The Annals of Statistics
1983, Vol. 11, No. 3, 770–782

REPRODUCTIVE EXPONENTIAL FAMILIES

BY O. BARNDORFF-NIELSEN AND P. BLÆSILD

Aarhus University

Consider a full and steep exponential model $\mathscr{M}$ with model function $a(\theta)b(x)\exp\{\theta\cdot t(x)\}$ and a sample $x_1, \cdots, x_n$ from $\mathscr{M}$. Let $\bar{t} = \{t(x_1) + \cdots + t(x_n)\}/n$ and let $\bar{t} = (\bar{t}_1, \bar{t}_2)$ be a partition of the canonical statistic $\bar{t}$. We say that $\mathscr{M}$ is *reproductive* in t_2 if there exists a function H independent of n such that for every n the marginal model for $\bar{t}_2$ is exponential with $n\theta$ as canonical parameter and $(H(\bar{t}_2), \bar{t}_2)$ as canonical statistic. Furthermore we call $\mathscr{M}$ *strongly reproductive* if these marginal models are all contained in that for $n = 1$. Conditions for these properties to hold are discussed. Reproductive exponential models are shown to allow of a decomposition theorem analogous to the standard decomposition theorem for χ^2-distributed quadratic forms in normal variates. A number of new exponential models are adduced that illustrate the concepts and also seem of some independent interest. In particular, a combination of the inverse Gaussian distributions and the Gaussian distributions is discussed in detail.

1. Introduction. Consider an exponential model $\mathscr{M}$ given by

$$a(\theta)b(x)^{\theta\cdot t(x)} \tag{1.1}$$

where the canonical parameter θ and the canonical statistic t are vectors of dimension k. The mean value of t under (1.1) will be denoted by τ, and Θ will stand for the domain of variation of θ. Furthermore, vectors are taken to be row vectors and the transpose of a vector v is denoted by v^*.

Let $\theta = (\theta_1, \theta_2)$, and $\tau = (\tau_1, \tau_2)$ and $t = (t_1, t_2)$ be similar partitions of θ, τ and t, and let the common dimension of θ_i, τ_i and t_i be denoted by k_i, $i = 1, 2$. Our interest in this paper is with cases where the marginal distributions of t_2 constitute an exponential family that has θ and $(H(t_2), t_2)$, for some vector function H, as corresponding canonical variates, a property that we express by writing

$$t_2 \sim \mathrm{EM}((H(t_2), t_2); \theta) \tag{1.2}$$

where EM is an abbreviation for "exponential model".

In general, t_2 does not follow an exponential model, and even when it does that model need not be of the form (1.2), as shown by the following counterexample.

EXAMPLE 1.1. Let u and v be positive random variables with joint probability density function

$$p(u, v; \phi, \psi) = \frac{1}{\sqrt{2\pi}}(\phi + \sqrt{\psi})v^{-3/2}u\exp(-(½)v^{-1}u^2 - (½)\psi v - \phi u)$$

where $\psi \geq 0$ and $\phi > -\sqrt{\psi}$. Here u follows the negative exponential distribution with parameter $\phi + \sqrt{\psi}$, i.e.

$$p(u; \phi, \psi) = (\phi + \sqrt{\psi})\exp(-(\phi + \sqrt{\psi})u),$$

while the conditional distribution of v given u is inverse Gaussian. We have $\theta = (-½\psi, -\phi)$ and $t = (v, u)$ and

Received September 1982; revised March 1982.

AMS 1980 *subject classifications.* Primary, 62E15; secondary, 62F99.

Key words and phrases. Affine foliations, decomposition, exact tests, generalized linear models, independence, inverse Gaussian distribution.

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

(1.3) $$u \sim \mathrm{EM}(u; -(\phi + \sqrt{\psi})),$$

but (1.3) cannot be recast in the form (1.2). □

In Section 2 we show that, provided $\mathscr{M}$ is full and steep, (1.2) is equivalent to

(1.4) $$t_1 - H(t_2) \perp t_2,$$

where $\perp$ denotes stochastic independence, and also equivalent to

(1.5) $$t_1 - H(t_2) \sim \mathrm{EM}(t_1 - H(t_2); \theta_1).$$

One notes that in (1.5) the distribution of $t_1 - H(t_2)$ depends on θ_1 only.

The models discussed in the next examples 1.2–1.4 have, in fact, the property that (1.2) holds for any sample size and with a fixed H, and this occasions the following definition. For a sample $x_1, \cdots, x_n$ from (1.1) we set $\bar{t} = n^{-1}(t(x_1) + \cdots + t(x_n))$, and we say that (1.1) is *reproductive* in t_2 if

(1.6) $$\bar{t}_2 \sim \mathrm{EM}((H(\bar{t}_2), \bar{t}_2); n\theta)$$

for every $n = 1, 2, \cdots$ and some H independent of n. Since the model for the sample $x_1, \cdots, x_n$ is of the form (1.1) with $\bar{t}$ and $n\theta$ as canonical variates it follows from the equivalence of (1.2), (1.4) and (1.5) that for models which are reproductive in t_2 one has

(1.7) $$\bar{t}_1 - H(\bar{t}_2) \perp \bar{t}_2$$

and

(1.8) $$\bar{t}_1 - H(\bar{t}_2) \sim \mathrm{EM}(\bar{t}_1 - H(\bar{t}_2); n\theta_1).$$

EXAMPLE 1.2. For the normal $N(\xi, \sigma^2)$ distribution we have an exponential representation (1.1) with $t = (x^2, x)$ and $\theta = (-1/(2\sigma^2), \xi/\sigma^2)$. Since, for a sample $x_1, \cdots, x_n$, the mean $\bar{x}$ is distributed as $N(\xi, \sigma^2/n)$, the model is reproductive in x, and (1.7) expresses the independence of s^2 and $\bar{x}$. □

EXAMPLE 1.3. Let $\Gamma(\lambda, \alpha)$ denote the gamma distribution whose probability density function is

(1.9) $$\frac{\alpha^\lambda}{\Gamma(\lambda)} x^{\lambda-1}\exp(-\alpha x).$$

The family of these distributions, both parameters α and λ being considered unknown, is reproductive in x. In fact, the family is exponential with $t = (\ln x, x)$ and $\theta = (\lambda, -\alpha)$ as the canonical variates, and if $x_1, \cdots, x_n$ is a sample from (1.9) then $x. = x_1 + \cdots + x_n$ follows the $\Gamma(n\lambda, \alpha)$ distribution, and this may be paraphrased as $\bar{x} \sim \mathrm{EM}((\ln \bar{x}, \bar{x}), n\theta)$. Furthermore, the well known independence of $\bar{x}$ and $\tilde{x}/\bar{x}$, where $\tilde{x}$ denotes the geometric mean of the observations, may be seen as a consequence of (1.7). □

