Probab. Th. Rel. Fields 76, 509–522 (1987)

Probability Theory and Related Fields

Characterizations of Natural Exponential Families with Power Variance Functions by Zero Regression Properties

Shaul K. Bar-Lev* and Osnat Stramer

Department of Statistics, University of Haifa, Haifa 31999, Israel

Summary. Series of new characterizations by zero regression properties are derived for the distributions in the class of natural exponential families with power variance functions. Such a class of distributions has been introduced in Bar-Lev and Enis (1986) in the context of an investigation of reproducible exponential families. This class is broad and includes the following families: normal, Poisson-type, gamma, all families generated by stable distributions with characteristic exponent an element of the unit interval (among these are the inverse Gaussian, Modified Bessel-type, and Whittaker-type distributions), and families of compound Poisson distributions generated by gamma variates. The characterizations by zero regression properties are obtained in a unified approach and are based on certain relations which hold among the cumulants of the distributions in this class. Some remarks are made indicating how the techniques used here can be extended to obtain characterizations of general exponential families.

1. Introduction

Let $\underline{X}=(X_1,\dots,X_n)$, where $X_1,\dots,X_n$ are i.i.d. r.v.'s with common distribution F, and $T_i=T_i(\underline{X})$, $i=1,2$, be two statistics. Characterizations of F by means of zero regression of T_1 on T_2 have been thoroughly discussed in Lukacs and Laha (1964) and Kagan, Linnik, and Rao (1973). (The latter references are hereafter referred to as LL and KLR, respectively.) Other recent works on this topic are those of Gordon (1973), Heller (1983), and Seshadri (1983). The majority of such characterizations are those for which T_1 is some polynomial and T_2 is $\sum_{i=1}^{n} X_i$ or the sample mean $\bar{X}_n$. In Lukacs (1963), LL, KLR, and Gordon (1973), T_1 is at most of fourth degree, whereas in Heller (1983) it is of a higher degree. The only

* The work of this author was performed while he was a visitor in the Department of Statistics, State University of New York at Buffalo

distributions so characterized are the normal, gamma, Poisson-type, Cauchy, geometric, binomial, and negative binomial. In Seshadri (1983), T_1 is not a polynomial and a series of characterizations of the inverse Gaussian distribution is obtained.

As noted in LL (p. 109) and Heller (1983), the problem of characterizing distributions by use of zero regression of an arbitrarily high degree polynomial statistic on the sample mean is of a complicated nature with difficulties of two kinds. First, it involves solving a differential equation of high degree in terms of the characteristic function (c.f.) of the distribution being characterized and, secondly, determining which solutions are c.f.'s.

A referee has pointed out that Heller (1979, 1983) attempted to show that there is a linkage between methods of obtaining polynomial statistics which have constant regression on the mean, for a given distribution, methods of proof for characterization theorems, and special function theory (placed within the context of group representation theory). Within this framework, different types of special functions (and thus of characterizations) are shown to derive from different Lie algebras or different irreducible representations of the same Lie algebra, or different realizations of the same representation of the same Lie algebra.

In this paper, we derive series of new characterizations by zero regression properties for the distributions in the class of natural exponential families having power variance functions (abbreviated NEF-PVF's). Such a class has been introduced independently and in different contexts by several authors. Tweedie (1984) and Jørgensen (1987) have treated this class in the context of generalized linear models and exponential dispersion models. Bar-Lev and Enis (1986) (BLE) have presented it in the context of an investigation of reproducible exponential families. This class includes as particular cases the following families: normal, Poisson-type, gamma, all families generated by stable distributions with characteristic exponent an element of the unit interval, and families of compound Poisson distributions generated by gamma variates.

For the case where T_1 is a suitably chosen polynomial statistic of arbitrarily high degree and T_2 is $\sum_{i=1}^{n} X_i$, we provide a series of new characterizations for NEF-PVF distributions. In obtaining such characterizations, we present a unified approach both in finding the polynomial statistics which have zero regression on $\sum_1^n X_i$, and in proving the subsequent characterization theorems. This approach, which is based on certain relations holding among cumulants of NEF-PVF distributions, can also be exploited for obtaining similar characterizations for other exponential families.

In Sect. 2, we briefly review some basic properties of NEF-PVF distributions and reparameterize the corresponding c.f.'s in terms of cumulants of order r and $r+1$, $r=1,2,\ldots$. Such reparameterizations are useful in proving the subsequent theorems. In Sect. 3, we present characterization theorems for NEF-PVF distributions, while their proofs are relegated to Sect. 4.

2. NEF-PVF's: Some Preliminaries

For later use, we present in this section a basic review (excerpted from BLE) of NEF-PVF's and reparameterizations of such families. Let $\mathscr{F}=\{F_\theta:\theta\in\Theta\subset R\}$ be a linear exponential family of order 1 (hereafter, called a natural exponential family (NEF)) with members

$$dF_{\theta}(x) = h(x)\exp\{\theta x + c(\theta)\}\, dv(x), \tag{2.1}$$

where v is a σ-finite measure on some Borel set of the real line. $\mathscr{F}$ is assumed to be minimal and steep. For $\theta \in \operatorname{int}\Theta$, let $\mu = \mu(\theta) = -dc(\theta)/d\theta$, $\Omega = \mu(\operatorname{int}\Theta)$, and $(V(\mu), \Omega)$ be the mean value, the mean parameter space, and the variance function, respectively, corresponding to (2.1).

Definition 2.1. A NEF $\mathscr{F}$ is said to have a power variance function if its variance function is of the form $V(\mu) = \alpha\mu^{\gamma}$, $\mu \in \Omega$, for some constants $\alpha \neq 0$ and γ, called the scale and power parameters, respectively. □

NEF-PVF's were shown in BLE to have Ω of the form R, R^{+}, or R^{-}. For a NEF-PVF, $\gamma = 0$ iff $\Omega = R$. A NEF-PVF with variance function $(\alpha\mu^{\gamma}, \Omega = R^{-})$ can be considered as the reflection about the origin of a NEF-PVF with variance function $(\alpha\mu^{\gamma}, \Omega = R^{+})$. Accordingly, we restrict our attention throughout the sequel to results concerning Ω of the form R or R^{+}. Since $\mathscr{F}$ is steep, Ω is equal to the interior of the convex support of (2.1); i.e., the members of a NEF-PVF $\mathscr{F}$ are concentrated on R (if $\gamma = 0$) or on R^{+} (if $\gamma \neq 0$). The following theorem provides a complete identification of NEF-PVF distributions.

