The Annals of Statistics
1989, Vol. 17, No. 2, 741–748

RATIONAL VARIANCE FUNCTIONS

BY SHAUL K. BAR-LEV AND DAOUD BSHOUTY[1]

University of Haifa and Technion–Israel Institute of Technology

Exponential dispersion models play an important role in the context of generalized linear models, where error distributions, other than the normal, are considered. Any statistical model expressible in terms of a variance-mean relation (V, Ω) leads to an exponential dispersion model provided that (V, Ω) is a variance function of a natural exponential family: Here Ω is the domain of means and V is the variance function of the natural exponential family. Therefore, it is of a particular interest to examine whether a pair (V, Ω) can serve as the variance function of a natural exponential family. In this study we consider the case where Ω is bounded and examine whether V can be the restriction to Ω of a rational function vanishing at the boundary points of Ω. The class of such functions is large and contains the important subclass of polynomials. It is shown that, apart from the binomial family (possessing a quadratic variance function) and affine transformations thereof, there exists no natural exponential family with variance function belonging to this class. Such a result implies, in particular, that the only variance functions of natural exponential families among polynomials of at least third degree are those restricted to unbounded domains Ω.

1. Introduction. Statistical models expressed in terms of variance-mean relations have been extensively discussed in the past 15 years. Scatter diagrams indicating an association between the mean and the variance for biological survival data appeared in numerous papers [see Jørgensen (1987) for references]. Wedderburn (1974) used the variance-mean relation, or the variance function (VF), to define quasilikelihoods. Quasilikelihoods have the advantage of imposing only second moment assumptions, whereas distributional assumptions are required to define likelihoods.

For a natural exponential family (NEF), the VF characterizes the family among the class of NEF's of the same order [Morris (1982) and Jørgensen (1987)]. Any NEF leads to an exponential dispersion model [Jørgensen (1987)], in which case the quasilikelihood reduces to a log-likelihood. Exponential dispersion models themselves have an important role in the context of generalized linear models, where error distributions, other than the normal, are considered [cf. McCullagh and Nelder (1983) and Jørgensen (1987)]. Accordingly, it is of a particular interest to delineate cases where a given variance-mean relation can serve as the VF of a NEF, and consequently, leads to an appropriate exponential dispersion model.

Before specifying the main problem of this study we recall the definitions of a NEF and its VF. Let ν be a positive, σ-finite measure on $\mathbb{R}$ which is not

Received March 1987; revised July 1988.

[1]Work supported by a Promotion of Research Grant from the Technion.

AMS 1980 *subject classifications.* Primary 62E10; secondary 62J02.

Key words and phrases. Variance function, rational variance function, exponential dispersion model, natural exponential family, entire function, meromorphic function.

concentrated on one point. The Laplace transform and effective domain of ν are given, respectively, by

$$T(\theta) = \int_{\mathbb{R}} \exp(\theta x)\, d\nu(x)$$

and

$$\{\theta \in \mathbb{R} : T(\theta) < \infty\}.$$

Let Θ denote the interior of the effective domain of ν and henceforth assume that Θ is nonempty. For $\theta \in \Theta$, define

$$dF_\theta(x) = (T(\theta))^{-1} \exp(\theta x)\, d\nu(x).$$

Then the family of probability distributions $\mathscr{F} = \{F_\theta : \theta \in \Theta\}$ is called a NEF generated by ν [or a linear exponential family of order 1, in the terminology of Barndorff-Nielsen (1978)]. The mean function of $\mathscr{F}$ is the mapping defined on Θ by

$$\mu(\theta) = \int_{\mathbb{R}} x\, dF_\theta(x).$$

Let Ω stand for $\mu(\Theta)$. Ω is called the mean domain of $\mathscr{F}$. It is known [Barndorff-Nielsen (1978), page 121)] that Ω is an open interval and that μ is a one-to-one, both ways continuously differentiable mapping between the two open intervals Θ and Ω. Denote by $\theta = \theta(\mu)$ its inverse function and let the function V on Ω be defined by

