OPTIMAL UNBIASED MEANS, POLYNOMIAL MOMENT RELATIONS, AND NATURAL EXPONENTIAL FAMILIES

L. Mattner

This is a preprint version last modified on October 27, 1995. The definitive version of this article appeared, under the same title, in *Statistics & Decisions* **14**, 343-352 (1996).

Abstract

Assume that a statistical model $\mathcal{P}$ on the real line possesses moments of every order and is such that, for some positive integer k, the moment of order $k+1$ is a polynomial function p of the k moments of lower order. Then the UMVU–property of the sample mean in a corresponding IID–model implies that $\mathcal{P}$ is a k–parametric model, parametrized through the first k moments and characterized by p. In each example treated here, $\mathcal{P}$ turns out to be essentially a k–parametric exponential family.

Our results extend and unify work of Kagan, Linnik & Rao and Bondesson on UMVU sample means in location or scale parameter families, Morris on natural exponential families with quadratic variance functions, and Plachky on the Poisson distribution.

1 Introduction and outline of results

Let $\mathcal{P} = \{P_\vartheta : \vartheta \in \Theta\}$ denote a family of probability distributions on the real line $\mathbf{R}$, with an arbitrary parameter set Θ, and consider observing independent copies $X_1, \ldots, X_n$ of a random variable X distributed acccording to P_ϑ with some unknown $\vartheta \in \Theta$. What can be said about $\mathcal{P}$ if the sample mean $\overline{X}_n = \frac{1}{n}\sum_{\nu=1}^{n} X_\nu$ is a uniformly minimum variance unbiased (UMVU) estimator for $E_\vartheta[X]$?

This question has been resolved quite satisfactorily in the sixties and seventies under the additional assumption that $\mathcal{P}$ is a location, scale, or location–scale model. See [Kagan, Linnik & Rao (1973), Chapter 7] and Bondesson (1974), or the more specific references given below.

[0] *1991Mathematics Subject Classification:* 62F10, 60E10.

Keywords and phrases: Characteristic function, characterizations, location parameter, scale parameter, shift model, UMVUE

Here we assume instead that $\mathcal{P}$ fulfills a polynomial moment relation (see (1) below). Under the further assumption of the existence of all moments, our basic Theorem 1 states in part a), among other things, that the UMVU property of $\overline{X}_n$ implies that $\mathcal{P}$ is a parametric model, characterized through the polynomial function p occuring in (1). In each example treated here, $\mathcal{P}$ is essentially an exponential family. This is anticipated in part b) of Theorem 1.

For the particular moment relations expressing the variance as an at most quadratic polynomial in the expectation, we obtain in Theorem 2 a generalization of the corresponding classification of the natural exponential families due to Morris (1982). Except for our assumption on the finiteness of all moments, the results on location and scale parameter families mentioned above turn out to be special cases of Theorem 2. Another special case, concerning the Poisson distribution and due to Plachky (1993), motivated the present study.

Finally, we give in Theorem 3, as an especially straightforward application of Theorem 1, a characterization of the two–parameter Gaussian family implicitly contained in Bondesson (1973 b).

2 Results

We first state our basic result. Some readers may wish to turn immediately to two most typical applications of Theorem 1, namely the proof of Theorem 3 and the proof of Theorem 2, case 3. We write $\mathbf{N} = \{1, 2, \ldots\}$ and $\mathbf{N}_0 = \{0, 1, 2, \ldots\}$.

Theorem 1 *Let $k \in \mathbf{N}$ and $p : \mathbf{R}^k \to \mathbf{R}$ be a polynomial of degree at most $d \in \mathbf{N}_0$,*

$$p(x_1, \ldots, x_k) = \sum_{\alpha \in \mathbf{N}_0^k, |\alpha| \le d} c_\alpha \cdot x_1^{\alpha_1} \cdot \ldots \cdot x_k^{\alpha_k}.$$