Theorem 2.1 (BLE, Sect. 4). *Let $\mathscr{F}$ be a NEF-PVF with power parameter γ, then*

(i) *it is necessary that $\gamma \notin (-\infty, 0) \cup (0, 1)$;*

(ii) *the γ-values* 0, 1, *and* 2 *correspond to the families of normal, Poisson-type, and gamma distributions, respectively;*

(iii) *for any fixed $\gamma \in (1, 2)$, $\mathscr{F}$ is a family of compound Poisson distributions generated by gamma variates; and*

(iv) *for any fixed $\gamma \in (2, \infty)$, $\mathscr{F}$ is the family generated by a stable distribution concentrated on* $(0, \infty)$ *and possessing a characteristic exponent equal to* $(2-\gamma)/(1-\gamma)$. □

Families of distributions which are NEF-PVF's constitute a rich class, as for each $\gamma \in \{0\} \cup [1, \infty)$ there corresponds an exponential family. The subclass indicated in (iv) of Theorem 2.1 contains as special cases the families of modified Bessel-type ($\gamma = 2.5$), inverse Gaussian ($\gamma = 3$) and Whittaker-type ($\gamma = 4$) distributions. Hougaard (1986) considered the utility of the latter subclass of distributions as survival distributions in the context of constructing life table methods for heterogeneous populations.

All NEF-PVF distributions are infinitely divisible. NEF-PVF's with power parameter $\gamma = 0, 1$ or $\gamma \in (1, 2)$ are regular with parameter space $\Theta = R$ if $\gamma = 0, 1$, and $\Theta = R^{-}$ if $\gamma \in (1, 2)$. NEF-PVF's with $\gamma \in (2, \infty)$ are steep (but nonregular) with $\operatorname{int}\Theta = R^{-}$. Any NEF-PVF $\mathscr{F}$ (with $\Omega = R^{+}$ or $\Omega = R$) can be considered as being a two-parameter family parameterized by α (the scale parameter) and θ, where $(\alpha, \theta) \in (0, \infty) \times \Theta$ (see BLE, Sect. 3). We henceforth consider $\mathscr{F}$ as being a two-parameter family.

Let κ_j be the j-th cumulant of a NEF-PVF distribution. For further use, it is more convenient to work with parameterizations of $\mathscr{F}$ in terms of (κ_r, κ_{r+1}), $r = 1, 2, \ldots$, rather than (α, θ). We first reparameterize $\mathscr{F}$ in terms of (κ_1, κ_2). By using BLE (Sects. 2 and 6), we obtain the following results: if $\gamma=0$ ($\Omega=R$) then $(\kappa_1,\kappa_2)\in R\times R^+$, $\alpha=\kappa_2$ and $\theta=\kappa_2/\kappa_1$; while if $\gamma\geqq 1$ ($\Omega=R^+$) then $(\kappa_1,\kappa_2)\in R^+\times R^+$, $\alpha=\kappa_2/\kappa_1^\gamma$ and $\theta=\kappa_1/[\kappa_2(1-\gamma)]$ for $\gamma>1$, or $\theta=(\kappa_1/\kappa_2)\log\kappa_1$ for $\gamma=1$. The mapping $(\alpha,\theta)\to(\kappa_1,\kappa_2)$ is one-to-one between the sets $R^+\times\operatorname{int}\Theta$ and $R^+\times R^+$. Let $f^*(t:\gamma)$ and $h^*(t:\gamma)=\log f^*(t:\gamma)$ be the c.f. and the cumulant c.f., respectively, corresponding to a NEF-PVF distribution with power parameter γ. Then, by use of the above results and BLE (Sect. 2) we can represent $h^*(t:\gamma)$ in terms of (κ_1,κ_2) as

$$h^*(t:\gamma)=\begin{cases} i\kappa_1 t-\kappa_2^2t^2/2, & \gamma=0\\ (\kappa_1^2/\kappa_2)[e^{it\kappa_2/\kappa_1}-1], & \gamma=1\\ -(\kappa_1^2/\kappa_2)\log[1-it(\kappa_1/\kappa_2)], & \gamma=2\\ \dfrac{\kappa_1^2}{\kappa_2(2-\gamma)}\{[1+it(1-\gamma)(\kappa_2/\kappa_1)]^\varrho-1\}, & \gamma\in(1,2)\cup(2,\infty), \end{cases} \tag{2.2}$$

where, $\varrho\equiv(2-\gamma)/(1-\gamma)$ for $\gamma\in(1,2)\cup(2,\infty)$.

For $\gamma=0$ (the normal case), $\kappa_j=0$ for $j=3,4,\ldots$. Accordingly, we express $h^*(t:\gamma)$ in terms of (κ_r,κ_{r+1}), $r=2,3,\ldots$, only for $\gamma\geqq 1$. To obtain such expressions, we employ some of the results of BLE (Sect. 6). Substituting $r=1$ into (6.5), and $(q,p)=(r-1,r-2)$ with $r\geqq 2$ into (6.6), respectively, of BLE, we obtain the following relations

$$\kappa_{r+2}\kappa_r-\beta_r(\gamma)\kappa_{r+1}^2=0,\qquad \gamma\geqq 1,\ r=1,2,\ldots, \tag{2.3}$$

where

$$\beta_r(\gamma)=[r\gamma-(r-1)]/[(r-1)\gamma-(r-2)]. \tag{2.4}$$

[We also define $\beta_0(\gamma)\equiv 1$ for all γ, and occasionally write β_r instead of $\beta_r(\gamma)$, for simplicity.] Use of (2.3) with $r=1$, yields $\kappa_3\kappa_1=\beta_1(\gamma)\kappa_2^2$. Since κ_1, κ_2, and $\beta_1(\gamma)$ are positive, κ_3 is positive, too. Substituting $\kappa_1=\beta_1(\gamma)\kappa_2^2/\kappa_3$ in (2.2), we obtain an expression for $h^*(t:\gamma)$ in terms of (κ_2,κ_3). By successively repeating the same argument for $r=2,3,\ldots$, it follows that the κ_r's are positive and