$$V(\mu) = \int_{\mathbb{R}} (x - \mu)^2\, dF_{\theta(\mu)}(x).$$

The pair (V, Ω) is called the VF of $\mathscr{F}$. As noted above, (V, Ω) characterizes $\mathscr{F}$. Moreover, V is a positive, real analytic function on Ω, i.e., it is the restriction to Ω of an analytic function on some domain D of the complex plane containing Ω. Thus, for examining the question whether a given analytic function f on some domain D can be used to construct a VF (V, Ω) of a NEF, one should consider, as possible forms for Ω, the largest open intervals of $\mathbb{R} \cap D$ on which f is positive [in this connection see Letac and Mora (1986), Theorem 2.3].

Such a question has been considered by several authors. Morris (1982) identified all VF's corresponding to NEF's among polynomials of at most second degree. Mora (1986) identified all VF's of NEF's among third degree polynomials. Tweedie (1984), Bar-Lev and Enis (1986) and Jørgensen (1987) treated the case where V is a power function of μ. For the case $\Omega \subset \mathbb{R}^+$, Jørgensen (1984) provided necessary and sufficient conditions for (V, Ω) to be a VF of a NEF. As noted in Bar-Lev (1987), however, Jørgensen's conditions are extremely difficult to check if Ω is bounded, even for a VF as simple as that of the binomial family.

In this study we consider the case where Ω is bounded and examine whether V can be the restriction to Ω of a rational function vanishing at the boundary points of Ω. The class $\mathscr{M}$ of such functions is large and contains as a subclass all polynomials restricted to bounded intervals Ω. The importance of polynomials VF's, in general, is linked with the fact that an empirical variance-mean relation of a data set in a form of a polynomial of some degree can always be established.

In Section 2 we show (Lemma 2.1) that the class of NEF's with VF's belonging to $\mathscr{M}$ coincides with the class of NEF's having bounded supports and rational VF's. We then establish in Theorem 2.1 a quite surprising result. It is shown that, apart from the binomial family (possessing a quadratic VF) and affine transformations thereof, there exist no NEF's with VF's belonging to $\mathscr{M}$. This result, in addition to providing a characterization of the binomial family, is of interest in its own right. It significantly reduces the work required to delineate VF's of NEF's among polynomials of a certain degree, and particularly implies that the only VF's among polynomials of at least third degree are those restricted to unbounded intervals.

Mora (1986), while making her classification of cubic VF's of NEF's, showed that certain cubic functions restricted to bounded intervals Ω are not VF's of NEF's [see also Letac and Mora (1987)]. Bar-Lev and Enis (1987), in their investigation of dual-conjugate families, also showed that certain cubic functions restricted to bounded intervals are not VF's of NEF's. The method used by the latter authors motivated the present study.

2. The main results. Let $\mathscr{F}$ be a NEF on $\mathbb{R}$. Without loss of generality, we assume throughout the sequel that $0 \in \Theta$ and $T(0) = 1$, so that ν $(= F_0)$ is a probability measure. Let $\psi(\theta) = \log T(\theta)$, $\theta \in \Theta$, and define $\mu_0 = \psi'(0)$. Let (V, Ω) be the VF of $\mathscr{F}$. Then θ and $T(\theta)$ can be expressed by

$$\theta = \int_{\mu_0}^{\mu} dt/V(t), \qquad T(\theta) = \exp\left\{\int_{\mu_0}^{\mu} t\,dt/V(t)\right\}. \tag{1}$$

We exploit (1) as well as the relation

$$\mu'(\theta) = V(\mu(\theta)) \tag{2}$$

in the proofs of the following lemmas.

LEMMA 2.1. *Let $\mathscr{F}$ be a NEF with rational VF (V, Ω). Denote by H the closed convex hull of the common support of $\mathscr{F}$. Then, the following two conditions are equivalent:*

(i) *$\mathscr{F}$ has a bounded support with* $\operatorname{int} H = (c, d)$.
(ii) $\Omega = (c, d)$ *and* $\lim_{m \to c+} V(m) = \lim_{m \to d-} V(m) = 0$.