EXAMPLE 1.4. The inverse Gaussian distribution with probability density function

(1.10) $$\frac{\sqrt{\chi}}{\sqrt{2\pi}} \exp(\sqrt{\chi\psi})x^{-3/2}\exp(-(½)\{\chi x^{-1} + \psi x\})$$

will be denoted by $N^-(\chi, \psi)$. Considering both $\chi > 0$ and $\psi \geq 0$ as unknown, we have an exponential family with $t = (x^{-1}, x)$ and $\theta = -½(\chi, \psi)$, and this is reproductive in x because the mean $\bar{x}$ of a sample $x_1, \cdots, x_n$ from (1.10) has the inverse Gaussian distribution $N^-(n\chi, n\psi)$. The independence of $n^{-1}\Sigma(x_i^{-1} - \bar{x}^{-1})$ and $\bar{x}$ (Tweedie, 1957) is therefore an instance of (1.7). □

Let $D(\theta)$ denote the marginal distribution of t_2 under (1.1) and suppose that (1.2) is fulfilled. It may happen that

$$\bar{t}_2 \sim D(n\theta)$$

for every $n = 1, 2, \cdots$ and we then say that (1.1) is *strongly reproductive* in t_2. This property, which obviously implies reproductivity, actually holds in examples 1.2–1.4. (Examples of models which are reproductive without being strongly reproductive will be discussed in Section 4.)

Let $\kappa(\theta) = -\ln a(\theta)$ denote the cumulant transform of the model $\mathscr{M}$. It may be noted that strong reproductivity of $\mathscr{M}$ in t_2 can be expressed as

$$\kappa_2(\eta; n\theta) = n\kappa_2(n^{-1}\eta; \theta)$$

for $n = 1, 2, \cdots$ and where $\kappa_2(\eta; \theta) = \kappa(\theta + (0, \eta)) - \kappa(\theta)$ is the cumulant transform of t_2 under P_θ. That is, κ_2 considered as a function of both η and θ has a homogeneity property of order one.

From now on we assume that the family of distributions (1.1) is full and steep (in the sense of Barndorff-Nielsen, 1978).

Suppose t_1 and t_2 are of the form $t_1(x) = H(x)$ for some continuous vector valued function H and $t_2(x) = x$, respectively. It is then possible to show that (1.1) is strongly reproductive if and only if $c \text{ int } \Theta \subset \text{int } \Theta$ for every scalar $c > 1$ and θ_2 is of the form

$$\theta_2 = -\theta_1 h(\tau_2)$$

for some $k_1 \times k_2$ matrix-valued function h. The proof will be given at the end of Section 3. Essentially, this result generalises Theorem 3.1 of Bar-Lev and Reiser (1981). These authors proved the validity of the result for $k = 2$, in which case the condition $c \text{ int } \Theta \subset \Theta$ may be deleted. Furthermore, when combined with Theorem 2.1 below the result provides an extension of the other main conclusions of Bar-Lev and Reiser's (1981) paper, as given in (ii) and (iii) of their Theorem 3.2.

The normal, gamma and inverse Gaussian models can also be used as building stones in the construction of interesting examples of reproductive and strongly reproductive models of higher exponential order, cf. Section 4.

For a further example we consider the Wishart distribution.

EXAMPLE 1.5. Let S be a random $d \times d$ dimensional and positive definite matrix that follows the Wishart distribution

$$p(S; \Sigma) = c(f, d) |\Delta|^{f/2} |S|^{(f-d-1)/2} \exp(-\text{tr}(\Delta S)/2) \tag{1.11}$$

where $c(f, d)$ is a norming constant, $\Sigma = f^{-1}ES$, $\Delta = \Sigma^{-1}$ and $f \geq d$, and let S be partitioned as

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}.$$

As is well known (and simple to establish by means of Basu's theorem) (S_{12}, S_{22}) is independent of $S_{11\cdot} = S_{11} - S_{12}S_{22}^{-1}S_{21}$ and this property is an instance of (1.4), with $t_1 = S_{11\cdot}$, $t_2 = (S_{12}, S_{22})$ and $H(t_2) = S_{12}S_{22}^{-1}S_{21}$. The equivalent formulations (1.2) and (1.5) express other well known facts about the Wishart distribution. □

We consider reproductive exponential models in Section 3 and show that such models allow of a decomposition theorem analogous to the standard decomposition theorem for χ^2-distributed quadratic forms in normal variates. We also show that, under a mild regularity condition, the independence result (1.7) may be reformulated as

$$\hat{\theta}_1 \perp \hat{\tau}_2$$

i.e. the two components of the maximum likelihood estimator of the mixed parameter (θ_1, τ_2) are independent. Quite generally, $\hat{\theta}_1$ and $\hat{\tau}_2$ are asymptotically independent, whether we have reproductivity or not, cf. Barndorff-Nielsen (1978), Section 9.8 (vi).

How does one determine whether an exponential model is reproductive or not? If the relevant component t_2 is known then it is often simple to check reproductivity in t_2 by inspection of the cumulant transform $\kappa(\theta) = -\ln a(\theta)$, cf. the examples in Section 4. Another type of condition for reproductivity in t_2 is provided by the relation

$$\tau_1 = m(\theta_1) + H(\tau_2), \tag{1.12}$$

i.e. τ_1 can be written as the sum of a function m of θ_1 and a function H of τ_2 (where H is the same function as occurs in (1.6)). In fact, as will be proved in Section 3, (1.12) is a necessary condition for reproductivity in t_2. We conjecture that this condition is also sufficient, and can show that this is actually the case under certain additional assumptions. However, we have no clue how to prove it in general.

Certain related results are discussed in Barndorff-Nielsen and Blæsild (1983).