$$h^*(t:\gamma)=\begin{cases} (\kappa_r^{r+1}/\kappa_{r+1}^r)[e^{it\kappa_{r+1}/\kappa_r}-1], & \gamma=1\\ -\left[\prod_{j=0}^{r-1}\beta_j^{j+1}\right](\kappa_r^{r+1}/\kappa_{r+1}^r)\log\left[1-\dfrac{it(\kappa_{r+1}/\kappa_r)}{\prod_{j=0}^{r-1}\beta_j}\right], & \gamma=2\\ \dfrac{1}{2-\gamma}\left[\prod_{j=0}^{r-1}\beta_j^{j+1}\right]\dfrac{\kappa_r^{r+1}}{\kappa_{r+1}^r}\left\{\left[1+\dfrac{it(1-\gamma)\kappa_{r+1}}{\left(\prod_{j=0}^{r-1}\beta_j\right)\kappa_r}\right]^\varrho-1\right\}, & \gamma\in(1,2)\cup(2,\infty). \end{cases} \tag{2.5}$$

Note that the mappings $(\kappa_r,\kappa_{r+1})\to(\kappa_{r+1},\kappa_{r+2})$, $r=1,2,\ldots$, are one-to-one from $R^+\times R^+$ onto itself.

3. Characterization Theorems for NEF-PVF Distributions

Let F be a distribution function and let $f(t), h(t)=\log f(t)$, μ_i, and κ_i be the c.f., the cumulant c.f., the i-th moment, and the i-th cumulant, respectively, associated with F. We assume that the p-th moment of F exists. Let $(X_1, \ldots, X_n)$ be a random sample of size n taken from F, and define $L_j = \sum_{k=1}^{n} X_k^j$, $j=1,2,\ldots$. We construct a polynomial statistic, $\hat{T}_{q,p}$, whose expectation is $\kappa_p\kappa_q$, $q \leqq p$. We then use $\hat{T}_{q,p}$ to construct an estimator, $\hat{S}_r(\gamma)$, whose expectation is

$$S_r(\gamma)=\kappa_{r+2}\kappa_r-\beta_r(\gamma)\kappa_{r+1}^2, \qquad r=1,2,\ldots,\gamma \geqq 1,$$

where $\beta_r(\gamma)$ is defined in (2.4). [We note here that, following Sect. 2, if F is a NEF-PVF distribution then: (i) if $\gamma=0$, $\kappa_p\kappa_q=0$, $p=3,4,\ldots$, and (ii) if $\gamma \geqq 1$, $S_r(\gamma)=0$, $r=1,2,\ldots$.]

By requiring $\hat{T}_{q,p}$ and $\hat{S}_r(\gamma)$ to have zero regression on L_1 and imposing mild conditions on F, we obtain easily solved equations in terms of the derivatives of f whose solutions can be identified as the c.f.'s corresponding to NEF-PVF distributions.

The construction of $\hat{T}_{q,p}$ is made as follows. Let $f^{(p)}(t)=d^p f(t)/dt^p$ and $h^{(p)}(t) = d^p h(t)/dt^p$. By LL (p. 99), the product $h^{(p)}(t)\,h^{(q)}(t)$ can be represented in some δ-neighborhood of zero (denoted by N_δ) as

$$h^{(p)}(t)\,h^{(q)}(t) = \sum \lambda_{s_1\ldots s_p}^{(p)} \lambda_{l_1\ldots l_q}^{(q)} \left\{\prod_{j=1}^{p} [f^{(j)}(t)/f(t)]^{s_j}\right\}\left\{\prod_{j=1}^{q} [f^{(j)}(t)/f(t)]^{l_j}\right\}, \tag{3.1}$$

where the summation in (3.1) is extended over all nonnegative integers $s_1, \ldots, s_p$ and $l_1, \ldots, l_q$ satisfying

$$\sum_{j=1}^{p} j s_j = p, \qquad \sum_{j=1}^{q} j l_j = q. \tag{3.2}$$

Substituting $t=0$ in (3.1) and noting that $h^{(j)}(0)=i^j\kappa_j$, we obtain

$$\kappa_p\kappa_q=\sum \lambda_{s_1\ldots s_p}^{(p)} \lambda_{l_1\ldots l_q}^{(q)} [\mu_1^{s_1+l_1}\mu_2^{s_2+l_2}\ldots\mu_q^{s_q+l_q}\mu_{q+1}^{s_{q+1}}\ldots\mu_p^{s_p}]. \tag{3.3}$$

Define $m_0 \equiv 0$, $m_r = \sum_{j=1}^{r} (s_j+l_j)$, $r=1,\ldots,q$, and $m_{q+r}=m_q+\sum_{j=1}^{r} s_{q+j}$, $r=1,\ldots,p-q$. Let $\hat{T}_{q,p}$ denote the statistic obtained by replacing in (3.3) the term in square brackets with

$$(1/n_{(m_p)})\sum{}^* \left\{\prod_{d=1}^{p} \prod_{k=m_{d-1}+1}^{m_d} X_{j_k}^d\right\}, \quad \text{where} \quad n_{(m_p)} \equiv n(n-1)\ldots(n-(m_p-1))$$

and the summation $\sum^*$ runs over all distinct indices j_i, $i=1,\ldots,m_p$, satisfying $1 \leqq j_i \leqq n$; i.e.,

$$\hat{T}_{q,p}=\sum \lambda_{s_1\ldots s_p}^{(p)} \lambda_{l_1\ldots l_q}^{(q)} (1/n_{(m_p)}) \left(\sum{}^* \left\{\prod_{d=1}^{p} \prod_{k=m_{d-1}+1}^{m_d} X_{j_k}^d\right\}\right). \tag{3.4}$$

The following lemma presents some properties of $\hat{T}_{q,p}$.

