The Annals of Statistics
1982, Vol. 10, No. 3, 979–989

AN EXPONENTIAL SUBFAMILY WHICH ADMITS UMPU TESTS BASED ON A SINGLE TEST STATISTIC

BY SHAUL K. BAR-LEV AND BENJAMIN REISER[1]

University of Haifa, Israel

Let $f(x:\theta) = a(x)\exp\{\theta_1 u_1(x) + \theta_2 u_2(x) + c(\theta)\}$, $\theta = (\theta_1, \theta_2) \in \Theta \subset R^2$, be a density with respect to the Lebesgue measure on the real line which characterizes a two-parameter exponential family of distributions. Let (θ_1, η_2) be the mixed parameters, where $\eta_2 = E\{u_2(X)\}$. Assume that θ_2 can be represented as $\theta_2 = -\theta_1\varphi'(\eta_2)$ where $\varphi'(\eta_2) = d\varphi(\eta_2)/d\eta_2$ for some function $\varphi(\eta_2)$. Let $(X_1, \cdots, X_n)$ be independent random variables having a common density $f(x:\theta)$ and set $T_i = \sum_{j=1}^n u_i(X_j)$, $i = 1, 2$. It is shown that if $u_2(x)$ is a 1-1 function then the random variables T_2 and $Z_n = T_1 - n\varphi(T_2/n)$ are independent and that the statistic Z_n is ancillary for θ_2 in the presence of θ_1 (i.e. the density of Z_n depends on θ_1 only). Furthermore, the density of Z_n belongs to the one-parameter exponential family with natural parameter θ_1. These results enable us to construct uniformly most powerful unbiased (UMPU) tests for various hypotheses concerning the parameter θ_1 which are based on the statistic Z_n.

1. Introduction. Let

$$f(x:\theta) = a(x)\exp\{\theta_1 u_1(x) + \theta_2 u_2(x) + c(\theta)\}, \quad \theta = (\theta_1, \theta_2), \tag{1.1}$$

be a density function with respect to (w.r.t.) the Lebesgue measure on the real line which characterizes a two-parameter exponential family of distributions. In this case it is required that $a(x) > 0$ and that $u_i(x)$, $i = 1, 2$ be absolutely continuous on the interior of the convex-support of (1.1) with $u_i'(x) = du_i(x)/dx \not\equiv 0$, $i = 1, 2$. The natural parameter space Θ is defined as the effective domain of $k(\theta)$, denoted by dom $k(\theta)$ (Barndorff-Nielsen, 1978, page 77), where

$$k(\theta) = \exp\{-c(\theta)\} = \int a(x)\exp\{\theta_1 u_1(x) + \theta_2 u_2(x)\}\, dx.$$

It is assumed that Θ has a non-empty interior in R^2 and that u_1 and u_2 are affinely independent w.r.t. the Lebesgue measure.

Let $(X_1, \cdots, X_n)$, $n > 1$, be independent r.v.'s having a common density of the form (1.1), and set $T_i = \sum_{j=1}^n u_i(X_j)$, $i = 1, 2$. Then the following results can be shown to hold (c.f. Lehmann, 1959):

(i) The joint distribution of (T_1, T_2) is a two-parameter exponential family with density (w.r.t. the Lebesgue measure on R^2) of the form

$$f_{T_1,T_2}(t_1, t_2:\theta) = h(t)\exp\{\theta_1 t_1 + \theta_2 t_2 + nc(\theta_1, \theta_2)\}, \quad t = (t_1, t_2).$$

(ii) The conditional distribution of T_1 given $T_2 = t_2$ is a one-parameter exponential family with conditional density function of the form

$$f_{T_1|T_2}(t_1 \mid t_2:\theta_1) = h(t)\exp\{\theta_1 t_1 - \log b(\theta_1, t_2)\}, \quad b(\theta_1, t_2) = \int h(t)\exp(\theta_1 t_1)\, dt_1. \tag{1.2}$$

Received January 1981; revised December 1981.

[1] This research was carried out while the authors were at the Technion-Israel Institute of Technology.

AMS 1970 *subject classifications.* Primary 62F05; secondary 62E10, 62E15.

Key words and phrases. Exponential family, ancillary, uniformly most powerful unbiased tests.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Statistics*. www.jstor.org

(iii) The marginal density function of T_2 is of the form

$$f_{T_2}(t_2:\theta) = \exp\{\theta_2 t_2 + nc(\theta_1, \theta_2) + \log b(\theta_1, t_2)\}. \tag{1.3}$$

Lehmann and Scheffe (1950, 1955) developed procedures for constructing UMPU tests for testing various composite hypotheses concerning one of the natural parameters, say θ_1. Those UMPU tests are essentially derived by using the conditional density function of T_1 given $T_2 = t_2$. However, Lehmann (1959, Chapter 5) remarked that such derivations of UMPU tests "turn out to be inconvenient in application to normal and certain other families of continuous distributions." He stated that in these applications the tests can be given a more convenient form in which they no longer appear as conditional tests in terms of T_1 given $T_2 = t_2$, but are expressed in terms of a single test statistic. For example, if $(X_1, \cdots, X_n)$ are independent r.v.'s having a common $N(\mu, \sigma^2)$ density, then for testing $\sigma^2 \le \sigma_0^2$ vs. $\sigma^2 > \sigma_0^2$ (σ_0^2 is specified), one might reduce the UMPU test to depend on the marginal distribution of S^2 rather than the conditional distribution of $\sum_{i=1}^n X_i^2$ given $\bar{X}_n$.

For identifying the cases where such reduction is valid, Lehmann introduced a theorem (Lehmann, 1959, Theorem 1 in Chapter 5) which gives sufficient conditions for this purpose and enables us to achieve such optimal tests through the marginal distribution of a single test statistic. Such a reduction essentially depends on the existence of a statistic $Z_n = h(T_1, T_2)$ which is independent of T_2 for any value of θ belonging to the boundary parameter set (in the above example, S^2 is independent of $\bar{X}_n$ for any values of μ and σ^2), and such that $h(t_1, t_2)$ is a monotone increasing function of t_1 for each value t_2. However, applications of such a theorem are known in the literature only for particular distributions.