2. Criteria for and implications of $t_2 \sim \mathrm{EM}((H(t_2), t_2); \theta)$. We shall repeatedly use the fact that if P_θ denotes the probability measure given by (1.1), if u is a statistic, and if $p(u;\theta)$ is the density of the lifted measure uP_θ with respect to some σ-finite measure dominating the class $\{uP_\theta : \theta \in \Theta\}$ of marginal distributions of u then, for any elements θ_0 and θ of Θ, we have

$$p(u;\theta) = E_{\theta_0}\left\{\frac{dP_\theta}{dP_{\theta_0}} \,\middle|\, u\right\} p(u;\theta_0), \tag{2.1}$$

cf. Barndorff-Nielsen (1978) Section 8.2 (iii).

In order to establish the equivalence of (1.2), (1.4) and (1.5) we need the following lemma. In essence, this result is well known (cf. for instance Patil, 1965) but a fully general and explicit formulation does not seem to be available in the literature.

LEMMA 2.1. *Let $\mathscr{P} = \{P_\theta : \theta \in \Theta\}$ be a parameterised class of probability measures with Θ a subset of R^k, let t be a k-dimensional statistic, and suppose that there exists a function r on Θ such that for every $\theta_0 \in \Theta$ and $\theta \in \Theta$ we have*

$$E_{\theta_0}\exp((\theta - \theta_0)\cdot t) = r(\theta)/r(\theta_0). \tag{2.2}$$

Suppose furthermore that Θ contains an open subset of R^k. Then the family of distributions of t under $\mathscr{P}$ is exponential and

$$\frac{d(tP_\theta)}{d(tP_{\theta_0})} = \frac{r(\theta_0)}{r(\theta)}\exp((\theta - \theta_0)\cdot t). \tag{2.3}$$

□

PROOF. Let θ_0 and θ be any elements of Θ and define a measure Q_θ by

$$dQ_\theta = \frac{r(\theta_0)}{r(\theta)}\exp((\theta - \theta_0)\cdot t)\; d(tP_{\theta_0}).$$

(Potentially, Q_θ may depend on θ_0.) For any $\acute{\theta} \in \Theta$ we then have

$$E_{Q_\theta}\exp((\acute{\theta} - \theta)\cdot t) = \frac{r(\theta_0)}{r(\theta)} E_{\theta_0}\exp((\acute{\theta} - \theta_0)\cdot t)$$

and hence, by (2.2),

$$E_{Q_\theta}\exp((\acute{\theta} - \theta)\cdot t) = r(\acute{\theta})/r(\theta).$$

Invoking the uniqueness theorem for Laplace transforms, we find that Q_θ equals tP_θ. □

The content of the following theorem is virtually the same as that of Theorem 2.1 in Bar-Lev (1983). The results were derived independently and the proofs are somewhat different. We wish, however, to acknowledge that our inspiration for the results in the theorem derived from the paper by Bar-Lev and Reiser (1982), that we had occasion to see in manuscript form.

THEOREM 2.1. *Under the exponential model* (1.1), *let $t = (t_1, t_2)$ be a partition of the canonical statistic and consider a k_1-dimensional statistic of the form $H(t_2)$ for some* *function H. Then*

$$t_2 \sim \text{EM}((H(t_2), t_2); \theta) \Leftrightarrow t_1 - H(t_2) \perp t_2 \Leftrightarrow t_1 - H(t_2) \sim \text{EM}(t_1 - H(t_2); \theta_1).$$

In this case the distribution of $t_1 - H(t_2)$ depends on θ_1 only, and, with

$$p(t_2; \theta) = a_0(\theta)\exp(\theta \cdot (H(t_2), t_2)) \tag{2.4}$$

as an exponential representation of the model for t_2, the Laplace transform of $t_1 - H(t_2)$ may be expressed as

$$E_{\theta_1}\exp(\zeta \cdot \{t_1 - H(t_2)\}) = \frac{a_0(\theta + (\zeta, 0))}{a_0(\theta)} \frac{a(\theta)}{a(\theta + (\zeta, 0))}, \tag{2.5}$$

where the right hand side, in fact, depends on θ through θ_1 only. □

PROOF. Using (2.1) we find

$$\begin{aligned} &E_{\theta_0}\{\exp((\theta_1 - \theta_{01}) \cdot (t_1 - H(t_2))) \mid t_2\} \\ &= \left[\frac{a(\theta)}{a(\theta_0)} \exp((\theta_1 - \theta_{01}) \cdot H(t_2) + (\theta_2 - \theta_{02}) \cdot t_2)\right]^{-1} \frac{p(t_2; \theta)}{p(t_2; \theta_0)}. \end{aligned} \tag{2.6}$$

The left hand side of (2.6) is the Laplace transform of $t_1 - H(t_2)$ given t_2. Hence $t_1 - H(t_2) \perp t_2$ if and only if the right hand side of (2.6) does not depend on t_2 which is the same as $t_2 \sim \text{EM}((H(t_2), t_2); \theta)$. The latter relation implies that the right hand side of (2.6) is of the form $r(\theta)/r(\theta_0)$ where

$$r(\theta) = a_0(\theta)/a(\theta) \tag{2.7}$$

and $a_0(\theta)$ is the norming constant in (2.4). Furthermore, since the left hand side of (2.6) does not depend on θ_2 we must have that r is a function of θ_1 only, i.e. the distribution of $t_1 - H(t_2)$ depends on θ_1 only and

$$E_{\theta_{01}}\{\exp((\theta_1 - \theta_{01}) \cdot (t_1 - H(t_2)))\} = \frac{r(\theta_1)}{r(\theta_{01})}. \tag{2.8}$$

The implication $t_2 \sim \text{EM}((H(t_2), t_2); \theta) \Leftrightarrow t_1 - H(t_2) \sim \text{EM}(t_1 - H(t_2); \theta_1)$ now follows by Lemma 2.1, and (2.5) is a consequence of (2.7) and (2.8). It remains only to prove that $t_1 - H(t_2) \sim \text{EM}(t_1 - H(t_2); \theta_1)$ implies $t_1 - H(t_2) \perp t_2$. For brevity, write $w = t_1 - H(t_2)$. By assumption we have that $d(wP_\theta)/dP(wP_{\theta_0})$ is of the form

$$\frac{d(wP_\theta)}{d(wP_{\theta_0})} = \frac{r(\theta_1)}{r(\theta_{01})} \exp((\theta_1 - \theta_{01}) \cdot w) \tag{2.9}$$

for some function r. On the other hand we obtain from (2.1)

$$\frac{d(wP_\theta)}{d(wP_{\theta_0})} = \frac{a(\theta)}{a(\theta_0)} \exp((\theta_1 - \theta_{01}) \cdot w) E_{\theta_0}\{\exp((\theta_1 - \theta_{01}) \cdot H(t_2) + (\theta_2 - \theta_{02}) \cdot t_2) \mid w\}, \tag{2.10}$$

and comparing (2.9) and (2.10) we find that $w \perp t_2$. □

Inspection of the proof shows that, in fact, when $t_2 \sim \text{EM}((H(t_2), t_2), \theta)$ we have

$$E_{\theta_{01}}\exp((\theta_1 - \theta_{01}) \cdot (t_1 - H(t_2))) = \frac{a_0(\theta)}{a_0(\theta_0)} \frac{a(\theta_0)}{a(\theta)} \tag{2.11}$$

for any θ_0 and θ in Θ. This generalizes (2.5). (Note that the right hand side of (2.11) does not depend on θ_{02} and θ_2.)