Lemma 3.1. *Assume that the p-th cumulant of F exists and* $n \geqq p+q$. *Let* $\hat{T}_{q,p}$ *be given by* (3.4), *then*

(i) $E(\hat{T}_{q,p}e^{itL_1}) = i^{-(p+q)}h^{(p)}(t)\,h^{(q)}(t)\,f^n(t), \qquad t \in N_\delta,$
(ii) $E(\hat{T}_{q,p}) = \kappa_q \kappa_p.$

Proof. (i) Using the i.i.d. property of the X_j's and the fact that $E(X_{jk}^d e^{itX_{jk}}) = i^{-d}f^{(d)}(t)$, $d=1,\dots,p$, we obtain

$$\sum{}^* E\left[e^{itL_1}\prod_{d=1}^{p}\prod_{k=m_{d-1}+1}^{m_d} X_{jk}^d\right] = \sum{}^* f^{n-m_p}(t)\prod_{d=1}^{p}\left[i^{-d}f^{(d)}(t)\right]^{m_d-m_{d-1}}$$
$$= n_{(m_p)} i^{-(p+q)} f^n(t)\left\{\prod_{d=1}^{q}\left[f^{(d)}(t)/f(t)\right]^{s_d+l_d}\right\}\left\{\prod_{d=q+1}^{p}\left[f^{(d)}(t)/f(t)\right]^{s_d}\right\}.$$

Hence, for $t \in N_\delta$ we have

$$E(\hat{T}_{q,p}e^{itL_1}) = i^{-(p+q)}f^n(t)\sum \lambda^{(p)}_{s_1\dots s_p}\lambda^{(q)}_{l_1\dots l_q}\left\{\prod_{d=1}^{q}\left[f^{(d)}(t)/f(t)\right]^{s_d+l_d}\right\}$$
$$\cdot\left\{\prod_{d=q+1}^{p}\left[f^{(d)}(t)/f(t)\right]^{s_d}\right\},$$

which is the desired result.

(ii) This statement is verified by substituting $t=0$ in (i). □

We now define

$$\hat{S}_r(\gamma) = \hat{T}_{r,r+2} - \beta_r(\gamma)\,\hat{T}_{r+1,r+1}, \qquad r = 1,2,\dots, \gamma \geqq 1. \tag{3.5}$$

The next corollary is an immediate application of Lemma 3.1.

Corollary 3.1. *Assume that F possesses a finite moment of order* $r+2$ *and that* $n \geqq 2r+2$. *Let* $\hat{S}_r(\gamma)$ *be given by* (3.5), *then for* $r=1,2,\dots$ *and* $\gamma \geqq 1$,

(i) $E(\hat{S}_r(\gamma)e^{itL_1}) = i^{-(2r+2)}f^n(t)\,[h^{(r)}(t)\,h^{(r+2)}(t) - \beta_r(\gamma)\,(h^{(r+1)}(t))^2],\ t \in N_\delta,$
(ii) $E(\hat{S}_r(\gamma)) = S_r(\gamma).$ □

We exemplify the form of $\hat{S}_r(\gamma)$ for the case $r=1$. Here, $\beta_1(\gamma)=\gamma$ and

$$\hat{T}_{1,3} = (1/n_{(2)})\sum X_j^3 X_k - (3/n_{(3)})\sum X_j^2 X_k X_l + (2/n_{(4)})\sum X_j X_k X_l X_m,$$
$$\hat{T}_{2,2} = (1/n_{(2)})\sum X_j^2 X_k^2 - (2/n_{(3)})\sum X_j^2 X_k X_l + (1/n_{(4)})\sum X_j X_k X_l X_m,$$

where the summations are all taken over all distinct indices, j, k, l, m. In terms of the L_j's, $\hat{S}_1(\gamma)$ has the form

$$\hat{S}_1(\gamma) = (1/n_{(4)})\{(n^2+n+4)L_3L_1 - (n^2+n)L_4 - 3(n+1)L_2L_1^2$$
$$+ 3(n-1)L_2^2 + 2L_1^4 + \gamma[(n^2-3n+3)L_2^2 - (n^2-n)L_4$$
$$- 2nL_2L_1^2 + 4(n-1)L_3L_1 + L_1^4]\}.$$

We now present characterization theorems for NEF-PVF distributions which are based on the statistics $\hat{T}_{q,p}$ and $\hat{S}_r(\gamma)$. The proofs of these theorems are presented in Sect. 4. In what follows, we assume that the X_j's have a common nondegenerate distribution F. Theorem 3.1 characterizes the normal distribution.

Theorem 3.1. *Let p and q be any two integers satisfying $1 \leqq q \leqq p$, $p \geqq 3$, and let $n \geqq p+q$. Assume that F possesses a finite p-th moment. Then $\hat{T}_{q,p}$ has zero regression on L_1 iff F is the normal distribution.* □

We note that we can generalize and obtain additional characterizations of the normal distribution, of the type presented in Theorem 3.1. We construct an unbiased estimator of $\kappa_q \kappa_p \kappa_v (v \geqq 3)$, say $\hat{T}_{q,p,v}$, in the same manner as $\hat{T}_{q,p}$ was constructed. We are then able to prove that $\hat{T}_{q,p,v}$ has zero regression on L_1 iff F is normal. Such kind of results are, in a sense, a generalization of Theorem 6.2.5 of LL. In the latter theorem, the authors obtained a characterization of the normal distribution by requiring the p-th statistic (an unbiased polynomial statistic for κ_p), $p \geqq 3$, to have zero regression on L_1.

In the following theorems, we consider characterizations of NEF-PVF distributions with power parameter $\gamma \geqq 1$.

Theorem 3.2. *Assume that F has a finite third moment with $\kappa_1 > 0$. For $n \geqq 4$ and $\gamma \geqq 1$, $\hat{S}_1(\gamma)$ has zero regression on L_1 iff F is a NEF-PVF distribution with power parameter γ.* □

Before presenting the next theorem, we introduce a notation taken from Lukacs (1970, p. 149). We write lext [F] = 0 if $F(x)=0$ for $x<0$, and $F(x)>0$ for $x>0$.