Although our interest in conditions for the existence of the statistic Z_n was originally motivated by hypothesis testing considerations, the problem of delineating cases where statistics are available whose distributions depend only on a subparameter is of more general interest.

In this paper we introduce a subfamily of exponential distributions which satisfy the appropriate Lehmann's conditions simultaneously for various composite hypotheses connected with θ_1 and has interesting properties of itself. This subfamily is characterized by easily checkable functional relationships between its parameters and includes many of the usually discussed members of (1.1). In characterizing this subfamily we use the mixed parameterization (Barndorff-Nielsen, 1978) of the exponential family.

Let $\eta_2 = E\{u_2(X)\}$. If θ_2 can be represented as a function of the mixed parameters θ_1 and η_2 such that $\theta_2 = -\theta_1\varphi'(\eta_2)$, where $\varphi'(\eta_2) = d\varphi(\eta_2)/d\eta_2$ for some function $\varphi(\eta_2)$, then it is proved in the sequel that the r.v.'s T_2 and $Z_n = T_1 - n\varphi(T_2/n)$ are independent such that the marginal distribution of Z_n depends only on θ_1 and possesses a density function which characterizes a one-parameter exponential family. These properties of Z_n are shown to satisfy Lehmann's conditions for the construction of UMPU tests based on its marginal distribution.

For the representation of this subfamily we introduce in Section 2 some of the known properties of the exponential family that we need for our purposes. In Sections 3 and 4 we describe our model and discuss its main properties. Several illustrative examples will appear in Section 5. In Section 6 these results are applied to the construction of UMPU tests based on the marginal distribution of Z_n.

A further application of the results of this paper in the direction of the asymptotic properties of maximum conditional likelihood estimates is presented by Bar-Lev (1981).

2. Preliminaries. Many references deal with the properties of the exponential family. We shall make use of those appearing in Lehmann (1959), Bildikar and Patil (1968), Berk (1972) and Barndorff-Nielsen (1978).

ASSUMPTION 2.1. $\Theta = \text{int}\ \Theta$.

Although we assume that Θ is open, i.e. that (1.1) is regular in the sense of Barndorff-Nielsen (1978), our results still hold for steep exponential families on int Θ (see the remark at the end of Section 4).

LEMMA 2.1. (Lehmann, 1959; Barndorff-Nielsen, 1978; Bildikar and Patil, 1968).
(i) *Θ is a convex set in R^2.*
(ii) *$c(\theta)$, $\theta = (\theta_1, \theta_2)$ is an analytic function of $\theta \in \Theta$.*
(iii) *For any $\theta \in \Theta$ the statistic $u = (u_2, u_2)$ has moments of all orders. In particular*

$$E_\theta(u) = (\eta_1, \eta_2), \quad \eta_i = -\partial c(\theta)/\partial\theta_i, \quad i = 1, 2. \tag{2.1}$$

Let $V_\theta(u)$ denote the covariance matrix of u. Then $V_\theta(u)$ is positive definite (p.d.) *for all $\theta \in \Theta$ and is given by*

$$V_\theta(u) = (\sigma_{ij}), \quad \sigma_{ij} = \mathrm{cov}_\theta(u_i, u_j) = -\,\partial^2 c(\theta)/\partial\theta_i\partial\theta_j, \quad i, j = 1, 2. \tag{2.2}$$

(iv) *Let $T = (T_1, T_2)$, then*

$$E_\theta(T) = (n\eta_1, n\eta_2), \quad V_\theta(T) = (n\sigma_{ij}), \quad i, j = 1, 2. \tag{2.3}$$

(v) *The following two conditions are equivalent:*
  (1) $\mathrm{cov}_\theta(T_1, T_2) = -\,n\partial^2 c(\theta)/\partial\theta_1\partial\theta_2 \equiv 0, \quad \theta \in \Theta.$
  (2) *The* r.v.'s *T_1 and T_2 are independent, and each constitutes a one-parameter exponential family with the natural parameter θ_1 and θ_2 respectively.* □

In order to avoid trivial cases which follow when one of the conditions in (v) is satisfied, we make the following assumption,

ASSUMPTION 2.2. $\qquad \mathrm{cov}_\theta(T_1, T_2) \not\equiv 0, \quad \theta \in \Theta;$

otherwise the marginal density of T_1 itself depends only on θ_1 and is independent of T_2.

LEMMA 2.2. (Barndorff-Nielsen, 1978). *Under Assumption 2.1 the following results hold:*
(i) *The mapping $(\theta_1, \theta_2) \to (\theta_1, \eta_2)$ defined on Θ is a homeomorphism.* (*Thus, (θ_1, η_2) furnishes a parameterization of the exponential family called a mixed parameterization*).
(ii) *The components (θ_1, η_2) in the mixed parameterization are variation independent.* □

Thus the domain of variation of (θ_1, η_2) is of the form $\Theta_1 \times H_2$, where H_2 is connected (using the connectedness of Θ) and $\Theta_1 \times H_2$ is open.

3. The model and the main results.

ASSUMPTION 3.1. θ_2 can be represented as:

$$\theta_2 = -\,\theta_1\varphi'(\eta_2). \tag{3.1}$$

The existence of such a representation for θ_2 is easily checked since by Lemma 2.1 (iii), $\eta_2 = -\partial c(\theta_1, \theta_2)/\partial\theta_2$ is a strictly monotone increasing function of θ_2 for each fixed $\theta_1 \in \Theta_1$, and thus an expression for θ_2 is easily derived.

REMARK 1. If θ_2 is of the form $-(\theta_1 + d)\varphi'(\eta_2) + e$, with d and e arbitrary constants then the results in the sequel hold for this case too, by redefining $\theta_1^* = \theta_1 + d$, $\theta_2^* = \theta_2 - e$ as our new natural parameters.