3. Conditions for reproductivity. Recall that the exponential model $\mathscr{M}$ is said to be reproductive in t_2 if for every $n = 1, 2, \cdots$ we have

$$\bar{t}_2 \sim \text{EM}((H(\bar{t}_2), \bar{t}_2); n\theta) \tag{3.1}$$

for a certain function H which does not depend on n and which takes values in R^{k_1}. In consequence of Theorem 2.1 the relation (3.1) is equivalent to

$$\bar{t}_1 - H(\bar{t}_2) \perp \bar{t}_2 \tag{3.2}$$

and also equivalent to

$$\bar{t}_1 - H(\bar{t}_2) \sim \mathrm{EM}(\bar{t}_1 - H(\bar{t}_2); n\theta_1). \tag{3.3}$$

LEMMA 3.1. *Suppose $\mathscr{M}$ is reproductive, with H continuous. Then τ_1 is of the form*

$$\tau_1 = m(\theta_1) + H(\tau_2) \tag{3.4}$$

for some function m. □

PROOF. For $\theta \in \operatorname{int} \Theta$ and $n \to \infty$ we have $\bar{t} \to \tau = \tau(\theta)$ a.s. and hence, by the assumed continuity of H,

$$\bar{t}_1 - H(\bar{t}_2) \to \tau_1 - H(\tau_2).$$

On the other hand, from Theorem 2.1 we have that the distribution of $\bar{t}_1 - H(\bar{t}_2)$ depends on θ_1 only and therefore $\tau_1 - H(\tau_2)$ must depend on θ through θ_1 only. □

In the terminology of Barndorff-Nielsen and Blæsild (1983), the relation (3.4) means that $\mathscr{M}$ possesses a τ-parallel foliation, cf. Theorem 5.1 of that paper.

We conjecture that (3.4) is, in fact, not only a necessary but also a sufficient condition for reproductivity of $\mathscr{M}$. Although we have been unable to prove this in full generality, the sufficiency is established under certain additional assumptions in Corollary 5.4 of Barndorff-Nielsen and Blæsild (1983).

It follows from (3.4) that m is a one-to-one, continuous function of θ_1 and that m is the gradient of some real valued function M on int Θ_1, where Θ_1 denotes the set of possible values of θ_1. It also follows that H possesses continuous partial derivatives with respect to the coordinates of τ_2. We set

$$h(\tau_2) = \frac{\partial H^*}{\partial \tau_2},$$

and

$$\check{H}(\tau_2) = \tau_2 h^*(\tau_2) - H(\tau_2)$$

(i.e. $\check{H}$ is the Legendre transform of H, cf. Section 2 of Barndorff-Nielsen and Blæsild, 1983).

LEMMA 3.2. *Suppose $\mathscr{M}$ is reproductive, with H continuous. Let*

$$p = t_1 - t_2 h^*(\tau_2) + \check{H}(\tau_2). \tag{3.5}$$

The distribution of p depends on θ_1 only and the Laplace transform of p is of the form

$$\mathrm{E}_{\theta_1} \exp(\zeta \cdot p) = \exp(M(\theta_1 + \zeta) - M(\theta_1)). \tag{3.6}$$ □

These properties hold, in fact, for any steep model (1.1) such that $\tau_1 = m(\theta_1) + H(\tau_2)$, whether the model is reproductive or not, see Section 5 of Barndorff-Nielsen and Blæsild (1983). We draw on this generality below.

Now, for a sample $x_1, \cdots, x_n$ of size n, let us consider the three variates

$$\bar{p} = \bar{t}_1 - \bar{t}_2 h^*(\tau_2) + \check{H}(\tau_2),$$

$$\acute{q} = H(\bar{t}_2) - \bar{t}_2 h^*(\tau_2) + \check{H}(\tau_2)$$

and

$$\acute{w} = \bar{t}_1 - H(\bar{t}_2).$$

We have

$$\acute{p} = \acute{q} + \acute{w} \tag{3.7}$$

and, by (3.2), $\acute{q}$ and $\acute{w}$ are independent. Actually, the relation (3.7) generalizes, as will be further discussed subsequently, the decomposition into independent components of quadratic forms in Gaussian variates.

The statistic $\acute{p}$ is the arithmetic mean of the n values of (3.5) determined by the observations $x_1, \cdots, x_n$, but may also be viewed as definition (3.5) applied to the distribution of $\bar{t}$. In view of Lemma 3.2 the Laplace transform of $\acute{p}$ is therefore

$$E_{\theta_1}\exp(\zeta \cdot \acute{p}) = \exp(n\{M(\theta_1 + n^{-1}\zeta) - M(\theta_1)\}). \tag{3.8}$$

Next, it will be shown that the marginal distribution of $\bar{t}_2$ satisfies a relation of the form (3.4). In this marginal distribution the role of $\bar{t}_1$ is taken over by $H(\bar{t}_2)$ (cf. (3.1)) and, writing

$$\tilde{\tau}_1 = E_\theta H(\bar{t}_2),$$

we have

$$\tilde{\tau}_1 = E_\theta \bar{t}_1 - E_{\theta_1}\acute{w} = m(\theta_1) + H(\tau_2) - E_{\theta_1}\acute{w} = \tilde{m}(n\theta_1) + H(\tau_2)$$

where we have used (3.4) and where

$$\tilde{m}(n\theta_1) = m(\theta_1) - E_{\theta_1}\acute{w}.$$

Since the function H of the relation

$$\tilde{\tau}_1 = \tilde{m}(n\theta_1) + H(\tau_2)$$

is the same as the function H in the relation (3.4), one sees that $\acute{q}$ may be obtained by applying the definition (3.5) to the marginal distribution of $\bar{t}_2$. Hence, using Lemma 3.2 and letting $\tilde{M}$ denote the indefinite integral of $\tilde{m}$, we find

$$E_{\theta_1}\exp(\zeta \cdot \acute{q}) = \exp(\tilde{M}(n\theta_1 + \zeta) - \tilde{M}(n\theta_1)). \tag{3.9}$$