Let $\mathcal{P}$ be such that, for every $\vartheta \in \Theta$, X possesses moments of every order and

$$E_\vartheta[X^{k+1}] = p(E_\vartheta[X], \ldots, E_\vartheta[X^k]) \tag{1}$$

holds. If $\overline{X}_n$ is UMVU for some $n \ge \max(1, d)$, then:

a) There is a domain $G \subset \mathbf{C}^{k+1}$ with $G \supset \{0\} \times \mathbf{C}^k$ and an analytic function

$$\Phi \,:\, G \to \mathbf{C}$$

such that, for every $\vartheta \in \Theta$, the characteristic function φ_ϑ of X satisfies

$$\varphi_\vartheta(t) = \Phi(t; E_\vartheta[X], \ldots, E_\vartheta[X^k])$$

in a neighbourhood of zero. The function $z \mapsto \Phi(z; a_1, \ldots, a_k)$ is for every $(a_1, \ldots a_k) \in \mathbf{C}^k$ the unique solution of the initial value problem

$$\left.\begin{array}{c} \varphi^{(k+1)} \cdot \varphi^{d-1} = i^{k+1} \sum c_\alpha \left(\frac{1}{i}\varphi'\right)^{\alpha_1} \cdot \ldots \cdot \left(\frac{1}{i^k}\varphi^{(k)}\right)^{\alpha_k} \cdot \varphi^{d-|\alpha|}, \\ \varphi(0) = 1, \quad \varphi'(0) = ia_1, \quad \ldots, \quad \varphi^{(k)}(0) = i^k a_k \end{array}\right\} \tag{2}$$

in a neighbourhood of zero. Every φ_ϑ is analytic in a strip containing the real axis.

b) If a complete exponential family $\mathcal{Q}$ in $x, T_2(x), \ldots, T_k(x)$ exists, such that the corresponding natural parameters are all in the interior of the natural parameter space, and such that (1) *holds when $\mathcal{Q}$ replaces $\mathcal{P}$, then $P_\vartheta \in \mathcal{Q}$ for every ϑ satisfying*

$$\Big(E_\vartheta[X], \ldots, E_\vartheta[X^k]\Big) \in \left\{ \Big(\int x \, dQ(x), \ldots, \int x^k \, dQ(x) \Big) : Q \in \mathcal{Q} \right\}. \tag{3}$$

We now give a complete discussion of the special case $k = 1$ and p quadratic, but choose to express the moment relation in terms of the variance.

Theorem 2 *Let q be a polynomial not identical zero and of degree at most 2. Assume that $\mathcal{P}$ is such that, for every $\vartheta \in \Theta$, X possesses moments of every order and*

$$Var_\vartheta(X) = q(E_\vartheta[X]). \tag{4}$$

If the sample mean is UMVU for some $n \geq 2$, then $\mathcal{P}$ is for some $r > 0$ a subset of an affine transformation of one of the following 6 families:

1. *The Gaussian location family*

$$\mathcal{P}_1 = \{N(a, 1) \, : \, a \in \mathbf{R}\}$$

   with $q(x) = 1$;

2. *the Poisson family*

$$\mathcal{P}_2 = \{\text{Poisson}(\lambda) \, : \, \lambda \geq 0\}$$

   with $q(x) = x$;

3. *a binomial or a Dirac family*

$$\mathcal{P}_3(r) = \begin{cases} \{\text{Bin}(r, p) \, : \, 0 \leq p \leq 1\} & \text{if } r \in \mathbf{N}, \\ \{\delta_0, \delta_r\} & \text{if } r \in (0, \infty) \setminus \mathbf{N}, \end{cases}$$

   with $q(x) = x \cdot (1 - \frac{x}{r})$;

4. *a negative binomial family*

$$\mathcal{P}_4(r) = \mathcal{P}_4^+(r) \cup \mathcal{P}_4^-(r)$$

   with $\mathcal{P}_4^+(r) = \{\text{Negbin}(r, p) \, : \, 0 < p \leq 1\}$, where $\text{Negbin}(r, p)$ is the distribution on $\mathbf{N}_0$ with the density $x \mapsto \binom{x+r-1}{x} p^r (1-p)^x$ with respect to counting measure, and $\mathcal{P}_4^-(r)$ is the image of $\mathcal{P}_4^+(r)$ under reflection at the point $-\frac{r}{2}$, here $q(x) = x \cdot (1 + \frac{x}{r})$;

5. *the Gamma family*

$$\mathcal{P}_5(r) = \{\Gamma(r, \alpha) \, : \, \alpha \in \mathbf{R}\}$$

   where $\Gamma(r, 1)$ has the Lebesgue density $x \mapsto \frac{1}{\Gamma(r)} x^{r-1} e^{-x} 1_{(0,\infty)}(x)$ and $\Gamma(r, \alpha)$ is the distribution of $\alpha \cdot Y$ with Y distributed according to $\Gamma(r, 1)$, with $q(x) = \frac{x^2}{r}$;