REMARK 2. It can be shown quite easily that if we are expressing θ_2, $c(\theta_1, \theta_2)$ and $\eta_1(\theta_1, \theta_2)$ by means of θ_1 and η_2, then these functions possess partial derivatives of all orders w.r.t. $(\theta_1, \eta_2) \in \Theta_1 \times H_2$. For the sake of brevity we omit the proof.

REMARK 3. A dual representation of θ_2 by means of θ_1 and η_2 as $\theta_2 = \alpha(\theta_1) + \beta(\eta_2)$ for some functions $\alpha(\cdot)$ and $\beta(\cdot)$ is given by Barndorff-Nielsen (1978, Theorem 10.4) in connection with the existence of cuts in the exponential family.

LEMMA 3.1.
(i) $\varphi'(\eta_2)$ *is not identically constant.*
(ii) *The functions* $c(\theta_1, \theta_2)$ *and* $\eta_1(\theta_1, \theta_2)$ *when expressed by means of the mixed parameters* θ_1 *and* η_2 *are of the following forms*:

$$c(\theta_1, \eta_2) = \theta_1 \chi(\eta_2) - M(\theta_1) \tag{3.2}$$

where $\chi(\eta_2) = \eta_2 \varphi'(\eta_2) - \varphi(\eta_2)$

$$\eta_1 = \varphi(\eta_2) + M'(\theta_1), \quad M'(\theta_1) = dM(\theta_1)/d\theta_1 \tag{3.3}$$

and $M(\theta_1)$ *is an infinitely differentiable function on* Θ_1 *for which*

$$M''(\theta_1) = d^2M(\theta_1)/d\theta_1^2 > 0, \quad \forall \theta_1 \in \Theta_1. \tag{3.4}$$

PROOF.
(i) Assuming that $\varphi'(\eta_2) \equiv c$ where c is an arbitrary constant, then by (3.1), $\theta_2 = -c\theta_1$ i.e. θ_1 and θ_2 are linearly dependent, in contradiction to the assumption that Θ is a non-empty open subset of R^2.
(ii) The general solution of the differential equation $\eta_2(\theta_1, \theta_2) = -\, \partial c(\theta_1, \theta_2)/\partial\theta_2$ is of the form

$$-c(\theta_1, \theta_2) = \int \eta_2(\theta_1, \theta_2)\, d\theta_2 + M(\theta_1). \tag{3.5}$$

Consider η_2 in the relation (3.1) as a function of (θ_1, θ_2), then differentiating w.r.t. θ_2 both sides of (3.1) results in

$$1 = -\theta_1 \varphi''(\eta_2)(\partial\eta_2/\partial\theta_2), \quad \partial\eta_2/\partial\theta_2 = \sigma_{22}(\theta) > 0, \quad \forall \theta \in \Theta. \tag{3.6}$$

Using (3.6) in (3.5) and changing the variable of integration in (3.5) to η_2 instead of θ_2, gives

$$-c(\theta_1, \theta_2) = -\theta_1 \int \eta_2 \varphi''(\eta_2)\, d\eta_2 + M(\theta_1)$$

and integration by parts of the integral in the right-hand side of the last equation yields

$$\begin{aligned} -c(\theta_1, \theta_2) = -\, \theta_1 \chi(\eta_2) + M(\theta_1) &= -\theta_1 \varphi'(\eta_2)\eta_2 + \theta_1 \varphi(\eta_2) + M(\theta_1) \\ &= \theta_2 \eta_2 + \theta_1 \varphi(\eta_2) + M(\theta_1). \end{aligned}$$

But

$$\eta_1 = -\partial c(\theta_1, \theta_2)/\partial\theta_1 = \theta_2(\partial\eta_2/\partial\theta_1) + \varphi(\eta_2) + \theta_1 \varphi'(\eta_2)(\partial\eta_2/\partial\theta_1) + M'(\theta_1).$$

The result (3.3) follows through use of Assumption 3.1.

As a first step towards proving (3.4) we show that

$$M''(\theta_1) = \sigma_{11}(\theta) - \sigma_{12}^2(\theta)/\sigma_{22}(\theta). \tag{3.7}$$

Using (3.3) we express $\sigma_{11}(\theta)$ as

$$\sigma_{11}(\theta) = \partial\eta_1(\theta)/\partial\theta_1 = \varphi'(\eta_2)(\partial\eta_2/\partial\theta_1) + M''(\theta_1). \tag{3.8}$$

Differentiation of both sides of (3.1) w.r.t. θ_1 yields $0 = -\varphi'(\eta_2) - \theta_1 \varphi''(\eta_2)(\partial\eta_2/\partial\theta_1)$ or

$$-\varphi'(\eta_2) = \theta_1 \varphi''(\eta_2)(\partial\eta_2/\partial\theta_1). \tag{3.9}$$

From (3.6), $-\theta_1 \varphi''(\eta_2) = 1/(\partial\eta_2(\theta)/\partial\theta_2)$. Substituting this into (3.9) one has

$$\varphi'(\eta_2) = \{\partial\eta_2(\theta)/\partial\theta_1\}/\{\partial\eta_2(\theta)/\partial\theta_2\} = \sigma_{12}(\theta)/\sigma_{22}(\theta)$$

or

$$\sigma_{12}(\theta) = \varphi'(\eta_2)\sigma_{22}(\theta). \tag{3.10}$$

Finally, by using (3.8) and (3.10) we obtain $\sigma_{11}(\theta) - \sigma_{12}^2(\theta)/\sigma_{22}(\theta)$ as exactly $M''(\theta_1)$. Since $V_\theta(u)$ is p.d. (Lemma 2.1) it follows that $M''(\theta_1) > 0$. □

It follows from (3.6) that θ_1 cannot be 0 and that $\theta_1 < 0(>0) \Leftrightarrow \varphi''(\eta_2) > 0(<0)$, and hence, by connectedness of $\Theta_1 \times H_2$, that either

$$\Theta_1 \subset R^-, \quad H_2 \subset \{\eta_2 : \varphi''(\eta_2) > 0\}$$

or

$$\Theta_1 \subset R^+, \quad H_2 \subset \{\eta_2 : \varphi''(\eta_2) < 0\}.$$

Thus the division by θ_1 of both sides of equation (3.1) is allowed, resulting in $-\varphi'(\eta_2) = \theta_2/\theta_1$; i.e., η_2 is a function of (θ_1, θ_2) only through the ratio θ_2/θ_1. This property of η_2 leads to the following theorem.