(It should be noted that the distribution of $\acute{q}$ depends on n but that we have partly suppressed this dependency in the notations.) Thus each of the three statistics in (3.7) has a distribution which depends on θ_1 only and due to the independence of $\acute{q}$ and $\acute{w}$ we see from (3.8) and (3.9) that

$$E_{\theta_1}\exp(\zeta \cdot \acute{w}) = \exp(n\{M(\theta_1 + n^{-1}\zeta) - M(\theta_1)\} - \{\tilde{M}(n\theta_1 + \zeta) - \tilde{M}(n\theta_1)\}). \tag{3.10}$$

It is immediate from the definitions that the (strong) reproductivity of (1.1) implies (strong) reproductivity in t_2 of the marginal model for t_2. Suppose now that (1.1) is itself the marginal model for t_2, i.e. $t_2(x) = x$ and (1.1) is of the form

$$a(\theta)b(x)\exp(\theta_1 \cdot H(x) + \theta_2 \cdot x). \tag{3.11}$$

If the reproductivity of (3.11) is strong then $\tilde{m}$ does not depend on n. Furthermore,

$$E_{\theta_1}\exp(\zeta \cdot \acute{q}) = \exp(M(n\theta_1 + \zeta) - M(n\theta_1))$$

and

$$E_{\theta_1}\exp(\zeta \cdot \acute{w}) = \exp(n\{M(\theta_1 + n^{-1}\zeta) - M(\theta_1)\} - \{M(n\theta_1 + \zeta) - M(n\theta_1)\}),$$

i.e. we may substitute M for $\tilde{M}$ in (3.9) and (3.10). Furthermore, in this case we have

$$\acute{w} = n^{-1}\textstyle\sum_{i=1}^n H(x_i) - H(\bar{x}).$$

We collect these results in:

Theorem 3.1. *Let the model $\mathscr{M}$ be reproductive in t_2, with H continuous. The variate*

$$\bar{p} = \bar{t}_1 - \bar{t}_2 h^*(\tau_2) + \check{H}(\tau_2)$$

is decomposable as

$$\bar{p} = \acute{q} + \acute{w} \tag{3.12}$$

where

$$\acute{q} = H(\bar{t}_2) - \bar{t}_2 h^*(\tau_2) + \check{H}(\tau_2), \quad \acute{w} = \bar{t}_1 - H(\bar{t}_2),$$

and $\acute{q}$ and $\acute{w}$ are independent. Furthermore, the distributions of $\bar{p}$, $\acute{q}$ and $\acute{w}$ depend on θ_1 only, and the Laplace transforms of $\bar{p}$, $\acute{q}$ and $\acute{w}$ are given by (3.8), (3.9) *and* (3.10).

In case $t_2(x) = x$ and $\mathscr{M}$ is strongly reproductive, we have $\acute{w} = n^{-1}\Sigma H(x_i) - H(\bar{x})$, and the factorisation of the Laplace transform of $\bar{p}$ corresponding to (3.12) *is of the form*

$$\begin{aligned} &\exp(n\{M(\theta_1 + n^{-1}\zeta) - M(\theta_1)\}) \\ &\quad = \exp(M(n\theta_1 + \zeta) - M(n\theta_1)) \\ &\qquad \cdot \exp(n\{M(\theta_1 + n^{-1}\zeta) - M(\theta_1)\} - \{M(n\theta_1 + \zeta) - M(n\theta_1)\}). \ \square \end{aligned} \tag{3.13}$$

Examples 1.2–1.4 are of the type discussed in the second part of the theorem and for these the decomposition (3.12) turns out as

$$n^{-1}\Sigma(x_i - \xi)^2 = (\bar{x} - \xi)^2 + n^{-1}\Sigma(x_i - \bar{x})^2,$$

$$-\ln(\tilde{x}\alpha/\lambda) + \bar{x}\alpha/\lambda - 1 = \{-\ln(\bar{x}\alpha/\lambda) + \bar{x}\alpha/\lambda - 1\} + \ln(\bar{x}/\tilde{x})$$

$$n^{-1}\Sigma(x_i^{-1/2} - \sqrt{\psi/\chi}x_i^{1/2})^2 = (\bar{x}^{-1/2} - \sqrt{\psi/\chi}\bar{x}^{1/2})^2 + n^{-1}\Sigma(x_i^{-1} - \bar{x}^{-1})$$

respectively. A further illustration of the decomposition appears in Example 4.1.

We add a remark on a relation between maximum likelihood estimation of the mixed parameter (θ_1, τ_2) and reproductivity. As a mild regularity condition, suppose there exists an n_0 such that $\bar{t} \in \text{int } C$ with probability 1 for $n \geq n_0$; here C denotes the closed convex hull of the marginal distribution of t. The maximum likelihood estimate $\hat{\theta}$ of θ exists uniquely and satisfies $\hat{\tau} = \tau(\hat{\theta}) = \bar{t}$, provided $\bar{t} \in \text{int } C$. Hence, in view of (3.4) we have, for $n \geq n_0$,

$$m(\hat{\theta}_1) = \bar{t}_1 - H(\bar{t}_2)$$

and since m is injective we find from Theorem 2.1 that

$$\hat{\theta}_1 \perp \hat{\tau}_2 \tag{3.14}$$

and that the distribution of $\hat{\theta}_1$ depends on θ_1 only. It can be shown that, on the other hand, (3.14) implies $\bar{t}_2 \sim \text{EM}((H(\bar{t}_2), \bar{t}_2); n\theta)$; see Barndorff-Nielsen and Blæsild (1983).

Finally, we present the criterion for strong reproductivity commented on in Section 1.