6. *the generalized hyberbolic secant family*

$$\mathcal{P}_6(r) = \{\mathrm{GHS}(r,\lambda) : \lambda \in \mathbf{R}\},$$

where $\mathrm{GHS}(r,\lambda)$ *is the probability distribution on* $\mathbf{R}$ *with characteristic function*

$$t \mapsto (\cosh t - i\lambda \sinh t)^{-r},$$

with $q(x) = r + \frac{x^2}{r}$.

Note that each of the above $\mathcal{P}_\nu$ or $\mathcal{P}_\nu(r)$, $\nu = 1, \ldots, 6$, can be written as a union of at most 4 homogeneous subfamilies. (Recall that a family of probability measures is called homogeneous if any two of its members are mutually absolutely continuous.) At most two of these are natural exponential families on $\mathbf{R}$ (i. e., exponential families in the identical statistic; as a precise definition we prefer the one given in [Letac & Mora (1990), p. 2] or [Letac (1992), p. 19], although the slight difference from the one given in Morris (1982) is irrelevant for the present paper), and the remaining ones are singletons containing a Dirac measure. Hence the following corollary to Theorem 2 is rather obvious.

Corollary *Let* $\mathcal{P}$ *be such that* (4) *holds for some polynomial* q, *of degree at most two. Suppose that* $\{E_\vartheta[X] : \vartheta \in \Theta\}$ *contains a nondegenerate interval. Then the following three statements are equivalent.*

i) $\mathcal{P}$ *is a subset of a natural exponential family.*

ii) $\mathcal{P}$ *is homogeneous, possesses moments of every order, and* $\overline{X}_n$ *is UMVU for some* $n \geq 2$.

iii) $\mathcal{P}$ *is a subfamily of an affine transformation of a family belonging to one of the six types listed in* Morris (1982).

The implication "*i*) $\Rightarrow$ *iii*)" was given in Morris (1982) and verified with a more rigorous proof in Letac & Mora (1990). The reader is invited to compare the latter proof with the proof of our Theorem 2.

We now consider special cases of Theorem 2. Suppose first that a location family $\mathcal{P}$ on $\mathbf{R}$ with finite nonzero variance is given and that $\overline{X}_n$ is UMVU for some $n \geq 2$. Then $\mathcal{P}$ is a Gaussian location family. This follows from [Bondesson (1975), Theorem 9.1] or, simpler but for $n \geq 3$ only, from Kagan, Linnik & Rao (1965) and [Bondesson (1975), Theorem 8.1]. Since the variance is constant in any location family, we see that Theorem 2, case 1 yields a more general result in the case that all moments are assumed to exist.

Now suppose that a scale parameter family $\mathcal{P}$ on $[0,\infty)$ with finite nonzero variances is given and that $\overline{X}_n$ is UMVU for some $n \geq 2$. Then $\mathcal{P}$ is a Gamma location family, see Bondesson (1973 a). Since $\mathrm{Var}_\vartheta(X) = \frac{1}{r}E_\vartheta[X]^2$ for some $r > 0$ must hold under each $P_\vartheta \in \mathcal{P}$, we see that Theorem 2, case 5 again yields a more general result in the case that all moments are finite.

As said before, the special case $q(x)=x$ yielding the Poisson distribution is due to Plachky (1993), with a different proof. Note that this case is peculiar in that $\mathcal{P}_2$ is the only one of the above families not invariant under any reflection.

We finally give an especially straightforward application of Theorem 1 in the case $k=2$. The result itself is implicitly given in [Bondesson (1973 b), p. 339].

Theorem 3 *Let $\mathcal{P}$ be such that, for every $\vartheta \in \Theta$, X possesses moments of every order and*

$$E_\vartheta[(X - E_\vartheta[X])^3] = 0.$$

If the sample mean is for some $n \geq 3$ UMVU, then $\mathcal{P}$ contains only Gaussian distributions (possibly degenerate).