THEOREM 3.1. *Under Assumptions* 2.1 *and* 2.2 *the distribution of the statistic* $\bar{T}_2 = T_2/n$ *is equivalent to that of* $u_2(X)$ *but with parameters* $n\theta_1$, $n\theta_2$ *(instead of* θ_1, θ_2, *respectively) if and only if* η_2 *is a function of* (θ_1, θ_2) *only through the ratio* θ_2/θ_1.

REMARK 4. Notice that when we are dealing practically with the construction of UMPU tests through the conditional distribution of T_1 given $T_2 = t_2$, then the main technical difficulty arises in finding the marginal distribution of T_2. Theorem 3.1 provides a simple solution since the marginal distribution of $u_2(X)$ can be found from (1.1) by a simple transformation.

PROOF. (i) Assume that η_2 is a function of (θ_1, θ_2) only through the ratio $\theta_2/\theta_1 = -\varphi'(\eta_2)$. Hence $c(\theta_1, \theta_2)$ is of the form (3.2):

$$c(\theta_1, \theta_2) = \theta_1\chi\{\eta_2(\theta_2/\theta_1)\} - M(\theta_1). \tag{3.11}$$

Denote by $\varphi_{u_2}(s)$, $\varphi_{T_2}(s)$ and $\varphi_{\bar{T}_2}(s)$ the characteristic functions (c.f.'s) of u_2, T_2 and $\bar{T}_2$ respectively. Using (1.1), one obtains the following general expressions of these c.f.'s:

$$\varphi_{u_2}(s) = \exp\{c(\theta_1, \theta_2) - c(\theta_1, \theta_2 + is)\} \tag{3.12}$$

$$\varphi_{T_2}(s) = \exp\{nc(\theta_1, \theta_2) - nc(\theta_1, \theta_2 + is)\}. \tag{3.13}$$

Now $\varphi_{\bar{T}_2}(s) = \varphi_{T_2}(s/n) = \exp\{nc(\theta_1, \theta_2) - nc(\theta_1, \theta_2 + is/n)\}$, but for $c(\theta_1, \theta_2)$ of the form (3.11) we have

$$\varphi_{u_2}(s) = \exp\{\theta_1\chi[\eta_2(\theta_2/\theta_1)] - \theta_1\chi[\eta_2((\theta_2 + is)/\theta_1)]\}$$

and

$$\varphi_{\bar{T}_2}(s) = \exp\left\{(n\theta_1)\chi\left[\eta_2\left(\frac{n\theta_2}{n\theta_1}\right)\right] - (n\theta_1)\chi\left[\eta_2\left(\frac{n\theta_2 + is}{n\theta_1}\right)\right]\right\}.$$

By comparing the last two c.f.'s one obtains the c.f. of $\bar{T}_2$ as that of $u_2(X)$ but with parameters $n\theta_1$ and $n\theta_2$ (instead of θ_1 and θ_2 respectively), and thus the desired result.
(ii) Assume that the c.f. of $\bar{T}_2$ is as that of u_2 but with parameters $n\theta_1$ and $n\theta_2$.
Using (3.12), it follows that

$$\varphi_{\bar{T}_2}(s) = \exp\{c(n\theta_1, n\theta_2) - c(n\theta_1, n\theta_2 + is)\}$$

and therefore

$$\varphi_{T_2}(s) = \varphi_{\bar{T}_2}(ns) = \exp\{c(n\theta_1, n\theta_2) - c(n\theta_1, n\theta_2 + nis)\}. \tag{3.14}$$

By comparing (3.13) and (3.14) it can be seen that

$$nc(\theta_1, \theta_2) - nc(\theta_1, \theta_2 + is) = c(n\theta_1, n\theta_2) - c(n\theta_1, n(\theta_2 + is)).$$

Differentiating w.r.t. s and then dividing both sides of the last equation by n gives

$$\frac{\partial c(\theta_1, \theta_2 + is)}{\partial(\theta_2 + is)} = \frac{\partial c(n\theta_1, n(\theta_2 + is))}{\partial[n(\theta_2 + is)]}.$$

In the last equation substitute $s = 0$, obtaining

$$\frac{\partial c(\theta_1, \theta_2)}{\partial \theta_2} = \frac{\partial c(u, v)}{\partial v}\bigg|_{u=n\theta_1, v=n\theta_2}.$$

Hence $\partial c(\theta_1, \theta_2)/\partial\theta_2$ is a homogeneous function of zero degree with general solution of the form (Franklin, 1940)

$$\partial c(\theta_1, \theta_2)/\partial\theta_2 = \psi(\theta_2/\theta_1).$$

The result follows since the left hand side of the last equation is exactly $-\eta_2(\theta_1, \theta_2)$. □

In the following we assume that $u_2(\cdot)$ is a 1-1 function. Nevertheless, notice that the above results hold without the restriction made by such an assumption.

ASSUMPTION 3.2. $u_2 = g(x)$ where $g(x)$ is a 1-1 function on the convex support of (1.1).

If we denote the appropriate inverse function by $g^{-1}(\cdot)$, it can be shown that the distribution of u_2 belongs to a two-parameter exponential family with density function of the form

$$f_{u_2}(u_2 : \theta) = r(u_2)\exp\{\theta_1 u_1[g^{-1}(u_2)] + \theta_2 u_2 + c(\theta_1, \theta_2)\}, \tag{3.15}$$

where $r(u_2) = a[g^{-1}(u_2)]\,|dg^{-1}(u_2)/du_2|$ and $c(\theta_1, \theta_2)$ is determined by (3.2).