Theorem 3.2. *Suppose t_1 and t_2 are of the form $t_1(x) = H(x)$ for some continuous vector-valued function H and $t_2(x) = x$, respectively. Then $\mathscr{M}$ is strongly reproductive if and only if c int $\Theta \subset$ int Θ for every scalar $c > 1$ and θ_2 is of the form*

$$\theta_2 = -\theta_1 h(\tau_2)$$

for some $k_1 \times k_2$ matrix-valued function h. (*It turns out that $h = \partial H^*/\partial\tau_2$.*) $\square$

Proof. The sufficiency of the condition follows from Corollary 5.4 of Barndorff-Nielsen and Blæsild (1983), while the necessity is a consequence of Theorem 5.6 of that same paper and of Lemma 3.1 above. $\square$

Combining this result with formula (5.4) in Theorem 5.1 of Barndorff-Nielsen and Blæsild (1983) one finds that, under the conditions of Theorem 3.2, the representation (1.1) takes the form

$$b(x)\exp(-M(\theta_1))\exp(-\theta_1 \cdot \{xh^*(\tau_2) - \check{H}(\tau_2) - H(x)\}). \tag{3.15}$$

For θ_1 one-dimensional, models of this kind may be used as elements within a multivariate extension of Nelder and Wedderburn's generalised linear models. More specifically, the latter models (Nelder and Wedderburn, 1972, see also Nelder, 1975) are for independent univariate observations $x_1, \cdots, x_n$ with model function for the single observation x_i of the form

$$b(x; \phi)\exp(-\phi\{\gamma_i x - \kappa(\gamma_i)\}), \tag{3.16}$$

the parameters γ_i being related by $\mu_i = f(\Sigma z_{ij}\beta_j)$, where $\mu_i = \kappa'(\gamma_i)$ is the mean value of x_i, the z_{ij} are known covariates, f is a known so-called "link function", and the β_j, $j = 1, \cdots, d$, are unknown parameters. Besides the exponential form of (3.16), it is an important feature of these models that when ϕ is considered as known, the log likelihood is proportional to ϕ. This implies, in particular, that the maximum likelihood estimate of $(\beta_1, \cdots, \beta_d)$ is the same whatever the value of ϕ. These properties would also hold if the x_i were multidimensional, each distributed as in (3.15) with θ_1 one-dimensional and independent of i and with mean value $\tau_{2i} = Ex_i = f(\Sigma z_{ij}\beta_j)$.

4. Examples of reproductive exponential families. Some instances of reproductive exponential families were given in Section 1. Here we wish to indicate a method for constructing new, higher dimensional examples from examples already known. We begin with a case of some particular independent interest.

EXAMPLE 4.1. *Inverse Gaussian—Gaussian.* Let u be the first passage time to level $c > 0$ of a Brownian motion starting at 0 and with diffusion coefficient ω and drift $\mu \geq 0$. Further, consider another, independent Brownian motion $y(\cdot)$ which also starts at 0 and which has a diffusion coefficient σ and a drift coefficient ξ, and define $v = y(u)$, i.e. v is the value of $y(\cdot)$ at the moment the first process reaches level c. Then $u \sim N^-(\chi, \psi)$ and $v \mid u \sim N(u\xi, u\sigma^2)$ where $\chi = c^2/\omega^2$ and $\psi = \mu^2/\omega^2$. The joint probability density function of u and v may therefore be written

$$\begin{aligned} &p(u, v; \chi, \kappa, \alpha, \beta) \\ &= \frac{1}{2\pi} \sqrt{\chi\kappa} \exp(\sqrt{\chi(\alpha - \beta^2/\kappa)})u^{-2}\exp(-\chi u^{-1}/2 - \kappa u^{-1}v^2/2 - \alpha u/2 + \beta v) \end{aligned} \tag{4.1}$$

where

$$\chi = c^2/\omega^2, \quad \kappa = 1/\sigma^2, \quad \alpha = \mu^2/\omega^2 + \xi^2/\sigma^2, \quad \beta = \xi/\sigma^2.$$

The distributions (4.1) obviously constitute an exponential model of order 4 and with

$$\begin{aligned} \theta &= (-\tfrac{1}{2}\chi, -\tfrac{1}{2}\kappa, -\tfrac{1}{2}\alpha, \beta), \\ t &= (u^{-1}, u^{-1}v^2, u, v), \\ \tau &= (\omega^2/c^2 + \mu/c, \sigma^2 + c\xi^2/\mu, c/\mu, c\xi/\mu) \end{aligned}$$

and

$$a(\theta) = \sqrt{\chi\kappa} \exp(\sqrt{\chi(\alpha - \beta^2/\kappa)}). \tag{4.2}$$

It follows at once from (4.2) that the Laplace transform of $t_2 = (u, v)$ is

$$E_\theta \exp(\delta u + \varepsilon v) = \exp(\sqrt{\chi\{(\alpha - 2\delta) - (\beta + \varepsilon)^2/\kappa\}} - \sqrt{\chi(\alpha - \beta^2/\kappa)}) \tag{4.3}$$

and hence $\bar{t}_2 = (\bar{u}, \bar{v})$ has Laplace transform

(4.4) $$E_\theta \exp(\delta\bar{u} + \varepsilon\bar{v}) = \exp(\sqrt{n\chi\{(n\alpha - 2\delta) - (n\beta + \varepsilon)^2/(n\kappa)\}} - \sqrt{n\chi(n\alpha - (n\beta)^2/(n\kappa))}).$$

Comparing (4.3) and (4.4) we find that (4.1) is strongly reproductive in (u, v). In other words, if we denote by $[N^-, N](\chi, \kappa, \alpha, \beta)$ the distribution (4.1) for (u, v) we have

$$(\bar{u}, \bar{v}) \sim [N^-, N](n\chi, n\kappa, n\alpha, n\beta)$$

or, equivalently,

$$\bar{u} \sim N^-(n\chi, n\psi), \quad \bar{v} \mid \bar{u} \sim N(\bar{u}\xi, \bar{u}\sigma^2/n)$$

for every $n = 1, 2, \cdots$. The corresponding function H is

$$H(u, v) = (u^{-1}, u^{-1}v^2)$$

and by (3.2) we find that the statistic

$$\acute{w} = (n^{-1}\Sigma u_i^{-1} - \bar{u}^{-1}, n^{-1}\Sigma u_i^{-1}v_i^2 - \bar{u}^{-1}\bar{v}^2)$$

is independent of $(\bar{u}, \bar{v})$. The Laplace transform of the distribution of $\acute{w}$ is derivable from (3.13) or by direct calculation, using well known properties of the inverse Gaussian distribution and the Gaussian distribution. It turns out that not only is $\acute{w}$ independent of $(\bar{u}, \bar{v})$ but, writing

$$r = n^{-1}\Sigma u_i^{-1} - \bar{u}^{-1}, \quad s = n^{-1}\Sigma u_i^{-1}v_i^2 - \bar{u}^{-1}\bar{v}^2,$$

we have that for $n > 1$ the three statistics r, s, and $(\bar{u}, \bar{v})$ are independent and