Theorem 3.3. *Assume that F has a finite moment of order $r+2$ $(r \geqq 2)$ and satisfies: $\kappa_r \neq 0$, $\kappa_{r+1} \neq 0$, and* lext $[F]=0$. *For $n \geqq 2r+2$ and $\gamma \geqq 1$, $\hat{S}_r(\gamma)$ has zero regression on L_1 iff F is a NEF-PVF distribution with power parameter γ.* □

Lukacs (1956, 1962) presents a characterization theorem for a family of distributions which includes the Poisson distribution as a special case. The next theorem generalizes the result of Lukacs and provides characterizations based on the zero regression of $\hat{S}_r(\gamma)$ on L_1. For each $\gamma \geqq 1$, the family of distributions so characterized includes NEF-PVF distributions with power parameter γ. Such a family possesses c.f.'s of the form

$$f(t:\gamma) = \exp\{h^*(t:\gamma) + P_k(t)\}, \tag{3.6}$$

where $h^*(t:\gamma)$ is the cumulant c.f. corresponding to a NEF-PVF distribution with power parameter γ and $P_k(t)$ is a polynomial in t with degree $k \leqq 3$, for $\gamma > 1$, and degree $k \leqq 2$, for $\gamma = 1$.

Theorem 3.4. *Assume that F has a finite moment of order $r+2$ $(r \geqq 3)$ with $\kappa_r > 0$ and $\kappa_{r+1} > 0$. For $n \geqq 2r+2$ and $\gamma \geqq 1$, $\hat{S}_r(\gamma)$ has zero regression on L_1 iff F has a c.f. of the form given by* (3.6). □

We end this section by making some comments concerning further characterizations of the type presented so far. The techniques used in this study have been based on the relation $S_r(\gamma)=\kappa_r\kappa_{r+2}-\beta_r(\gamma)\kappa_{r+1}^2=0$, which holds for NEF-PVF distributions. Adoption of similar techniques enables us to obtain characterizations for general NEF's. This is made as follows. Let $\mathscr{F}_0$ be a particular NEF we wish to characterize and express $\mathscr{F}_0$ by its variance function. Suitable differentiation of the latter can lead to various relations holding among the cumulants of $\mathscr{F}_0$ (e.g., BLE, Sect. 6). Such relations can always be given the general form $g(\kappa_{j_1},\ldots,\kappa_{j_m})=0$, where $(\kappa_{j_1},\ldots,\kappa_{j_m})$ is a certain set of m cumulants. Among these relations, we choose g_0 which is polynomial in the κ_{j_i}'s and construct an unbiased polynomial statistic $\hat{g}_0$ for g_0 in a manner analogous to that used for $S_r(\gamma)$. Then, by requiring $\hat{g}_0$ to have zero regression on L_1, we obtain a differential equation, one solution of which is the c.f. associated with $\mathscr{F}_0$. To show that the latter c.f. is the only feasible solution, we may have to impose some further conditions. The number of conditions added depends heavily on a suitable choice of g_0. Indeed, such a kind of procedure has been actually employed in most of the references cited in Sect. 1. The present authors have obtained further characterizations of the distributions appearing in those references by employing the above described technique. Details are omitted for the sake of brevity.

4. Proofs of Theorems 3.1–3.4

We prove only the necessity parts of Theorems 3.1–3.4, as the sufficiency parts are easily verified. The proofs are based on a fundamental lemma, frequently used in characterization problems (cf. Lemma 1.1.1 of KLR). This lemma states: If $E(Y)$ exists, then Y has zero regression on X iff

$$E(YE^{itX})=0, \qquad t\in R. \tag{4.1}$$

Proof of Theorem 3.1. Suppose that $\hat{T}_{q,p}$ has zero regression on L_1. It follows from (4.1) and Lemma 3.1 that

$$h^{(p)}(t)\,h^{(q)}(t)=0, \qquad t\in N_\delta. \tag{4.2}$$

We first show that relation (4.2) implies that $h^{(p)}(t)=0$ on N_δ. To this end assume there exists $t_0\in N_\delta$ for which $h^{(p)}(t_0)\neq 0$. Since the p-th moment of F exists, $h^{(p)}(t)$ is continuous and there exists a neighborhood of t_0, say $N_{\delta_1}(t_0)$, such that $N_{\delta_1}(t_0)\subset N_\delta$ and $h^{(p)}(t)\neq 0$ on $N_{\delta_1}(t_0)$. Hence, by (4.2) we must have, $h^{(q)}(t)=0$ on $N_{\delta_1}(t_0)$. But $q\leqq p$, thus we obtain $h^{(p)}(t)=0$ on $N_{\delta_1}(t_0)$, a contradiction. Therefore, $h^{(p)}(t)=0$ on N_δ, the general solution of which is a polynomial of degree at most $p-1$. By using the initial conditions ($h(0)=0$, $h^{(j)}(0)=i^j\kappa_j$, $j=1,\ldots,p-1$), the theory of analytic c.f.'s and the theorem of Marcinkiewicz (see Lukacs (1970), Corollary to Theorem 7.3.3) we obtain that F is normal. □

Before proving the remaining theorems, we present a lemma whose proof is omitted as it can easily be verified by induction.

Lemma 4.1. (i) *Let* $\gamma \in (1,2)\cup(2,\infty)$ *and* $r=1,2,\ldots,$ *then*

$$\varrho = r + 1/(1-\beta_r), \tag{4.3}$$

$$1-\gamma = (1-\beta_r)\prod_{j=0}^{r-1}\beta_j, \tag{4.4}$$

$$2-\gamma = \left[\prod_{j=0}^{r-1}\beta_j^{j+1}\right]\left[\prod_{j=1}^{r}(1-j(1-\beta_j))\right], \tag{4.5}$$

$$(1-\beta_r)^r \prod_{j=1}^{r}(j+1/(1-\beta_r)) = \prod_{j=1}^{r}(1+j(1-\beta_r)), \tag{4.6}$$

and

$$1 = \varrho_{(r)}\frac{(1-\gamma)^r}{(2-\gamma)}\frac{l_1}{l_2^r}, \tag{4.7}$$

where $\varrho_{(r)} = \varrho(\varrho-1)\varrho(\varrho-1)\ldots(\varrho-(r-1))$, $l_1 = \prod_{j=0}^{r-1}\beta_j^{j+1}$, *and* $l_2 = \prod_{j=0}^{r-1}\beta_j$.