3 Proofs

Proof of Theorem 1. a) The statistic

$$h(X_1,\ldots,X_n) := X_1^{k+1} - \sum_\alpha c_\alpha \prod_{\nu=1}^{\alpha_1} X_\nu \cdot \prod_{\nu=\alpha_1+1}^{\alpha_1+\alpha_2} X_\nu^2 \cdot \ldots \cdot \prod_{\nu=\alpha_1+\ldots+\alpha_{k-1}+1}^{\alpha_1+\ldots+\alpha_k} X_\nu^k \tag{5}$$

is, by (1), an unbiased estimator of zero with finite variance for the model under discussion. Since $\sum_{\nu=1}^n X_\nu$ is UMVU, a standard induction argument yields

$$E_\vartheta\left[\left(\sum_{\nu=1}^n X_\nu\right)^l h(X_1,\ldots,X_n)\right] = 0 \qquad (\vartheta\in\Theta,\, l\in\mathbf{N}_0) \tag{6}$$

(use the well–known UMVU criterion (see e.g. [Lehmann(1983), p. 77]) and the assumption that X_1 possesses moments of every order).

Fix $\vartheta\in\Theta$, write φ instead of φ_ϑ, and put

$$\Omega(t) := E_\vartheta\left[e^{it\sum_{\nu=1}^n X_\nu} h(X_1,\ldots,X_n)\right] \qquad (t\in\mathbf{R}).$$

Then Ω is infinitely often differentiable and (6) yields

$$\Omega^{(l)}(0) = 0 \qquad (l\in\mathbf{N}_0). \tag{7}$$

On the other hand, Ω is a differential polynomial in φ, namely

$$\Omega = \frac{1}{i^{k+1}}\varphi^{(k+1)}\cdot\varphi^{n-1} - \sum_\alpha c_\alpha \left(\frac{1}{i}\varphi'\right)^{\alpha_1}\cdot\left(\frac{1}{i^2}\varphi''\right)^{\alpha_2}\cdot\ldots\cdot\left(\frac{1}{i^k}\varphi^{(k)}\right)^{\alpha_k}\cdot\varphi^{n-|\alpha|}.$$

Hence φ is, in a neighbourhood of zero, a solution of the initial value problem

$$\left.\begin{array}{c} \varphi^{(k+1)} = \frac{i^{k+1}}{\varphi^{n-1}}\left(\Omega + \sum c_\alpha \left(\frac{1}{i}\varphi'\right)^{\alpha_1}\cdot\ldots\cdot\left(\frac{1}{i^k}\varphi^{(k)}\right)^{\alpha_k}\cdot\varphi^{n-|\alpha|}\right), \\ \varphi(0)=1, \quad \varphi'(0)=ia_1, \quad \ldots, \quad \varphi^{(k)}(0)=i^k a_k \end{array}\right\} \tag{8}$$

for $a_\nu = E_\vartheta[X^\nu]$, $\nu=1,\ldots k$.

Consider now the initial value problem

$$\text{Same as (8), except that } \Omega \equiv 0. \tag{9}$$

This has, in a neighbourhood of zero, a unique solution ψ, which is in fact analytic.

Now (7), (8) and (9) easily yield

$$\varphi^{(l)}(0) = \psi^{(l)}(0) \qquad (l \in \mathbf{N}_0).$$

This implies $\varphi = \psi$, since ψ is analytic and φ is a characteristic function (see e.g. [Billingsley(1986), p. 406, Proof of Theorem 30.1]). In particular, φ is analytic and solves (9) in a neighbourhood of zero. The analyticity of φ in a strip containing the the real axis follows from [Linnik & Ostrovskiĭ(1977), p. 25, Theorem 2.2.3].

The statement concerning Φ just expresses the analytic dependence of the solution of (9) on the initial conditions.

b) We may apply part a) to $\mathcal{Q}$, using standard facts concerning exponential families ("interior" is used here for deducing the existence of all moments of X). If $P_\vartheta \in \mathcal{P}$ is such that (3) holds, then φ_ϑ and the characteristic function ψ of some $Q \in \mathcal{Q}$ both satisfy (9) with $a_\nu = E_\vartheta[X^\nu]$, $\nu = 1, \ldots, k$, in a neighbourhood of zero. Hence $\varphi_\vartheta = \psi$ and thus $P_\vartheta \in \mathcal{Q}$.