(4.5) $$r \sim \Gamma((n-1)/2, n\chi/2), \quad s \sim \Gamma((n-1)/2, n\kappa/2).$$

Moreover,

$$\hat{\chi} = r^{-1}, \quad \hat{\kappa} = s^{-1}, \quad \hat{\mu} = c/\bar{u}, \quad \hat{\xi} = \bar{v}/\bar{u}.$$

We also note that in the decomposition $\bar{p} = \acute{q} + \acute{w}$ of Theorem 3.1 we have

$$\acute{q} = \bar{u}^{-1}((1 - (\mu/c)\bar{u})^2, (\bar{v} - \xi\bar{u})^2)$$

and that the two coordinates of $\acute{q}$ are independent and satisfy

$$\bar{u}^{-1}(1 - (\mu/c)\bar{u})^2 \sim \Gamma(1/2, n\chi/2), \quad \bar{u}^{-1}(\bar{v} - \xi\bar{u})^2 \sim \Gamma(1/2, n\kappa/2).$$

Though it is somewhat incidental to the theme of this paper, we wish to point out the existence of some natural exact tests for the $[N^-, N]$ model. Firstly, the independence of r and s together with the distribution results (4.5) shows that the hypothesis of identical diffusion coefficients, i.e. $\omega^2 = \sigma^2$ or equivalently $c^{-2}\chi = \kappa$, is testable by an exact F-test. Assuming $\omega^2 = \sigma^2$, we may proceed to test identity of the drift coefficients μ and ξ. From the properties of $\acute{q}$ noted above we find, writing now $\acute{q}_0$ instead of $\acute{q}$ to indicate the hypothesis $\omega^2 = \sigma^2$, that

$$\acute{q}_0 = c^2\bar{u}^{-1}(1 - (\mu/c)\bar{u})^2 + \bar{u}^{-1}(\bar{v} - \xi\bar{u})^2 = \bar{u}^{-1}\{(c - \mu\bar{u})^2 + (\bar{v} - \xi\bar{u})^2\}$$

follows a $\Gamma(1, n/(2\sigma^2))$-distribution. Supposing that $\xi = \mu$ we may rewrite $\acute{q}_0$ as

(4.6) $$\acute{q}_0 = (2\bar{u})^{-1}(\bar{v} - c)^2 + 2\bar{u}(\tilde{\xi} - \xi)^2$$

where

$$\tilde{\xi} = (c + \bar{v})/(2\bar{u}).$$

A direct calculation, by means of Laplace transforms, shows that the two terms on the right hand side of (4.6) are independent, each having a $\Gamma(1/2, n/(2\sigma^2))$-distribution. Further, we have from above that these two terms are independent of r and s and that

$$c^2 r + s \sim \Gamma(n - 1, n/(2\sigma^2)).$$

Thus the quotient

$$(2\bar{u})^{-1}(\bar{v}-c)^2/(c^2r+s) \tag{4.7}$$

yields an F-test on 1 and $2n-2$ degrees of freedom for the hypothesis $\xi=\mu$, while if this hypothesis is adopted, hypotheses on ξ may be tested by F-tests, with degrees of freedom 1 and $2n-1$, based on the quotient

$$2\bar{u}(\tilde{\xi}-\xi)^2/\{c^2r+s+(2\bar{u})^{-1}(\bar{v}-c)^2\}.$$

The mean value of the numerator in (4.7) is

$$E\{(2\bar{u})^{-1}(\bar{v}-c)^2\}=\sigma^2/n+\tfrac{1}{2}(c/\mu)(\xi-\mu)^2$$

and this is always greater than σ^2/n unless $\xi=\mu$.

(Suppose $\mu>0$ and let $\rho=\xi/\mu$. The hypothesis $\xi=\mu$ can then be rephrased as $\rho=1$ and one may ask whether the exact F-test for $\rho=1$ generalises immediately to any hypothesis of the form $\rho=\rho_0$. This is not the case. While it is still possible to decompose $\acute{q}_0$ into two quadratic terms, as

$$\acute{q}_0=\{(1+\rho_0^{-2})\bar{u}\}^{-1}(\bar{v}/\rho_0-c)^2+\{(1+\rho_0^{-2})\bar{u}\}(\tilde{\xi}-\xi)^2,$$

the first term no longer follows a (gamma-) distribution with σ as scale parameter.) □

A related example with similar properties is obtained by observing the time when the second Brownian motion $y(\cdot)$, of the previous example, first reaches level uc', rather than observing $y(u)$. Here c' is some positive constant and ξ is assumed to be nonnegative. Letting now v denote this second first passage time we arrive, in effect, at the following situation.

EXAMPLE 4.2. *Inverse Gaussian—inverse Gaussian.* Let $u\sim N^-(\chi,\psi)$ and $v\,|\,u\sim N^-(u^2\kappa,\lambda)$, where the four parameters satisfy $\chi>0$, $\psi\geq 0$, $\kappa>0$, $\lambda\geq 0$. (Expressed in terms of the two Brownian motions referred to above these parameters are $\chi=c^2/\omega^2$, $\psi=\mu^2/\omega^2$, $\kappa=c'^2/\sigma^2$ and $\lambda=\xi^2/\sigma^2$.) Setting

$$\alpha=\psi-2\sqrt{\kappa\lambda} \tag{4.8}$$

the probability density function of (u,v) may be written as

$$\begin{aligned} &p(u,v;\chi,\psi,\kappa,\lambda)\\ &\quad=\sqrt{\chi\kappa}\exp(\sqrt{\chi}\sqrt{\alpha+2\sqrt{\kappa\lambda}})b(u,v)\exp(-\{\chi u^{-1}+\kappa u^2v^{-1}+\alpha u+\lambda v\}/2)\end{aligned} \tag{4.9}$$

where $b(u,v)=(2\pi)^{-1}u^{-1/2}v^{-3/2}$. Thus (u,v) follows an exponential model of order 4 with

$$\theta=-\tfrac{1}{2}(\chi,\kappa,\alpha,\lambda),\quad t=(u^{-1},u^2v^{-1},u,v) \tag{4.10}$$

and

$$\tau=(\chi^{-1}+\sqrt{\psi/\chi},\kappa^{-1}+\sqrt{(\chi\lambda)/(\psi\kappa)},\sqrt{\chi/\psi},\sqrt{(\chi\kappa)/(\psi\lambda)}).$$