As a consequence of Theorem 3.1 it follows that the r.v. $\bar{T}_2$ has the following density function

$$f_{\bar{T}_2}(\bar{t}_2 : \theta) = r(\bar{t}_2)\exp\{n\theta_1 u_1[g^{-1}(\bar{t}_2)] + n\theta_2\bar{t}_2 + n\theta_1\chi(\eta_2) - M(n\theta_1)\}. \tag{3.16}$$

In the following theorem we present one of the main results of this paper.

THEOREM 3.2. *Under Assumptions* 3.1 *and* 3.2 *the following properties hold*:

(i)
$$u_1[g^{-1}(\bar{T}_2)] = \varphi(\bar{T}_2) \quad \text{a.s.}$$

(ii) *The distribution of the statistic* $Z_n = T_1 - n\varphi(\bar{T}_2)$ *belongs to a one-parameter exponential family with natural parameter* θ_1 *and density of the form,*

$$f_{Z_n}(z_n : \theta_1) = d(z_n)\exp\{\theta_1 z_n - [nM(\theta_1) - M(n\theta_1)]\}, \quad \theta_1 \in \Theta_1. \tag{3.17}$$

(iii) *For any* $\theta \in \Theta$ *the* r.v.'s Z_n *and* T_2 *are independent.*

PROOF.
(i). This result is proved below in Lemma 4.1.
(ii) and (iii). Denote by $\varphi_{T_1}(s : t_2)$ the c.f. of the conditional distribution of T_1 given $T_2 = t_2$. Then by (1.2) it follows that

$$\varphi_{T_1}(s : t_2) = E\{\exp(isT_1) \mid T_2 = t_2\} = b(\theta_1 + is, t_2)/b(\theta_1, t_2). \tag{3.18}$$

An expression for $b(\theta_1, t_2)$ in our model will be derived in the following manner. Using (3.16) one finds that the density function of T_2 can be written as

$$f_{T_2}(t_2 : \theta) = [r(\bar{t}_2)/n]\exp\{\theta_2 t_2 + n[\theta_1\chi(\eta_2) - M(\theta_1)] + n\theta_1 u_1[g^{-1}(\bar{t}_2)] + nM(\theta_1) - M(n\theta_1)\}.$$

Then, by comparing the last equation with (1.3), it follows that

$$\log b(\theta_1, t_2) = n\theta_1 u_1[g^{-1}(\bar{t}_2)] - [M(n\theta_1) - nM(\theta_1)] + \log[r(\bar{t}_2)/n]. \tag{3.19}$$

Substituting (3.19) in (3.18), and using the notation

$$H_n(\theta_1) = nM(\theta_1) - M(n\theta_1),$$

we obtain

$$\varphi_{T_1}(s:t_2) = E\{\exp(isT_1) \mid T_2 = t_2\} = \exp\{(is)nu_1[g^{-1}(\bar{t}_2)] + H_n(\theta_1 + is) - H_n(\theta_1)\}$$

(which holds for almost all values of t_2), and therefore

$$E\{\exp\{is[T_1 - nu_1[g^{-1}(\bar{T}_2)]]\} \mid T_2 = t_2\}$$
$$= E\{\exp(isZ_n) \mid T_2 = t_2\} = \exp\{H_n(\theta_1 + is) - H_n(\theta_1)\}.$$

Hence Z_n and T_2 are independent. From the last equation we obtain

$$\varphi_{Z_n}(s) = E\{E[\exp(isZ_n) \mid T_2]\} = \exp\{H_n(\theta_1 + is) - H_n(\theta_1)\}.$$

Therefore, the distribution of Z_n depends on θ_1 only. The marginal density function of Z_n can be shown to be of the form (3.17) by using the 1-1 two-dimensional transformation $(T_1, T_2) \to (Z_n, T_2)$. For the sake of brevity, details are omitted. □

4. Other properties of the statistic Z_n. In the next two lemmas we prove the existence of some additional properties of the functions $\varphi(\eta_2)$ and $M(\theta_1)$. Those properties enable us to characterize those of Z_n.

LEMMA 4.1. *Let C_2 denote the convex support of the marginal distribution of $\bar{T}_2$, which coincides with that of $u_2(X)$. Then, under Assumptions* 3.1 *and* 3.2,

(4.1) $$\varphi(x) = u_1[g^{-1}(x)] + c, \quad \forall x \in \text{int } C_2,$$

where c is an unspecified constant which can be chosen without loss of generality as 0.

PROOF. We first show that $H_2 = \text{int } C_2$. Denote by S the support of the joint distribution of $(\bar{T}_1, \bar{T}_2)$ $n > 1$, by C the convex support of S (i.e. $C = cl \text{ conv } S$) and by H, $\eta(\text{int } \Theta)$. Theorem 9.2 of Barndorff-Nielsen (1978) states that the exponential model given by (1.1) is steep iff $H = \text{int } C$. By Assumption 2.1, (1.1) is regular which implies the steepness of the model (Barndorff-Nielsen, 1978, Theorem 8.2). Let $\Pi_y(B)$ denote the projection of a set $B \subset R^2$ onto the coordinate y. Then C_2 and H_2 can be expressed as $C_2 = cl \text{ conv } \Pi_{\bar{T}_2}(S)$, $H_2 = \Pi_{\bar{T}_2}(\text{int } C) = \Pi_{\bar{T}_2}(\text{int conv } S)$. By applying Lemma 5.4 of Barndorff-Nielsen (1978) and noting that, from (i) of Section 1, int $S \neq \phi$, the desired result follows.

As noted above, the conditional distribution of T_1 for given $T_2 = t_2$ constitutes a one-parameter exponential family having conditional density function of the form (1.2). So that for each fixed t_2, $\log b(\theta_1, t_2)$ plays the same role in the conditional density (1.2) as that of the function $-nc(\theta)$ in the joint density of (T_1, T_2). Therefore we can immediately state the following properties of $\log b(\theta_1, t_2)$ analogous to the properties of $-nc(\theta)$ and valid for all $\theta_1 \in \Theta_1$ and almost all t_2:

$$E_{\theta_1}(T_1 \mid t_2) = \partial \log b(\theta_1, t_2)/\partial\theta_1,$$
$$V_{\theta_1}(T_1 \mid t_2) = \partial^2 \log b(\theta_1, t_2)/\partial\theta_1^2 > 0,$$

where the notations $E_{\theta_1}(T_1 \mid t_2)$, $V_{\theta_1}(T_1 \mid t_2)$ stand for $E_{\theta_1}(T_1 \mid T_2 = t_2)$, $V_{\theta_1}(T_1 \mid T_2 = t_2)$ respectively.