Proof of Theorem 2. Note that (4) can be written in the form (1) with $k = 1$ and $d = 2$. Hence Theorem 1 is applicable. Note also that all our assumptions on $\mathcal{P}$ are unaffected under affine transformations. In particular, if $\tilde{\mathcal{P}}$ is the model obtained from $\mathcal{P}$ via the transformation

$$\tilde{X} = aX + b$$

with $a \neq 0$, then $\tilde{\mathcal{P}}$ satisfies (4) with the polynomial q replaced by

$$\tilde{q}(x) = a^2 \cdot q(\frac{x-b}{a}). \tag{10}$$

This proof splits naturally into 6 cases, according to the coefficients and the zeros of the polynomial q. Let us write

$$q(x) = v_2 x^2 + v_1 x + v_0.$$

CASE 1: $v_2 = v_1 = 0$. Then $v_0 > 0$. By considering the transformation $\tilde{X} = \frac{1}{\sqrt{v_0}} X$ and (10), we may without loss of generality assume that $q(x) \equiv 1$. Put

$$\mathcal{Q} := \{N(a, 1) : a \in \mathbf{R}\}$$

and apply Theorem 1 b) to get $\mathcal{P} \subset \mathcal{Q} = \mathcal{P}_1$.

CASE 2: $v_2 = 0$, $v_1 \neq 0$. By the transformation $\tilde{X} = \frac{1}{v_1} X + \frac{v_0}{v_1^2}$ we may assume that $q(x) = x$. We apply Theorem 1 b) with

$$\mathcal{Q} = \{\text{Poisson}(\lambda) : \lambda > 0\}$$

to get $P_\vartheta \in \mathcal{Q}$ if $E_\vartheta[X] > 0$. But $E_\vartheta[X] < 0$ is impossible for the present q and $E_\vartheta[X] = 0$ implies that P_ϑ is the Dirac measure concentrated at zero, i. e., $P_\vartheta = \text{Poisson}(0)$. Hence $\mathcal{P} \subset \mathcal{P}_2$.

CASE 3: $v_2 < 0$ and q has two distinct real zeros. Then q can be written as

$$q(x) = \frac{1}{r} \cdot (\beta - x) \cdot (x - \alpha)$$

with α, $\beta \in \mathbf{R}$, $\alpha \neq \beta$, and $r > 0$. By the transformation $\tilde{X} = r\frac{X-\alpha}{\beta-\alpha}$ we may assume that

$$q(x) = x \cdot (1 - \frac{x}{r}). \tag{11}$$

Hence (4) is equivalent to

$$E_\vartheta[X^2] = p(E_\vartheta[X])$$

where

$$p(x) = q(x) + x^2 = x + \frac{r-1}{r}x^2.$$

Fix $\vartheta \in \Theta$ and put $a = E_\vartheta[X]$. By (11) and nonnegativity of variances, $0 \leq a \leq r$. By Theorem 1 a), the characteristic function φ_ϑ of X is in a neighbourhood of zero the solution of (2), i. e.

$$\left.\begin{array}{c} \varphi'' \cdot \varphi = i^2 \cdot \left(\frac{1}{i}\varphi' \cdot \varphi + \frac{r-1}{r}(\frac{1}{i}\varphi')^2\right), \\ \varphi(0) = 1, \quad \varphi'(0) = ia \end{array}\right\}. \tag{12}$$

Hence, easily checked,

$$\varphi_\vartheta(t) = \left(1 - \frac{a}{r} + \frac{a}{r}e^{it}\right)^r \tag{13}$$

in a neighbourhood of zero.

If $r \in \mathbf{N}$, then φ_ϑ is the characteristic function of the binomial distribution $\text{Bin}(r, \frac{a}{r})$. Since ϑ was arbitrary, $\mathcal{P} \subset \mathcal{P}(r)$ in this case.

If $r \notin \mathbf{N}$ and $0 < \frac{a}{r} < 1$, then it is well known that the right hand side in (13) is not a characteristic function. (If it were, then by analytic continuability of φ_ϑ along the whole imaginary axis and by [Linnik & Ostrovskiĭ(1977), p. 25, Theorem 2.2.3], φ_ϑ would extend to an entire function, which is impossible since $z \mapsto 1 - \frac{a}{r} + \frac{a}{r}e^{iz}$ has simple zeros in the complex plane.) Hence $a \in \{0, r\}$ and, by (11), $\mathcal{P} \subset \{\delta_0, \delta_r\}$.