(ii) *Let* $\gamma = 2$ *and* $r=1,2,\ldots,$ *then*

$$\prod_{j=0}^{r-1}\beta_j^{j+1} = r^r/(r-1)!\,. \quad \square \tag{4.8}$$

We now prove Theorems 3.2 and 3.3, simultaneously. In these proofs, we distinguish between three different cases relating to γ-values. These are A) $\gamma=1$, B) $\gamma\in(1,2)\cup(2,\infty)$, and C) $\gamma=2$. Note that the indices r of the $S_r(\gamma)$'s in Theorems 3.2 and 3.3 are $r=1$ and $r\geqq 2$, respectively. Accordingly, for each of the cases A, B, and C, we consider two subcases: (i) $r=1$, and (ii) $r\geqq 2$, and refer to the different assumptions on F linked with each such subcase.

Proof of Theorems 3.2 and 3.3. Let $\gamma\geqq 1$, $r=1,2,\ldots,$ and assume that $\hat{S}_r(\gamma)$ has zero regression on L_1. By (4.1) and Corollary 3.1, we have

$$h^{(r+2)}(t)\,h^{(r)}(t) = \beta_r[h^{(r+1)}(t)]^2, \qquad t\in N_\delta. \tag{4.9}$$

Since $\kappa_j \neq 0$ for $j=r,\ r+1$, and $h^{(j)}(\cdot)$ is continuous, there exists a common neighborhood of the origin in which $h^{(r)}(\cdot)$ and $h^{(r+1)}(\cdot)$ do not vanish. This neighborhood can be taken to be N_δ without loss of generality. Thus, by dividing both sides of (4.9) by $h^{(r)}(t)h^{(r+1)}(t)$, we get $(d/dt)\log\{h^{(r+1)}(t)/[h^{(r)}(t)]^{\beta_r}\}=0$, or

$$h^{(r+1)}(t)/[h^{(r)}(t)]^{\beta_r} = m, \qquad t\in N_\delta, \tag{4.10}$$

where m is a constant determined by the initial conditions

$$h(0)=0,\ h^{(j)}(0)=i^j\kappa_j, \qquad j=1,2,\ldots, \tag{4.11}$$

as

$$m = i^{r+1}\kappa_{r+1}/[i^r\kappa_r]^{\beta_r}. \tag{4.12}$$

We now consider the three cases, A, B, and C, separately.

A) $\gamma=1$. Here, $\beta_r=\beta_r(1)=1$, $r=1,2,\ldots$, and (4.12) and (4.10) become, respectively, $m=i\kappa_{r+1}/\kappa_r$ and

$$h^{(r+1)}(t)=mh^{(r)}(t), \qquad t\in N_\delta. \tag{4.13}$$

(4.13) is a linear differential equation of order $(r+1)$ whose general solution is

$$h(t)=c_0+\sum_{j=1}^{r-1} c_j(it)^j/j!+c_r e^{mt}, \qquad t\in N_\delta, \tag{4.14}$$

where the c_j's are determined by (4.11) as

$$c_0=0, c_j=\kappa_j-\kappa_r^{r+1-j}/\kappa_{r+1}^{r-j} \quad \text{for} \quad j=1,2,\ldots,r-1, \quad \text{and} \quad c_r=\kappa_r^{r+1}/\kappa_{r+1}^r.$$

From the theory of analytic c.f.'s (cf. Lukacs, 1956, Lemma 4.1), it follows that $h(t)$ in (4.14) can be extended to the whole real axis. We now consider below the two subcases relating to r-values.

i) $r=1$: for this subcase, (4.14) becomes, $h(t)=(\kappa_1^2/\kappa_2)[e^{it\kappa_2/\kappa_1}-1]$. Since $\kappa_2>0$ (F is nondegenerate) and $\kappa_1>0$ (an assumption of Theorem 3.2), we obtain by comparing $h(t)$ with $h^*(t:1)$ in (2.2) that the corresponding F is a NEF-PVF distribution with power parameter $\gamma=1$.

ii) $r\geqq 2$: we have

$$f(t)=\exp\left\{(\kappa_r^{r+1}/\kappa_{r+1}^r)[e^{it\kappa_{r+1}/\kappa_r}-1]+\sum_{j=1}^{r-1} c_j(it)^j/j!\right\}, \qquad t\in R. \tag{4.15}$$

Since $f(t)$ in (4.15) is analytic and lext $[F]=0$ (an assumption of Theorem 3.3), $f(t)$ satisfies (cf. Kawata, 1972, Theorem 11.5.6)

$$\limsup_{y\to\infty}[y^{-1}\log f(iy)]=0, \tag{4.16}$$

where

$$y^{-1}\log f(iy)$$
$$=y^{-1}\left\{(\kappa_r^{r+1}/\kappa_{r+1}^r)[e^{-y\kappa_{r+1}/\kappa_r}-1]+\sum_{j=1}^{r-1} c_j(-y)^j/j!\right\}, \qquad y>0. \tag{4.17}$$

We prove that (4.16) implies that the coefficients c_j, $j=1,\ldots,r-1$ vanish. For this purpose, we distinguish between two cases concerning κ_r and κ_{r+1}: 1) $\kappa_{r+1}/\kappa_r<0$, and 2) $\kappa_{r+1}/\kappa_r>0$. For 1), we divide $y^{-1}\log f(iy)$ by y^r and obtain by L'Hospital's rule that $\lim_{y\to\infty}[y^{-(r+1)}\log f(iy)]=\infty$. The latter result implies $\lim_{y\to\infty}[y^{-1}\log f(iy)]=\infty$, and this contradicts (4.16). Thus, case 1) is not feasible. For case 2), we clearly have, $\lim_{y\to\infty} y^{-1}(\kappa_r^{r+1}/\kappa_{r+1}^r)[e^{-y\kappa_{r+1}/\kappa_r}-1]=0$. The latter relation together with (4.16) and (4.17) imply that $\lim_{y\to\infty}\sum_{j=1}^{r-1} c_j(-1)^j y^{j-1}/j!=0$.