It is claimed that in general we have the following result:

(4.2) $$(1/n)(\partial \log b(\theta_1, T_2)/\partial\theta_1) \to_P \eta_1 \quad \text{as} \quad n \to \infty$$

whatever be the value of θ.

The convergence in probability in (4.2) is valid since

$$E\{(1/n)(\partial \log b(\theta_1, T_2)/\partial\theta_1)\} = (1/n)E_\theta\{E_{\theta_1}(T_1 \mid T_2)\} = E_\theta(u_1) = \eta_1,$$

and by using the known identity

$$V_\theta(T_1) = V_\theta\{E_{\theta_1}(T_1 \mid T_2)\} + E_\theta\{V_{\theta_1}(T_1 \mid T_2)\}$$

it follows that

$$V_\theta\{(1/n)(\partial \log b(\theta_1, T_2)/\partial\theta_1)\} = (1/n^2)V_\theta\{E_{\theta_1}(T_1 \mid T_2)\} \le (1/n^2)V_\theta(T_1) = \sigma_{11}(\theta)/n \to 0$$

as $n \to \infty$, which implies the result (4.2).

Under our model, $(1/n)\partial \log b(\theta_1, T_2)/\partial\theta_1$ has the form

$$(4.3)\quad (1/n)\partial \log b(\theta_1, T_2)/\partial\theta_1 = u_1[g^{-1}(\bar{T}_2)] + M'(\theta_1) - M'(n\theta_1), \quad M'(n\theta_1) = M'(x)\,|_{x=n\theta_1}.$$

Since $u_1(x)$ and $u_2(x)$ are continuous (see the discussion following equation (1.1)) and $u_2 = g(x)$is 1-1 (Assumption 3.2) then $u_1[g^{-1}(u_2)]$ is continuous on $H_2 = \text{int } C_2$, and it is implied that

$$(4.4)\quad u_1[g^{-1}(\bar{T}_2)] \to_{\text{a.s.}} u_1[g^{-1}(\eta_2)].$$

($\bar{T}_2 \to_{\text{a.s.}} \eta_2$ as the mean of independent r.v.'s having common expectation η_2). Finally by using (3.3) we conclude from (4.2) and (4.4) that
(i) $\lim_{n\to\infty} M'(n\theta_1) = F(\theta_1)$, say, exists for any $\theta_1 \in \Theta_1$;

$$\text{(ii)}\quad \varphi(\eta_2) = u_1[g^{-1}(\eta_2)] - F(\theta_1).$$

It follows from (ii) that $F(\theta_1) \equiv c$, an arbitrary constant; otherwise η_2 can be expressed as a function of θ_1 only, and this fact implies that $\sigma_{22}(\theta) = \partial\eta_2/\partial\theta_2 \equiv 0$ which contradicts our assumption that Θ has a non-empty interior in R^2, i.e. $u_1[g^{-1}(\eta_2)] = \varphi(\eta_2) + c$. If c does not equal 0 it can be made so by the following construction of the exponential family: substituting in (3.15), $u_1[g^{-1}(u_2)] = \varphi(u_2) + c$, $u_2 \in \text{int } C_2 = H_2$, one gets

$$f_{u_2}(u_2:\theta) = r(u_2)\exp\{\theta_1\varphi(u_2) + \theta_2 u_2 + \theta_1\chi(\eta_2) - M_1(\theta_1)\},$$

where $M_1(\theta_1) = M(\theta_1) - c\theta_1$. If we now repeat the whole process that leads to the conclusion (i), one obtains $\lim_{n\to\infty} M_1'(n\theta_1) = 0$. Thus if $c \neq 0$ we assume that such a construction has been made. □

The following result deals with the form of the convex support of the distribution of Z_n and uses the fact that either $\Theta_1 \subset R^-$ or $\Theta_1 \subset R^+$.

LEMMA 4.2.

(i) *If* $\Theta_1 \subset R^-$ *then* $Z_n > 0$ a.s. (ii) *If* $\Theta_1 \subset R^+$ *then* $Z_n < 0$ a.s.

PROOF. (i) If $\Theta_1 \subset R^-$ then H_2 is a non-empty open interval in R which is contained in $\{\eta_2:\varphi''(\eta_2) > 0\}$ (see discussion immediately following Lemma 3.1). Thus $\varphi(\eta_2)$ is a strictly convex function on H_2. Let u_2^i, $i = 1, \cdots, n$ be independent observations from (3.15). Since $H_2 = \text{int } C_2$ it follows that with probability one, $u_2^i \in H_2$, $i = 1, \cdots, n$, and $\sum_{i=1}^n u_2^i/n \in H_2$. Using the fact that $\varphi(\eta_2)$ is a strictly convex function on H_2, one gets that, with probability one,

$$\varphi(\bar{t}_2) = \varphi\{\textstyle\sum_{i=1}^n u_2^i/n\} < (1/n)\{\sum_{i=1}^n \varphi(u_2^i)\} = (1/n)\{\sum_{i=1}^n u_1^i\} = t_1/n$$

i.e.

$$Z_n = T_1 - n\varphi(\bar{T}_2) > 0 \quad \text{a.s.}$$

(ii) The proof of (ii) is analogous, using the fact that $\varphi(\eta_2)$ is a strictly concave function on H_2. □

REMARK. All the results derived above under the assumption that the model is regular (Θ open) hold on int Θ for steep models. For steep models it can be shown that (Barndorff-Nielsen, 1978, Theorem 9.1 (ii), Theorem 9.3 and the discussion immediately preceeding Theorem 5.34) the mapping $(\theta_1, \theta_2) \to (\theta_1, \eta_2)$ defined on int Θ = int(dom $k(\theta)$) is a homeomorphism and that θ_1 and η_2 are variation independent. These conditions are sufficient for our derivations.