PROOF. Assume that (i) holds. Then $\mathscr{F}$ is regular with $\theta(\Omega) = \mathbb{R}$ and hence $\mathscr{F}$ is steep with $\Omega = \operatorname{int} H$ [see Barndorff-Nielsen (1978), Theorem 9.2, page 142]. Consequently,

$$\lim_{\mu \to c^+} \int_{\mu}^{\mu_0} dt/V(t) = \lim_{\mu \to d^-} \int_{\mu_0}^{\mu} dt/V(t) = \infty, \tag{3}$$

and hence $\lim_{t \to c^+} V(t) = \lim_{t \to d^-} V(t) = 0$.

Assume next that (ii) holds. Since V is rational, $1/V(t)$ has poles (of at least first order) at c and d. Hence (3) holds and $\theta(\Omega) = \mathbb{R}$. Therefore $\mathscr{F}$ is steep with int $H = \Omega$. □

REMARK 2.1. Note that for the important case where V is a polynomial, boundedness of Ω implies that V vanishes at the boundary points of Ω.

REMARK 2.2. There are many examples of NEF's with rational VF's restricted to bounded intervals Ω that violate condition (ii) of Lemma 2.1. Indeed, a wider version of a result in Bar-Lev (1989), presented in Corollary 3.3 of Letac and Mora (1989), enables us to easily construct such examples. For instance, the latter corollary implies that for $\Omega = (0,1)$,

$$V_k(\mu) = \mu(1-\mu^k)^{-1} = \sum_{j=0}^{\infty} \mu^{jk+1}, \qquad k = 1,2,\ldots,$$

are VF's of infinitely divisible NEF's. Moreover, since infinitely divisible distributions have unbounded supports, it follows that the NEF's corresponding to V_k, $k = 1,2,\ldots,$ are nonsteep.

LEMMA 2.2. *Let (V,Ω) be the VF of a NEF $\mathscr{F}$ with V being the restriction to Ω of a rational function $f = p_m/q_n$ in $\mathbb{C}$, where p_m and q_n are relatively prime polynomials of degree m and n, respectively, and μ be the restriction to Θ of a meromorphic function w in $\mathbb{C}$. Then*:

(i) *$w(z)$ is the solution of the differential equation*

$$r'(z) = f(r(z)), \qquad z \in \mathbb{C}. \tag{4}$$

(ii) *$f \circ w$ does not vanish in $\mathbb{C}$.*

(iii) *f may vanish at most at two points.*

(iv) *If $f(w_1) = f(w_2) = 0$, then w does not attain the values w_1 and w_2 and $m - n = 2$.*

PROOF. (i) Since f is rational and w is meromorphic in $\mathbb{C}$ then $f \circ w$ is meromorphic in $\mathbb{C}$. The restriction of $f \circ w$ to Θ coincides with μ' which itself is the restriction of w' to Θ. Hence, by (2) and the uniqueness of the analytic continuation we have $w' = f \circ w$ in $\mathbb{C}$.

(ii) Assume to the contrary that $f(w(z_0)) = 0$ at some point $z_0 \in \mathbb{C}$. By (i), $f(w(z_0)) = w'(z_0) = 0$, so that z_0 is a regular point of w and $f^{(k)}(w(z_0))$, $k = 1,2,\ldots,$ are finite. Since w is a solution of (4), we obtain by differentiating (4),

$$w''(z_0) = f'(w(z_0))w'(z_0) = 0.$$

By successive differentiation of (4), we obtain that $w^{(k)}(z_0) = 0$, $k = 1,2,\ldots$. Hence $w \equiv c_1$, where c_1 is a constant, and $\Omega = \mu(\Theta) = \{c_1\}$, a contradiction.

(iii) If f has at least three zeroes, then, by the little Picard theorem [see Nevanlinna (1970), page 15], w would reach one of them, a contradiction with (ii).

