gaussian distributions on $\mathbb{R}$ by conditional moments; it also uses results of Barndorff-Nielsen and Koudou (1995) on the projection of NEF.

Clearly, Pommeret’s result yields a characterization of Gaussian and Poisson distributions on $\mathbb{R}$, and the aim of this paper is to generalize it in the following direction: if $F_1, \ldots, F_n$ are NEF on $\mathbb{R}^d$ such that their convolution product is a NEF, then each of them has an affine variance function and, as a consequence (cf. Letac, 1989) is a Poisson–Gaussian family.

One could think of two possible manners of demonstrating the result. The first would be to construct a $d$-dimensional version of the proof of Pommeret (1999), which needs an extension of Bryc’s characterization to multidimensional Poisson–Gaussian families and some tedious technical calculus. A second and more elegant approach, that we present in this paper, is to make use of the characterization, by their variance function, of NEF on $\mathbb{R}^d$ whose marginal on $\mathbb{R}^k$, $k < d$, is a NEF. The latter characterization, established by Bar-Lev et al. (1994), is recalled in Section 1 which contains also a summary of the main notation about NEF and the definition of Poisson–Gaussian families on $\mathbb{R}^d$. Section 2 is devoted to the result of the present paper.

1. PRELIMINARIES

1.1. Natural Exponential Families

Let us first recall some notation about natural exponential families. Let $\mu$, be a positive radon measure on $\mathbb{R}^d$ and consider its Laplace transform

$$L_\mu(\theta) := \int_{\mathbb{R}^d} \exp\langle \theta, x \rangle \mu(dx),$$

where, for $\theta = (\theta_1, \ldots, \theta_d)$ and $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, $\langle \theta, x \rangle = \sum_{i=1}^d \theta_i x_i$.

Let $\Theta(\mu)$ denote the interior of the convex set $\{\theta \in \mathbb{R}^d; L_\mu(\theta) < \infty\}$, and consider $k_\mu := \log L_\mu$ defined on $\Theta(\mu)$.

The set of $\mu$’s that are not concentrated on some strict affine subspace of $\mathbb{R}^d$ and such that $\Theta(\mu)$ is not empty is denoted by $\mathcal{M}_d$. For $\mu$ in $\mathcal{M}_d$, one
defines the natural exponential family (NEF) generated by \( \mu \), as the set 
\[ F(\mu) = \{ P(\theta, \mu)(dx) := \exp(\langle \theta, x \rangle - k_\mu(\theta)) \mu(dx) ; \theta \in \Theta(\mu) \}. \]
The function \( k_\mu \) is strictly convex and real-analytic, and its gradient \( k'_\mu \) defines a diffeomorphism from \( \Theta(\mu) \) to an open subset \( M_F \) of \( \mathbb{R}^d \). Since 
\[ k'_\mu(\theta) = \int_{\mathbb{R}^d} xP(\theta, \mu)(dx), \]
\( M_F \) is also called the mean domain of \( F \). Let \( \psi_\mu \) be its inverse function, and for \( m \in M_F \), consider 
\[ P(m, F) := P(\psi_\mu(m), \mu). \]
For \( m \in M_F \), denote by \( V_F(m) \) the covariance matrix of \( P(m, F) \). The map \( m \mapsto V_F(m) \) is called the variance function of \( F \). Note that \( M_{\psi_\mu}(F) = f(M_F) \) and the variance function of the family \( f(F) \) is given by 
\[ V_{f(F)}(M) = AV_F(A^{-1}(M-B)) A', \]
where \( A' \) means the transpose of the matrix \( A \).
1.3. Poisson–Gaussian Families

It follows from the definitions that the variance function of a Gaussian NEF on $\mathbb{R}$ is constant; for a Poisson NEF, the variance function is of the form $V_F(m) = bm$, $b \in \mathbb{R}$. Letac (1989) characterized NEF’s on $\mathbb{R}^d$ whose variance function is affine. More precisely:

**Proposition 1.2.** Consider a NEF $F$ on $\mathbb{R}^d$. The following assertions are equivalent:

(i) There exist a linear map $B$ from $\mathbb{R}^d$ to the set $S_d$ of $d \times d$ real symmetric matrices and $C \in S_d$ such that

$$V_F(m) = B(m) + C$$

for all $m \in M_F$.