5. Example

EXAMPLE 5.1. *Normal distribution.*
(i) $f(x:\mu, \sigma^2) = (2\pi)^{-1/2}\exp\{-(x-\mu)^2/2\sigma^2\}$, $x, \mu \in R$, $\sigma^2 \in R^+$;
(ii) $\theta_1 = -1/2\sigma^2$, $\theta_2 = \mu/\sigma^2$, $\Theta = R^- \times R$, $u_1(X) = X^2$, $u_2(X) = X$, $T_1 = \sum_{i=1}^n X_i^2$, $T_2 = \sum_{i=1}^n X_i$;
(iii) $c(\theta_1, \theta_2) = \theta_2^2/4\theta_1 + (½)\log(-2\theta_1)$;
(iv) $\eta_2 = -\theta_2/2\theta_1$, $\theta_2 = -2\theta_1\eta_2$;
(v) $\eta_1 = \eta_2^2 - 1/2\theta_1$, $\varphi(\eta_2) = \eta_2^2$, $M'(\theta_1) = -1/2\theta_1$;
(vi) $Z_n = T_1 - n\varphi(\bar{T}_2) = (n-1)S^2 > 0$ a.s., Z_n and T_2 are independent and the distribution of Z_n depends on $\theta_1 = -1/2\sigma^2$ only. □

EXAMPLE 5.2. *Gamma distribution.*
(i) $f(x:\lambda, \alpha) = (-\lambda)^\alpha x^{\alpha-1}e^{\lambda x}/\Gamma(\alpha)$, $x, \alpha \in R^+, \lambda \in R^-$;
(ii) $\theta_1 = \alpha$, $\theta_2 = \lambda$, $\Theta = R^+ \times R^-$, $u_1(X) = \log X$, $u_2(X) = X$;
(iii) $c(\alpha, \lambda) = \alpha\log(-\lambda) - \log\Gamma(\alpha)$;
(iv) $\eta_2 = -\alpha/\lambda$, $\lambda = -\alpha/\eta_2$, $\eta_1 = \log\eta_2 + d\log\Gamma(\alpha)/d\alpha - \log\alpha$, $\varphi(\eta_2) = \log\eta_2$, $M'(\alpha) = d\log\Gamma(\alpha)/d\alpha - \log\alpha$;
(v) $Z_n = T_1 - n\varphi(T_2/n) = n\cdot\log\{[\Pi_{i=1}^n X_i]^{1/n}/\bar{X}_n\} < 0$ a.s. and Z_n and $\bar{X}_n$ are independent. The independence of the sample mean and the ratio between the geometric mean and the sample mean for the Gamma distribution is well known. (Cox and Lewis, 1966; Glaser, 1976; Bain and Engelhardt, 1975). □

EXAMPLE 5.3. *Inverse Gaussian distribution.* We write the distribution as
(i) $f(x:\alpha, \lambda) = (2\pi)^{-1/2}x^{-3/2}\lambda^{1/2}\exp\{-\alpha x/2 - \lambda/2x + (\alpha\lambda)^{1/2}\}$, $x, \lambda \in R^+, \alpha \in R^+ \cup \{0\}$;
(ii) $\theta_1 = -\lambda/2$, $\theta_2 = -\alpha/2$, $\Theta = R^- \times (R^- \cup \{0\})$, $u_1(X) = 1/X$, $u_2(X) = X$;
(iii) $c(\theta_1, \theta_2) = 2(\theta_1\theta_2)^{1/2} + (½)\log(-2\theta_1)$;
(iv) $\eta_2 = -(\theta_1/\theta_2)^{1/2}$, $\theta_2 = \theta_1/\eta_2^2$, $\eta_1 = (1/\eta_2) - (1/2\theta_1)$, $\varphi(\eta_2) = 1/\eta_2$, $M'(\theta_1) = -1/2\theta_1$;
(v) $Z_n = \sum_{i=1}^n (1/X_i) - (n/\bar{X}_n) > 0$ a.s., Z_n and $\bar{X}_n$ are independent for all $\theta \in \text{int}\,\Theta$ (i.e. $\alpha > 0, \lambda > 0$) using the steepness of the model (Barndorff-Nielsen, 1978, Example 8.4). This property of the Inverse Gaussian distribution has been shown by Tweedie (1957a, 1957b). □

Other examples include the Log-Normal, Log-Gamma and random walk distribution (Johnson and Kotz, 1970, page 149).

The referee has conjectured that the above three examples (and of course 1-1 transformations of them) exhaust the domain of application of Theorem 3.2. This is equivalent to saying that $u_2(X)$ must have one of those three distributions for Theorem 3.2 to hold. We have been unable to confirm or disprove this conjecture.

It can be shown that similar results to those developed above in Sections 3 and 4 hold under analogous assumptions for the bivariate case, i.e. when (1.1) is a bivariate density function w.r.t. the Lebesgue measure on R^2. Assume that (X, Y) is a random vector which has a bivariate density function of the form

$$f_{X,Y}(x, y:\theta) = a(x, y)\exp\{\theta_1 u_1(x, y) + \theta_2 u_2(x, y) + c(\theta)\}, \tag{5.1}$$

where T_i stands for $\sum_{j=1}^n u_i(X_j, Y_j), i = 1, 2,$ with $\{(X_j, Y_j)\}$ $\{(X_j, Y_j)\}_{j=1}^n$ a random sample from (5.1). We illustrate such results by the following example.

EXAMPLE 5.4. *Bivariate Gamma distribution.* A detailed discussion on this distribution is found in Mihram and Hultquist (1967) who termed it a bivariate warning time/failure time distribution.
(i) $f(x, y:a, p, q) = \{a^{p+q}/\Gamma(p)\Gamma(q)\}x^{p-1}(y-x)^{q-1}e^{-ay}$, $y > x > 0$, $a, p, q > 0$. We assume that a and q are the unknown parameters (p is known). Similar results hold when a and p are the unknown parameters and q is known (see (vi) below).