Hence the c_j's must vanish, in particular

$$c_1 = \kappa_1 - \kappa_r^r/\kappa_{r+1}^{r-1} = 0, \qquad c_2 = \kappa_2 - \kappa_r^{r-1}/\kappa_{r+1}^{r-2} = 0. \tag{4.18}$$

Since $\kappa_2 > 0$ and $\kappa_1 > 0$ (lext $[F]=0$), it follows from (4.18) that $\kappa_j > 0, j = r, r+1$. Finally, substituting in (4.15), $c_j = 0, j = 1, \ldots, r-1$, and comparing the resulting expression with (2.5) (for $\gamma = 1$), we get the desired result.

B) $\gamma \in (1,2) \cup (2, \infty)$. Setting $y = h^{(r)}(t)$ in (4.10), we obtain $dy/y^{\beta_r} = m\,dt$ or $h^{(r)}(t) = [(1-\beta_r)(mt + c_r)]^{1/(1-\beta_r)}$ for $t \in N_\delta$, where m is given in (4.12) and

$$c_r = (i^r \kappa_r)^{1-\beta_r}/(1-\beta_r), \qquad \beta_r \neq 1, r = 1, 2, \ldots. \tag{4.19}$$

Hence, by using (4.3), we can write

$$h(t) = D(t) + \sum_{j=0}^{r-1} c_j (it)^j / j!, \qquad t \in N_\delta, \tag{4.20}$$

where

$$D(t) = \frac{[(1-\beta_r)c_r]^\varrho \, [(m/c_r)t + 1]^\varrho}{m^r (1-\beta_r)^r \prod_{j=1}^{r} (j + 1/(1-\beta_r))}. \tag{4.21}$$

We express $D(t)$ differently. By using (4.3), (4.4), (4.12), and (4.19), we obtain

$$\frac{m}{c_r} = \frac{i(1-\gamma)(\kappa_{r+1}/\kappa_r)}{\prod_{j=0}^{r-1} \beta_j} e^{2\pi i u \beta_r}, \qquad u = 0, \pm 1, \pm 2, \ldots, \tag{4.22}$$

and

$$\frac{[(1-\beta_r)c_r]^\varrho}{m^r} = \frac{\kappa_r^{r+1}}{\kappa_{r+1}^r} e^{2\pi i \varrho \beta_r v} e^{2\pi i \varrho w}, \qquad v, w = 0, \pm 1, \pm 2, \ldots. \tag{4.23}$$

From (4.5) and (4.6) we have

$$(1-\beta_r)^r \prod_{j=1}^{r} (j + 1/(1-\beta_r)) = (2-\gamma) \prod_{j=0}^{r-1} \beta_j^{j+1}. \tag{4.24}$$

Substituting (4.22), (4.23), and (4.24) into (4.21), and then the resulting expression into (4.20), we obtain

$$h(t) = \frac{1}{2-\gamma} \left[\prod_{j=0}^{r-1} \beta_j^{j+1} \right] [\kappa_r^{r+1}/\kappa_{r+1}^r] e^{2\pi i \varrho \beta_r v} e^{2\pi i \varrho w}$$
$$\times \left\{ \left[1 + \frac{it(1-\gamma)(\kappa_{r+1}/\kappa_r) e^{2\pi i u \beta_r}}{\prod_{j=0}^{r-1} \beta_j} \right]^\varrho - 1 \right\} + \sum_{j=1}^{r-1} c_j (it)^j / j!, \tag{4.25}$$

$$u, v, w = 0, \pm 1, \pm 2, \ldots, t \in N_\delta.$$

We show that only one of the solutions given by (4.25) is feasible. To this end, we differentiate (4.25) to obtain $h^{(r)}(t)$ and $h^{(r+1)}(t)$, and then substitute $h^j(0)=i^j\kappa_j, j=r, r+1$, we find

$$i^j\kappa_j = i^j\varrho_{(j)}[(1-\gamma)^j/(2-\gamma)]\,(l_1/l_2^j)\kappa_j e^{2\pi i\varrho\beta_r v} e^{2\pi i\varrho w}[e^{2\pi i u\beta_r}]^j,$$

for $j=r,\ r+1$, and $u,v,w=0,\pm 1,\pm 2,\dots$, where $\varrho_{(j)}$, l_1, and l_2 are defined in Lemma 4.1. Use of (4.7) in the latter relations, yields for $j=r,\ r+1$, $e^{2\pi i\varrho\beta_r v}e^{2\pi i\varrho w}[e^{2\pi i u\beta_r}]^r=1$, $e^{2\pi i\varrho\beta_r v}e^{2\pi i\varrho w}[e^{2\pi i u\beta_r}]^{r+1}=1$, from which we obtain that $e^{2\pi i u\beta_r}=1$ and $e^{2\pi i\varrho\beta_r v}e^{2\pi i\varrho w}=1$. Substituting the two latter identities in (4.25), we get

$$h(t)=\frac{1}{2-\gamma}\left[\prod_{j=0}^{r-1}\beta_j^{j+1}\right][\kappa_r^{r+1}/\kappa_{r+1}^r]\quad\left\{\left[1+\frac{it(1-\gamma)}{\prod_{j=0}^{r-1}\beta_j}(\kappa_{r+1}/\kappa_r)\right]^{\varrho}-1\right\}$$
$$+\sum_{j=1}^{r-1}c_j(it)^j/j!,\qquad t\in N_\delta, \tag{4.26}$$

where the c_j's are determined by (4.11) as

$$c_j=\kappa_j-\varrho_{(j)}[(1-\gamma)^j/(2-\gamma)]\,(l_1/l_2^j)\,[\kappa_r^{r+1-j}/\kappa_{r+1}^{r-j}],\qquad j=1,\dots,r-1.$$

Clearly, $h(t)$ in (4.26) is analytic and can be extended to the whole real line. The proofs of Theorems 3.2 and 3.3 can now be completed in manners analogous to those of the two subcases $r=1$ and $r\geq 2$ of case A. Details are omitted for the sake of brevity.

C) $\gamma=2$. By noting that $r+1/(1-\beta_r)=0$ for this case, we immediately obtain from (4.10)

$$h(t)=E(t)+\sum_{j=0}^{r-1}c_j(it)^j/j!,\qquad r=1,2,\dots,t\in N_\delta, \tag{4.27}$$

where

$$E(t)=\frac{1}{(1-\beta_r)^r m^r\prod_{j=1}^{r-1}(j+1/(1-\beta_r))}\log\{(1-\beta_r)c_r[(m/c_r)t+1]\}, \tag{4.28}$$

m and c_r are given by (4.12) and (4.19), respectively, and $\prod_{j=1}^{0}$ is defined to be 1.