(ii) Up to an affine transformation on $\mathbb{R}^d$, there exists $k$ in $\{0, 1, \ldots, d\}$ such that, for all $m = (m_1, \ldots, m_d) \in M_F$,

$$V_F(m) = \text{diag}(m_1, \ldots, m_k, 1, 1, \ldots, 1)$$

(with a slight abuse of notation in the case $k = 0$, where $V_F(m)$ is the unit matrix).

In other words, a NEF on $\mathbb{R}^d$ has an affine variance function if, and only if, up to an affine transformation, it is the product of $k$ univariate Poisson families and $d-k$ Gaussian families, $k = 0, \ldots, d$. Such a family will be called a Poisson–Gaussian family.

1.4. Projection of NEF

If $F$ is a NEF on $\mathbb{R}^d$ ($d \geq 2$) and if we consider the family $p(F)$ of the margins on $\mathbb{R}^k$ ($1 \leq k < d$) of the elements of $F$, then $p(F)$ is not necessarily a NEF on $\mathbb{R}^k$. Bar-Lev et al. (1994), proved that if $p(F)$ is still a NEF, then the variance function of $F$ fulfills the condition mentioned in the following Proposition 1.3. Other characterizations of such NEF’s, using for instance Laplace transforms of conditional distributions, can be found in Barndorff-Nielsen and Koudou (1995).

**Proposition 1.3** (Bar-Lev et al., 1994). Let $F$ be a NEF on $\mathbb{R}^d$ ($d \geq 2$). For $k = 1, \ldots, d-1$, consider the projection $p(F)$ of $F$ under the map $\mathbb{R}^d \to \mathbb{R}^k$; $(x_1, \ldots, x_d) \mapsto (x_1, \ldots, x_k)$. $p(F)$ is a NEF on $\mathbb{R}^k$ if, and only if, for all $(m_1, \ldots, m_d) \in M_F$, the principal $k \times k$ submatrix of the matrix $V_F(m_1, \ldots, m_d)$ does not depend on $(m_{k+1}, \ldots, m_d)$. The latter submatrix is then the variance function of the family $p(F)$. 

2. NEF'S OBTAINED BY CONVOLUTION OF NEF'S

We now give the announced result on the convolution of NEF, extending Letac (1992) and Pommeret (1999).

**Theorem 2.1.** Let \( F_1, \ldots, F_n \) (\( n \geq 2 \)) be NEF's on \( \mathbb{R}^d \). Then the following statements are equivalent:

(i) The set
\[
F_1 * F_2 * \cdots * F_n := \{(\mu_1 * \cdots * \mu_n); (\mu_1, \ldots, \mu_n) \in F_1 \times F_n\}
\]
is a NEF (on \( \mathbb{R}^d \)).

(ii) \( F_i \) is a Poisson–Gaussian family for each \( i \in \{1, \ldots, n\} \) and there exist a linear map \( B \), not depending on \( i \), from \( \mathbb{R}^d \) to the set \( \mathcal{S}_d \) of real symmetric \( d \times d \) matrices, and \( c_1, \ldots, c_n \in \mathcal{S}_d \) such that
\[
V_{F_i}(m_i) = B(m_i) + c_i
\]
for all \( i = 1, \ldots, n \) and for all \( m_i \in M_{F_i} \).

(iii) \( F_1 * F_2 * \cdots * F_n \) is a Poisson–Gaussian family.

(iv) \( F_{i_1} * F_{i_2} * \cdots * F_{i_k} \) is a Poisson–Gaussian family for all \( i_1, \ldots, i_k \in \{1, \ldots, n\} \).

(v) \( F_{i_1} * F_{i_2} * \cdots * F_{i_k} \) is a NEF for all \( i_1, \ldots, i_k \in \{1, \ldots, n\} \).