(iv) Assume that $f(w_1) = f(w_2) = 0$ and that w attains one of these values, say w_1, at some point z_1. Then $f(w(z_1)) = 0$ and this contradicts (ii). Thus, w does not attain any of the values w_1 and w_2. Since $w \not\equiv$ constant, it follows by the little Picard theorem that w attains the value infinity at some point z_2. Thus, w has at z_2 a pole of some order j. Therefore, w' and $f \circ w$ have poles at z_2 of orders $j + 1$ and $(m - n)j$, respectively. By (4), $(m - n)j = j + 1$, or $1 + j^{-1} = m - n$, and thus $j = 1$ and $m - n = 2$. □

Before stating our main result we recall Rouché's theorem in the following setting. Let D be a simply connected domain whose boundary γ is a Jordan curve. Let f be analytic in $\overline{D}$. The winding number of $f(\gamma)$ around $\zeta_0 \in \mathbb{C}$ is defined by

$$\operatorname{Ind}(\zeta_0, f(\gamma)) = \frac{1}{2\pi i}\int_{f(\gamma)} \frac{d\zeta}{\zeta - \zeta_0} = \frac{1}{2\pi i}\int_{\gamma} \frac{f'(z)}{f(z) - \zeta_0}\, dz.$$

Geometrically, the winding number is the number of times that $f(\gamma)$ encompasses ζ_0 in the positive direction. By the argument principle it is also the number of times that $f(z)$ attains in D the value ζ_0 including multiplicities. A reformulation of Rouché's theorem [see Rudin (1987), Theorems 10.10 and 10.43] can be stated as follows.

Lemma 2.3 (Rouché's theorem). *Let D be a simply connected domain whose boundary γ is a Jordan curve. Let f and g be analytic in $\overline{D}$ and suppose that $|f| < |g|$ on γ. Then*

$$\operatorname{Ind}(0, g(\gamma)) = \operatorname{Ind}(0, (f + g)(\gamma)).$$

Theorem 2.1. *Let $\mathscr{F}$ be a NEF with VF (V, Ω), where $\Omega = (c, d)$ is finite. Let V be the restriction to Ω of a rational function $f = p_m/q_n$, where p_m and q_n are relatively prime polynomials of degree m and n, respectively, such that $f(c) = f(d) = 0$. Then there exists a positive integer j such that for $\mu \in (c, d)$,*

$$V(\mu) = j^{-1}(\mu - c)(d - \mu). \tag{5}$$

Proof. Since $f(c) = f(d) = 0$, we have $\lim_{t \to c^+} V(t) = \lim_{t \to d^-} V(t) = 0$. Hence, by Lemma 2.1, $\mathscr{F}$ is steep with $\Theta = \mathbb{R}$ and the members of $\mathscr{F}$ possess entire characteristic functions. Consequently, the Laplace transform $T(\theta)$ of F_0 admits an analytic continuation to the whole complex plane denoted by $L(z)$, $z = \theta + i\eta$, $\theta, \eta \in \mathbb{R}$. Since $L(z)$ is entire,

$$w(z) \equiv L'(z)/L(z) \tag{6}$$

is meromorphic in $\mathbb{C}$. Also, since

$$\mu(\theta) = T'(\theta)/T(\theta), \qquad \theta \in \mathbb{R},$$

it follows that μ is the restriction of w to Θ. By Lemma 2.2, we have

$$w'(z) = f(w(z)), \tag{7}$$

with $w(0) = \mu_0$. Hence, by the uniqueness of the solution of a first order differential equation, w satisfies

$$z = \int_{\mu_0}^{w} \frac{ds}{f(s)}, \tag{8}$$

and this relation must define w as a one-valued meromorphic function in $\mathbf{C}$ independently of the path of integration. Similarly, by using (6), we obtain that $L(z)$ satisfies

$$L(z) = \exp\left\{ \int_{\mu_0}^{w} \frac{s\,ds}{f(s)} \right\}, \tag{9}$$

and this relation must define $L(z)$ uniquely as an entire function in $\mathbf{C}$. By our assumption on f it follows from Lemma 2.2 that $\dot{m} = n + 2$ and f has exactly two distinct zeroes $w_1 = c$ and $w_2 = d$. Accordingly, by partial fractions techniques, we have