CASE 4: $v_2 > 0$ and q has two distinct real zeros. Then, similar to case 3, we may assume that q can be written as

$$q(x) = x \cdot (1 + \frac{x}{r}) \tag{14}$$

for some $r > 0$. Put $\mathcal{Q}_+ := \mathcal{P}_4^+(r) \setminus \{\delta_0\}$ and $\mathcal{Q}_- := \mathcal{P}_4^-(r) \setminus \{\delta_{-r}\}$. Observe that the mean and the variance of $\text{Negbin}(r, p)$ are $r\frac{1-p}{p}$ and $r\frac{1-p}{p^2}$. Hence an application of Theorem 1 b) to the complete exponential family $\mathcal{Q}_+$ yields $P_\vartheta \in \mathcal{P}_4(r)$ for $E_\vartheta[X] > 0$.

Since the mean and the variance of the distribution obtained from $\text{Negbin}(r, p)$ under reflection at $-\frac{r}{2}$ are $-\frac{r}{p}$ and $r\frac{1-p}{p^2}$, an analogous application of Theorem 1 b) to $\mathcal{Q}_-$ yields $P_\vartheta \in \mathcal{P}_4(r)$ for $E_\vartheta[X] < -r$.

For $E_\vartheta[X] \in \{0, -r\}$ we have $P_\vartheta \in \{\delta_0, \delta_{-r}\} \subset \mathcal{P}_4(r)$, from (14).

In the case of $E_\vartheta[X] \in (-r, 0)$, (14) would yield a negative variance.

CASE 5: q has a double real zero. By a translation we may assume that p can be written as

$$q(x) = \frac{x^2}{r} \tag{15}$$

for some $r > 0$. Put $\mathcal{Q}_+ := \{\Gamma(r, \alpha) : \alpha > 0\}$ and $\mathcal{Q}_- := \{\Gamma(r, \alpha) : \alpha < 0\}$. Observe that the mean and variance of $\Gamma(r, \alpha)$ are $r\alpha$ and $r\alpha^2$, and apply Theorem 1 b) to the complete exponential family $\mathcal{Q}_+$. It follows that $P_\vartheta \in \mathcal{P}_5(r)$ if $E_\vartheta[X] > 0$. The same argument with $\mathcal{Q}_-$ in place of $\mathcal{Q}_+$ yields $P_\vartheta \in \mathcal{P}_5(r)$ if $E_\vartheta[X] < 0$. Finally, if $E_\vartheta[X] = 0$, then (15) yields $\mathrm{Var}_\vartheta(X) = 0$, hence $P_\vartheta = \delta_0 = \Gamma(r, 0) \in \mathcal{P}_5(r)$.

CASE 6: $v_2 = 0$ and q has no real zeros. Then q can be written as

$$q(x) = \frac{r}{a^2} + \frac{(x-c)^2}{r}$$

for some $c \in \mathbf{R}$ and $r, a > 0$. By the transformation $\tilde{X} = aX - ac$ we may assume that

$$q(x) = r + \frac{x^2}{r}. \tag{16}$$

Since $\mathcal{Q} := \mathcal{P}_6$ is a complete exponential family in the identity with

$$\left\{\int x\, dQ(x) : Q \in \mathcal{Q}\right\} = \mathbf{R}$$

(see Morris (1982)), Theorem 1 b) implies $\mathcal{P} \subset \mathcal{P}_6$.

Proof of Corollary. $i) \Rightarrow ii)$: Obvious from the well-known general properties of one–parameter exponential families.

$ii) \Rightarrow iii)$: Obvious from inspection of the maximal homogeneous subfamilies of the $\mathcal{P}_\nu$ or $\mathcal{P}_\nu(r)$ given in Theorem 2.

$iii) \Rightarrow i)$: Trivial.

Proof of Theorem 3. We may apply Theorem 1 with $k = 2$, $d = 3$ and

$$p(x_1, x_2) = 3x_1x_2 - 2x_1^3.$$

Put $\mathcal{Q} := \{N(a, \sigma^2) : a \in \mathbf{R}, \sigma^2 > 0\}$. Then $\mathcal{Q}$ is a complete exponential family in x and x^2. Theorem 1 b) yields $P_\vartheta \in \mathcal{Q}$ if $\mathrm{Var}_\vartheta(X) > 0$. But those P_ϑ with $\mathrm{Var}_\vartheta(X_1) = 0$ are degenerate gaussian.