**Proof.** Throughout the sequel we will make an abuse of notation: \( \partial k/\partial \theta_1 \) will stand for the gradient of the map \( \theta_1 \mapsto k(\theta_1, \theta_2) \) for fixed \( \theta_2 \), even if \( \theta_1 \) is a vector, and \( k' \) will denote the gradient of \( k \) although \( k \) is a function of several variables.

Consider \( (\mu_1, \ldots, \mu_n) \in (F_1 \times \cdots \times F_n) \). Let \( X_1, \ldots, X_n \) be independent random vectors in \( \mathbb{R}^d \) such that the distribution of \( X_i \) is \( \mu_i \), \( i = 1, \ldots, n \). Denote by \( \mu \) the law on \( (\mathbb{R}^d)^n \) of the vector \( (X_1 + \cdots + X_n, X_2, \ldots, X_n) \) and by \( F \) the NEF generated by \( \mu \). Let us compute the principal \( d \times d \) submatrix of \( V_F(m_1, \ldots, m_n) \).

Since \( X_1, X_2, \ldots, X_n \) are independent, \( F \) is the image of the family
\[
G := F_1 \otimes F_2 \otimes \cdots \otimes F_n
\]
by the linear transformation \( \phi; (\mathbb{R}^d)^n \rightarrow (\mathbb{R}^d)^n; (x_1, \ldots, x_n) \mapsto (x_1 + \cdots + x_n, x_2, \ldots, x_n) \). The matrix of \( \phi \) in the canonical basis of \( (\mathbb{R}^d)^n \) is
\[
A = [A_{ij}]_{1 \leq i, j \leq n},
\]
where \( A_{ij} \) is a \( d \times d \) matrix equal to the unit matrix if \( i = 1 \) or \( i = j \) or the null matrix otherwise. Thus, by formula (1), the principal \( d \times d \) submatrix of \( V_{F}(m_{1}, \ldots, m_{n}) \) is

\[
V_{11} = \sum_{i,j=1}^{n} A_{ij} [V_{G}(A^{-1}(m_1, \ldots, m_n)')]_{ij} A_{1j}.
\]

Now remark that \( A^{-1}(m_1, \ldots, m_n)' = (m_1 - m_2 - \cdots - m_n, m_2, \ldots, m_n)' \) and that, \( G \) being a product of independent families, we have: for all \( (x_1, \ldots, x_n) \in M_G \), \( [V_{G}(x_1, \ldots, x_n)]_{ij} \) is equal to \( V_{F}(x_i) \) if \( i = j \) and vanishes otherwise. Therefore,

\[
V_{11} = V_{F1}(m_1 - m_2 - \cdots - m_n) + \sum_{j=2}^{n} V_{Fj}(m_j).
\] (3)

(i) \( \Rightarrow \) (ii) The main argument of the proof is the fact that the image of \( F \) by the projection \( (\mathbb{R}^d)^n \rightarrow \mathbb{R}^d; (x_1, \ldots, x_n) \mapsto x_i \) is \( F_1 \ast F_2 \ast \cdots \ast F_n \) and, by assumption, is a NEF, so that the variance function of \( F \) has the property of Proposition (1.3). Thus, \( V_{11}(m_1, \ldots, m_n) \) does not depend on \( (m_2, \ldots, m_n) \), so that the differentiation of (2) with respect to \( m_i, \ 2 \leq i \leq n \) yields

\[
-V_{F1}'(m_1 - m_2 - \cdots - m_n) + V_{Fj}'(m_j) = 0.
\] (4)

By differentiating (4) with respect to \( m_i \), we obtain that, for all \( (m_1, \ldots, m_n) \in M_F \),

\[
-V_{F1}''(m_1 - m_2 - \cdots - m_n) = 0.
\]

Since \( m_1 - m_2 - \cdots - m_n \) describes \( M_{F1} \) as \( (m_1, \ldots, m_n) \) runs over \( M_F \), this implies that

\[
F_{F1}'' = 0.
\]