$$1/f(s) = \sum_{i=1}^{2} \sum_{j=1}^{n_i} \alpha_{ij}(s - w_i)^{-j} \equiv t(s) + \sum_{i=1}^{2} \alpha_{i1}(s - w_i)^{-1}, \tag{10}$$

where n_i, $i = 1, 2$, denotes the multiplicity of w_i as a root of $p_m(w) = 0$ and $\alpha_{ij} \in \mathbf{C}$ with $\alpha_{in_i} \neq 0$ for $i = 1, 2$. Therefore, we have from (8) and (10),

$$z = c_2 + \int t(w)\,dw + \sum_{i=1}^{2} \alpha_{i1} \log(w - w_i)^{-1}, \tag{11}$$

where c_2 is a constant. It was already noted that w attains the value infinity at some point z_2. Since $\int t(w)\,dw$ is finite at z_2, we conclude that the last term on the right-hand side of (11) must be also finite at z_2 and this implies that $\alpha_{11} = -\alpha_{21}$.

We now show that $\int t(w)\,dw$ is identically constant. For this consider the Möbius transform

$$l = (w - w_1)(w - w_2)^{-1}.$$

By Lemma 2.2, w does not attain the values w_1 and w_2, and therefore $l \circ w$ is a nonvanishing entire function. Inserting the relations

$$(w - w_1)^{-1} = (l - 1)[(w_2 - w_1)l]^{-1} \quad \text{and} \quad (w - w_2)^{-1} = (l - 1)(w_2 - w_1)^{-1}$$

in (11) yields

$$z = h_{n_2 - 1}(l) - \sum_{j=2}^{n_1} \frac{\alpha_{1j}}{(j-1)} \left[\frac{l - 1}{(w_2 - w_1)l} \right]^{j-1} - \alpha_{11} \log l, \tag{12}$$

where $h_{n_2-1}(l)$ is a polynomial in l of degree $n_2 - 1$. [The middle summation on the right-hand side of (12) is void if $n_1 = 1$.] Since $l \circ w$ is a nonvanishing entire function, $\log l \circ w$ may be defined as an entire function. The positively oriented circle $|l| = L$, denoted by γ_L, corresponds to a closed curve δ_z in the z-plane. Now, l is a Möbius transform of w, hence the preimage λ_w of γ_L is, for sufficiently large L, a small circle in the w-plane encountering w_2 only once. Note that the positively oriented curve δ_z corresponds to the negatively oriented curve λ_w. On γ_L, for sufficiently large L, the middle and last terms of the right-hand side of (12) are smaller in absolute value than $|h_{n_2-1}(l)|$. Therefore, by Lemma 2.3, we have

$$\operatorname{Ind}(0, \delta_z) = \operatorname{Ind}\big(0, h_{n_2-1}(\gamma_L)\big) = n_2 - 1.$$

Hence, δ_z encompasses the origin exactly $n_2 - 1$ times if $n_2 > 1$. Assume that $n_2 > 1$. Then for $j = 1, \dots, n_2 - 1$, we obtain by using (7) and the residue theorem that

$$\begin{aligned} 0 &= \frac{1}{2\pi i}\int_{\delta_z}[l(w(z))]^{-j}\,dz \\ &= \frac{1}{2\pi i}\int_{\lambda_w}\left(\frac{w - w_2}{w - w_1}\right)^j \frac{1}{f(w)}\,dw = -\alpha_{2j+1}(w_2 - w_1)^{-j}. \end{aligned}$$

This is a contradiction and therefore $n_2 = 1$. By interchanging the roles of w_1 and w_2, we obtain that $n_1 = 1$. Therefore $\int t(w)\,dw$ is identically constant and we conclude that

$$f(w) = a(w - c)(w - d), \qquad a \in \mathbb{R}. \tag{13}$$

By using (13) and (9), we get

$$L(z) = b(w - c)^{\alpha/a}(w - d)^{(1-\alpha)/a}, \qquad \alpha, b \in \mathbb{R}.$$