4 Concluding remarks.

Let us first observe that part of the conclusion of Theorem 1 a) may be stated as follows:

$$E_\vartheta[h(X_1,\ldots,X_n)|\overline{X_n}] = 0 \qquad (\vartheta \in \Theta), \tag{17}$$

with h as in (5). In fact, (17) is known to be equivalent to the differential equation in (2), with φ_ϑ in place of φ. See, e. g., Laha & Lukacs (1960).

In the particular situation of Theorem 2, we have

$$h(X_1,\ldots,X_n) = X_1^2 - aX_1X_2 - bX_1 - c$$

for some constants a, b, c, so that we could have used [Laha & Lukacs (1960), Theorems 1–5] for proving Theorem 2. However, as the reader may check, performing the necessary identifications in detail would not yield to a very short proof of Theorem 2. This is why we have prefered the slightly longer but self–contained proof given above.

As seen in (the proof of) the Corollary, the Morris (1982) classification "i) $\Rightarrow$ iii)" follows easily from our Theorem 2. The converse is not true: Each of the families $\mathcal{P}_2$, $\mathcal{P}_3(r)$, $\mathcal{P}_4(r)$ and $\mathcal{P}_5(r)$ fulfils the assumption of Theorem 2, but none of them is an exponential family.

Notwithstanding the preceding paragraph, each of the families in Theorem 2 is a union of mutually singular exponential families or Dirac families.

Problem *Explain this fact.*

Note in this connection that the UMVU–property of $\overline{X_n}$ alone is far from sufficient to yield exponentiality, even for smooth, homogeneous models $\{P_\vartheta : \vartheta \in \Theta\}$. See Mattner (199?).

Acknowledgement. I thank D. Plachky for providing me with a copy of Plachky (1993) prior to publication, L. Bondesson for sending me a copy of Bondesson(1974), and G. Letac and the referees for their helpful comments on earlier versions of this paper.

References

[1] Billingsley, P. (1986). *Probability and Measure, 2nd Edition.* Wiley, New York.

[2] Bondesson, L. (1973 a). Characterizations of the Gamma distributions. *Theory Probab. Appl.* **18**, 367–369.

[3] Bondesson, L. (1973 b). Characterizations of the normal and the Gamma distributions. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **26**, 335-344.

[4] Bondesson, L. (1974). Characterizations of location and scale parameter families of distributions through optimality properties of certain estimators. Dissertation, Lund, 15 pp.

[5] BONDESSON, L. (1975). Uniformly minimum variance estimation in location parameter families. *Ann. Statist.* **3** 637–660.

[6] KAGAN, A.M., LINNIK, JU.V. & RAO, C.R. (1965). On a characterization of the normal law based on a property of the sample average. *Sankhyā Ser. A.* **27**, 405–406.

[7] KAGAN, A.M., LINNIK, JU.V. & RAO, C.R. (1973). *Characterization Problems in Mathematical Statistics.* Wiley, New York.

[8] LAHA, R.G. & LUKACS, E. (1960). On a problem connected with quadratic regression. *Biometrika* **47**, 335–343.

[9] LEHMANN, E.L. (1983). *Theory of Point Estimation.* Wiley, New York.

[10] LETAC, G. (1992). *Lectures on Natural Exponential Families and Their Variance Functions.* Monografias de Matemática No. 50. Instituto de Matemática Pura e Aplicada, Rio de Janeiro.

[11] LETAC, G. & MORA, M. (1990). Natural real exponential families with cubic variance functions. *Ann. Statist.* **18** 1–37.

[12] LINNIK, JU.V. & OSTROVSKIĬ, I.V. (1977). *Decomposition of Random Variables and Vectors.* Transl. Math. Monographs Vol. 48, Amer. Math. Soc., Providence, R.I.

[13] MATTNER, L. (199?). Complete order statistics in parametric models. *Ann. Statist.*

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

[15] PLACHKY, D. (1993). An estimation–theoretical characterization of the Poisson distribution. *Statistics and Decisions,* Supplement Issue No. **3**, 175–178.

Lutz Mattner
Universität Hamburg
Institut für Mathematische Stochastik
Bundesstr. 55
D–20146 Hamburg
Email: mattner@math.uni–hamburg.de