Therefore, \( V_{F1} \) is an affine function, and \( F_1 \) is a Poisson–Gaussian family by Proposition (1.2). Moreover, formula (4) implies that

\[
V_{F1}'(m_i) = V_{Fj}'(m_j)
\]

for all \( i, j \) so that the linear parts of the affine functions \( V_{F1}, \ldots, V_{Fn} \) are equal. This proves (ii).
(ii) ⇒ (iii) Using formula (2) and the linearity of $B$, we rewrite (3) to get

$$V_{11}(m_1, \ldots, m_n) = B(m_1) + \sum_{j=1}^{n} c_j$$  (5)

which proves that $V_{11}(m_1, \ldots, m_n)$ does not depend on $(m_2, \ldots, m_n)$. As a consequence, by the sufficient part of Proposition (1.3), $F_1 \ast F_2 \ast \cdots \ast F_n$ is a NEF. Furthermore, its variance function is $V_{11}$ by the same proposition. Since $V_{11}$ is affine by (5), it is the variance function of a Poisson–Gaussian family and (iii) is proved.

The sequel of the proof is straightforward: (iii) clearly implies (i) and thus (i), (ii) and (iii) are equivalent. Since this equivalence holds for any number of families, (iv) and (v) are equivalent to (i).

Concluding remarks.

- We have shown that Poisson–Gaussian families are the only ones obtained by the convolution of a finite number of independent NEF’s. Moreover, the families involved by the convolution product are the same families up to translations. Some statistical applications follow from this property. For instance, Cohen and Sackrowitz (1991) considered independent variables $X_1, \ldots, X_n$ with distributions in NEF’s such that the law of $\sum_{i \in I} X_i$ belongs to a NEF for all subsets $I \subset \{1, \ldots, n\}$.

- It is natural to ask whether the limit of $F_1 \ast \cdots \ast F_n$ is still a NEF when $n \to +\infty$. This question is solved with the following convergence result. Mora (1990) shows that if $(F(n))_{n \geq 1}$ is a sequence of NEFs on $\mathbb{R}^d$ such that

  (i) $\bigcap_{n \geq 1} M_{F(n)} = M_0$ is a non-empty open set,

  (ii) $\lim_{n \to +\infty} V_{F(n)}(m) = V(m)$ exists uniformly on compact subsets of $M_0$,

  (iii) $V(m)$ is positive definite on $M_0$,

then there exists a NEF with variance function $V(m)$ on $M_0$.

For example, if $(F_n)_{n \geq 1}$ are independent Poisson–Gaussian NEFs with variance functions $V_{F_n}(m_n) = B(m_n) + v_n$ and if the two series $m = \sum_{n \geq 1} m_n$ and $v = \sum_{n \geq 1} v_n$ converge on all compact subsets of $M_0$, then $F = \lim_{n \to +\infty} (F_1 \ast \cdots \ast F_n)$ is a Poisson–Gaussian family with variance function $V_F(m) = B(m) + v$. We give here a numerical example: Consider a Poisson Gaussian NEF, $F_1 = F(\mu)$, with variance function defined on $M_0$ by

$$V_{F_1}(m) = B(m) + C,$$
and, for all \( n \geq 1 \), define
\[
    \mu_n = \mu^{*1/\sigma^2},
\]
where \( * \) denotes the convolution, product. Writing \( F_n = F(\mu_n) \) it is straightforward that
\[
    M_{F_n} = M_{F_{1}}/n^2, \\
    V_{F_n}(m_n) = (1/n^2) V_{F_{1}}(n^2m_n) \\
    = B(m_n) + C/n^2.
\]
Then \( F = \lim_{n \to \infty} (F_1 * \cdots * F_n) \) is a Poisson–Gaussian family with variance function
\[
    V_F(m) = B(m) + v.
\]

REFERENCES


C. Kokonendji, Exponential families with variance-function in \( \mathbb{R} \) \( P(\sqrt{\lambda}) \): Seshadri’s class, \textit{Test} 2 (1994), 123–172.