We now investigate this case in a manner analogous to case B [see the lines following Eq. (4.21)]. By using (4.8), we obtain

$$(1-\beta_r)^r\prod_{j=1}^{r-1}(j+1/(1-\beta_r))=-1\Big/\prod_{j=0}^{r-1}\beta_j^{j+1}. \tag{4.29}$$

By noting that $\beta_r=1+1/r$ for $\gamma=2$, we can express m^r and m/c_r, as

$$m^r=\kappa_{r+1}^r/\kappa_r^{r+1}, \tag{4.30}$$

and

$$(m/c_r)=\left[-i\Big/\prod_{j=1}^{r-1}\beta_j\right](\kappa_{r+1}/\kappa_r)e^{2\pi iw/r},\qquad w=0,\pm 1,\pm 2,\ldots. \tag{4.31}$$

Substitute (4.29), (4.30), and (4.31) into (4.28) and then the resulting expression into (4.27). Then, a similar kind of argumentation as in case B can be used to show that $e^{2\pi iw/r}=1$. This results in

$$h(t)=-\left(\prod_{j=0}^{r-1}\beta_j^{j+1}\right)(\kappa_r^{r+1}/\kappa_{r+1}^r)\log\left[1-it(\kappa_{r+1}/\kappa_r)\Big/\prod_{j=0}^{r-1}\beta_j\right]$$
$$+\sum_{j=0}^{r-1}c_j(it)^j/j!,\qquad t\in N_\delta, \tag{4.32}$$

where the c_j's are determined by (4.11) as

$$c_0=0, c_j=\kappa_j-(j-1)!\,(\kappa_r^{r+1-j}/\kappa_{r+1}^{r-j})\left[\prod_{j=0}^{r-1}\beta_j^{j+1}\Big/\prod_{j=0}^{r-1}\beta_j\right]^j,\qquad j=1,\ldots,r-1.$$

The rest of the proof can be completed in a manner analogous to that of case A. Details are omitted. □

Proof of Theorem 3.4. For $\gamma\geqq 1$, we denote by $f(t:\gamma)$ the c.f. corresponding to F, and assume that $\hat{S}_r(\gamma)$ has zero regression on L_1. It is apparent, from the proofs of Theorems 3.2 and 3.3 and the assumption (of the present theorem) concerning the positiveness of κ_r and κ_{r+1}, that $f(t:\gamma)$ has the form, $f(t:\gamma)=\exp\{h^*(t:\gamma)+P_k(t)\}$, where $P_k(t)$ is polynomial of degree k, and $h^*(t:\gamma)$ is the cumulant c.f. corresponding to a NEF-PVF distribution with power parameter γ [see (2.5)]. Consider the function $m(z)=g(z)\exp[P_k(z)]$, where $g(z)$ is an infinitely divisible c.f. regular in a half plane $\mathrm{Im}(z)>-d(d>0)$. Christensen (1962, Theorem 4.1) showed that a necessary condition for $m(z)$ to be a c.f. is that $k\leqq 3$. In our case, $\exp[h^*(z:\gamma)]$ is infinitely divisible (Sect. 2) and, as easily can be shown, regular in a half plane, hence $k\leqq 3$. For $\gamma=1$ (the Poisson-type case), Lukacs (1970, Theorem 7.3.5) showed that $k\leqq 2$. □

Acknowledgements. We would like to thank Peter Enis for helpful suggestions and comments on this paper.

References

Bar-Lev, S.K., Enis, P.: Reproducibility and natural exponential families with power variance functions. Ann. Statist. **14**, 1507–1522 (1986)

Christensen, I.F.: Some further extensions of a theorem of Marcinkiewicz. Pacific J. Math. **12**, 59–67 (1962)

Gordon, F.S.: Characterizations of populations using regressions properties. Ann. Statist. **1**, 114–126 (1973)

Heller, B.: Special functions and characterizations of probability distributions by constant regression of polynomial statistics on the mean. Unpublished Ph. D. dissertation, Department of Statistics, University of Chicago 1979

Heller, B.: Special functions and characterizations of probability distributions by zero regression properties. J. Multi. Anal. **13**, 473–487 (1983)

Hougaard, P.: Survival models for heterogeneous populations derived from stable distributions. Biometrika **73**, 387–396 (1986)

Jørgensen, B.: Exponential dispersion models (with discussion). To appear in J. Roy Statist. Soc. Ser. B **49** (1987)

Kagan, A.M., Linnik, Yu. V., Rao, C.R.: Characterizations problems in mathematical statistics. New York: Wiley 1973

Kawata, T.: Fourier analysis in probability theory. New York London: Academic Press 1972

Lukacs, E.: Characterizations of populations by properties of suitable statistics. Proc. 3rd Berkeley Symp. on Math. Statist. and Probab. Theory **2**, 195–214 (1956)

Lukacs, E.: On the characterization of a family of populations which includes the Poisson population. Ann. Univ. Sci. Budapest Eotuos Sect. Math. **3–4**, 159–175 (1962)

Lukacs, E.: Characterization problems for discrete distributions. In: Patil, G. (ed.) Proceedings International Symposium on Classical and Contagious Discrete Distributions, pp. 65–73. Oxford New York: Pergamon 1963

Lukacs, E.: Characteristic functions, 2nd edn. New York: Hafner 1970

Lukacs, E., Laha, R.G.: Applications of characteristic functions. London: Griffin 1964

Seshadri, V.: The inverse Gaussian distributions: some properties and characterizations. Can. J. Statist. **11**, 131–136 (1983)

Tweedie, M.C.K.: An index which distinguishes between some important exponential families. In: Ghosh, J.K., Roy, J. (eds.) Statistics: applications and new directions. Proceedings of the Indian Statistical Institute Golden Jubilee International Conference, pp. 579–604. Calcutta: Indian Statistical Institute 1984

Received July 3, 1985; in revised form July 25, 1987