Let z_0 be a zero of $L(z)$ of some order j. Then z_0 is a first order pole of w. Therefore, near z_0,

$$L(z) = O\big((z - z_0)^{-\alpha/a}(z - z_0)^{-(1-\alpha)/a}\big) = O\big((z - z_0)^{-1/a}\big),$$

and this implies that $a = -j^{-1}$. This concludes the proof of the theorem. □

Note that the regular binomial family constitutes a NEF with VF of the form $(\mu(j - \mu)j^{-1}, \Omega = (0, j))$, for some $j \in \{1, 2, \dots\}$. By transforming $\mu \to \mu^* = (d - c)\mu j^{-1} + c$, we obtain (5), i.e., (5) is the VF of a NEF obtained from the regular binomial family by a suitable affine transformation.

As was outlined in detail by one of the referees, Theorem 2.1 can be extended to the case where V is the restriction to Ω of the quotient of two entire functions (rather than the quotient of two polynomials as in Theorem 2.1) vanishing at the boundary points of Ω. Interested readers can obtain an outline of this extension from the authors.

Acknowledgments. We are grateful to an Associate Editor, two referees and Peter Enis for helpful comments. We are particularly indebted to one of the referees for careful criticism and detailed suggestions which improved the quality of the paper considerably.

REFERENCES

BAR-LEV, S. K. (1987). Discussion of "Exponential dispersion models" by B. Jørgensen. *J. Roy. Statist. Soc. Ser. B* **49** 153–154.

BAR-LEV, S. K. and ENIS, P. (1986). Reproducibility and natural exponential families with power variance functions. *Ann. Statist.* **14** 1507–1522.

BAR-LEV, S. K. and ENIS, P. (1987). On dual-conjugate families. Technical Report 87-2, Dept. Statistics, SUNY at Buffalo.

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

JØRGENSEN, B. (1984). Exponential dispersion models. Reprint 9, Dept. Mathematics, Odense Univ.

JØRGENSEN, B. (1987). Exponential dispersion models (with discussion). *J. Roy. Statist. Soc. Ser. B* **49** 127–162.

LETAC, G. and MORA, M. (1986). On the variance functions of natural exponential families on $\mathbb{R}$. *C. R. Acad. Sci. Paris Sér. I. Math.* **302** 551–554.

LETAC, G. and MORA, M. (1987). Natural real exponential families with cubic variances. Unpublished manuscript.

LUKACS, E. (1970). *Characteristic Functions*, 2nd ed. Hafner, New York.

MCCULLAGH, P. and NELDER, J. A. (1983). *Generalized Linear Models.* Chapman and Hall, London.

MORA, M. (1986). A classification of cubic variance functions of exponential families on $\mathbb{R}$. *C. R. Acad. Sci. Paris Sér. I Math.* **302** 587–590.

MORRIS, C. N. (1982). Natural exponential families with quadratic variance functions. *Ann. Statist.* **10** 65–80.

NEVANLINNA, R. (1970). *Analytic Functions.* Springer, Berlin.

RUDIN, W. (1987). *Real and Complex Analysis*, 3rd ed. McGraw-Hill, New York.

TWEEDIE, M. C. K. (1984). An index which distinguishes between some important exponential families. In *Statistics: Applications and New Directions. Proceedings of the Indian Statistical Institute Golden Jubilee International Conference* (J. K. Ghosh and J. Roy, eds.) 579–604. Indian Statistical Institute, Calcutta.

WEDDERBURN, R. W. M. (1974). Quasilikelihood functions, generalized linear models and the Gauss–Newton method. *Biometrika* **61** 439–447.

DEPARTMENT OF STATISTICS
UNIVERSITY OF HAIFA
HAIFA 31999
ISRAEL

FACULTY OF MATHEMATICS
TECHNION–ISRAEL INSTITUTE OF TECHNOLOGY
HAIFA 32000
ISRAEL