(ii) $\theta_1 = q$, $\theta_2 = -a$, $\Theta = R^+ \times R^-$, $u_1(x, y) = \log(y - x)$, $u_2(x, y) = y$;
(iii) $c(\theta_1, \theta_2) = (p + \theta_1)\log(-\theta_2) - \log \Gamma(\theta_1)$;
(iv) $\eta_2 = -(p + \theta_1)/\theta_2$, $\theta_2 = -(\theta_1/\eta_2) - (p/\eta_2)$. Here θ_2 is of the form $-\theta_1\varphi'(\eta_2) - p\varphi'(\eta_2)$ (see Remark 1 under Assumption 3.1). We therefore define a new natural parameter θ_1^* as $\theta_1^* = (\theta_1 + p)$ so that $\theta_2 = -\theta_1^*\varphi'(\eta_2)$. In this case, $\varphi'(\eta_2) = 1/\eta_2$, $\varphi(\eta_2) = \log \eta_2$, $T_1 = \sum_{j=1}^n \log(Y_j - X_j)$, $T_2 = \sum_{j=1}^n Y_j$.
(v) $Z_n = T_1 - n\varphi(T_2/n) = \sum_{j=1}^n \log(Y_j - X_j) - n(\log \bar{Y}_n) < 0$ a.s. Z_n and $\bar{Y}_n$ are independent. Mihram and Hultquist (1967) have shown that X/Y and Y are independent. This entails the independence between $\log(1 - X/Y)$ and Y, thus implying our result for $n = 1$.
(vi) Assume now that a and p are the unknown parameters (q is known). For this case $u_1(x, y) = \log x$, $u_2(x, y) = y$, $\theta_1 = p + q$, $\theta_2 = -a$. An analogous development yields the independence between $Z_n = \sum_{j=1}^n \log X_j - n(\log \bar{Y}_n)$ and $\bar{Y}_n$, and furthermore that the distribution of Z_n depends on θ_1 only. □

6. UMPU tests based on the statistic Z_n. We are interested in testing the following four types of hypotheses concerning θ_1:

$$H_1\colon \theta_1 \le \theta_1^0 \qquad K_1\colon \theta_1 > \theta_1^0$$

$$H_2\colon \theta_1 \le \theta_1^1 \text{ or } \theta_1 \ge \theta_1^2 \qquad K_2\colon \theta_1^1 < \theta_1 < \theta_1^2$$

$$H_3\colon \theta_1^1 \le \theta_1 \le \theta_1^2 \qquad K_3\colon \theta_1 < \theta_1^1 \text{ or } \theta_1 > \theta_1^2$$

$$H_4\colon \theta_1 = \theta_1^0 \qquad K_4\colon \theta_1 \ne \theta_1^0$$

where $\theta_1^i \in \Theta_1$, $i = 0, 1, 2$ are specified values of θ_1. It can readily be seen that UMPU tests for these four hypotheses can be constructed based on the statistic Z_n since by use of the results derived above, Lehmann's (1959, Chapter 5) Theorem 1 applies. For Examples 5.1, 5.2 and 5.3 this results in tests that are available in the literature (see Lehmann, 1959; Bain and Engelhardt, 1975; Chhikara and Folks, 1976, respectively). These tests were developed separately for each case. Our approach provides a unifying element.

Acknowledgment. The authors are grateful to the referee for his constructive criticisms which resulted in improvements both in content and style.

REFERENCES

BAIN, L. G. and ENGELHARDT, M. (1975). A two-moment chi-square approximation for the statistic $\log(\bar{x}/\tilde{x})$. *J. Amer. Statist. Assoc.* **70** 948–950.

BAR-LEV, S. K. (1981). Asymptotic properties of the maximum conditional likelihood estimate for exponential subfamilies of distributions. Unpublished manuscript.

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

BERK, R. H. (1972). Consistency and asymptotic normality of M.L.E.'s for exponential models. *Ann. Math. Statist.* **43** 193–204.

BILDIKAR, S. and PATIL, G. P. (1968). Multivariate exponential-type distributions. *Ann. Math. Statist.* **39** 1316–1326.

CHHIKARA, R. S. and FOLKS, J. L. (1976). Optimum test procedures for the mean of the first passage time in Brownian motion with positive drift (inverse Gaussian distribution). *Technometrics* **18** 189–193.

COX, D. R. and LEWIS, P. A. W. (1966). *The Statistical Analysis of Series of Events.* Methuen, London.

FRANKLIN, P. (1944). *Methods of Advanced Calculus.* McGraw-Hill, Hightstown, New Jersey.

GLASER, R. E. (1976). The ratio of the geometric mean to the arithmetic mean for a random sample from a gamma distribution. *J. Amer. Statist. Assoc.* **71** 480–487.

JOHNSON, N. L. and KOTZ, S. (1970). *Continuous Univariate Distributions*-1. Houghton Mifflin, Boston.

LEHMANN, E. L. (1959). *Testing Statistical Hypotheses.* Wiley, New York.

LEHMANN, E. L. and SCHEFFE, H. (1950, 1955). Completeness, similar regions and unbiased estimation. *Sankhya* **10** 305–340 and **15** 219–236.

MIHRAM, G. A. and HULTQUIST, R. A. (1967). A bivariate warning-time/failure time distribution. *J. Amer. Statist. Assoc.* **62** 589–599.

TWEEDIE, M. C. K. (1957a). Statistical properties of inverse Gaussian distributions I. *Ann. Math. Statist.* **28** 362–377.

TWEEDIE, M. C. K. (1957b). Statistical properties of inverse Gaussian distributions II. *Ann. Math. Statist.* **28** 696–705.

DEPARTMENT OF STATISTICS
UNIVERSITY OF HAIFA
HAIFA 31999
ISRAEL