Let the distribution (4.9) be denoted by $[N^-,N^-](\chi,\psi,\kappa,\lambda)$ and set

$$r=n^{-1}\Sigma u_i^{-1}-\bar{u}^{-1},\quad s=n^{-1}\Sigma u_i^2v_i^{-1}-\bar{u}^2\bar{v}^{-1}.$$

Then, as observed in Barndorff-Nielsen (1983), the following properties hold. The three statistics r, s and $(\bar{u},\bar{v})$ are independent and

$$\begin{aligned}(\bar{u},\bar{v})&\sim[N^-,N^-](n\chi,n\psi,n\kappa,n\lambda),\\ r\sim\Gamma((n-1)/2,n\chi/2),&\quad s\sim\Gamma((n-1)/2,n\kappa/2).\end{aligned} \tag{4.11}$$

Moreover, $\hat{\chi}=r^{-1}$, $\hat{\psi}=(r\bar{u}^2)^{-1}$, $\hat{\kappa}=s^{-1}$ and $\hat{\lambda}=\bar{u}^2/(s\bar{v}^2)$.

Comparing (4.11), (4.9) and (4.8) we find that (4.9) is strongly reproductive in (u,v). The corresponding H-function is $H(u,v)=(u^{-1},u^2v^{-1})$, and $\acute{w}=(r,s)$. □

As is well known, the inverse Gaussian distribution allows of "hierarchical analysis of variance" in complete analogy with such analysis for normal variates, cf. Tweedie (1957). From the properties of $[N^-, N]$ and $[N^-, N^-]$, as discussed in Examples 4.1 and 4.2, it is seen that these two models similarly provide analysis of variance procedures. We refrain from discussing any details of this.

Next we present an example of a model which is reproductive without being strongly reproductive.

EXAMPLE 4.3. *Inverse Gaussian—gamma.* Suppose $u \sim N^-(\chi, \psi)$ and $v|u \sim \Gamma(uc, \lambda)$ where c is a known positive constant. The joint probability density function of u and v is

$$p(u, v; \chi, \psi, \lambda) = \frac{\sqrt{\chi}}{\sqrt{2\pi}} \exp(\sqrt{\chi\psi}) u^{-3/2} v^{cu-1} \frac{1}{\Gamma(cu)} \exp(-\chi u^{-1}/2 - \alpha u - \lambda v) \tag{4.12}$$

where

$$\alpha = \tfrac{1}{2}\psi - c \ln \lambda.$$

Again we have an exponential model for (u, v), this time of order 3 and with

$$\theta = -(\tfrac{1}{2}\chi, \alpha, \lambda), \quad t = (u^{-1}, u, v)$$

and

$$\tau = (\chi^{-1} + \sqrt{\psi/\chi}, \sqrt{\chi/\psi}, (c/\lambda)\sqrt{\chi/\psi}).$$

From the form of the cumulant transform

$$\kappa(\theta) = -\tfrac{1}{2} \ln \chi - \sqrt{2\chi(\alpha + c \ln \lambda)}$$

it appears that if we denote the distribution of (u, v) by $[N^-, \Gamma](\chi, \alpha, \lambda; c)$ then

$$(\bar{u}, \bar{v}) \sim [N^-, \Gamma](n\chi, n(\alpha - c \ln n), n\lambda; nc),$$

so that this model is reproductive in (u, v). The model is however not strongly reproductive since the transfer from (u, v) to $(\bar{u}, \bar{v})$ is accompanied by a change of index, from c to nc. □

Other examples of reproductive models that are not strongly reproductive are obtainable by assuming that part of the canonical parameter of a strongly reproductive model is known. For instance, if for $[N^-, N](\chi, \kappa, \alpha, \beta)$ the parameter κ is taken as known then one

TABLE 1

Reproductive exponential models for two-dimensional observations (u, v), obtained by suitably coupling two of the models N, Γ, N^-. The models are reproductive in (u, v). (c denotes a known positive constant.)

reproductive model	definition		strongly reproductive
	$u \sim$	$v/u \sim$	
$[N^-, N]$	$N^-(\chi, \psi)$	$N(u\xi, u\sigma^2)$	+
$[N, N^-]$	$N(\xi, \sigma^2)$	$N^-(u^2\kappa, \lambda)$	+
$[N^-, N^-]$	$N^-(\chi, \psi)$	$N^-(u^2\kappa, \lambda)$	+
$[\Gamma, N]$	$\Gamma(\lambda, \alpha)$	$N(u\xi, u\sigma^2)$	+
$[\Gamma, N^-]$	$\Gamma(\lambda, \alpha)$	$N^-(u^2\kappa, \beta)$	+
$[\Gamma, \Gamma]$	$\Gamma(\lambda, \alpha)$	$\Gamma(uc, \beta)$	−
$[N^-, \Gamma]$	$N^-(\chi, \psi)$	$\Gamma(uc, \beta)$	−

has a model which is reproductive, but not strongly reproductive, in (u, v). The model $[N^-, \Gamma]$ of Example 4.3 is, however, not of this type since (4.12) is no longer exponential if c is considered as unknown.

There are further ways of coupling two of the distributions N, Γ and N^- so as to obtain reproductive exponential models, and Table 1 indicates the various possibilities (ignoring the obvious $[N, N]$ combination). It may be noted that the model considered in Example 1.1 is the submodel of the model $[\Gamma, N^-]$ determined by $\lambda = \kappa = 1$. This illustrates the fact that a full exponential submodel of a strongly reproductive model need not be reproductive.

Higher order examples may be constructed by further coupling. Suppose, for instance, that $(u, v) \sim [N^-, N^-](\chi, \psi, \kappa, \lambda)$, and that $w \mid (u, v) \sim N^-(u^2\mu, \nu)$ or $w \mid (u, v) \sim N^-(v^2\mu, \nu)$ or $w \mid (u, v) \sim N^-((u^2 + v^2)\mu, \nu)$. Each of these three combinations gives rise to an exponential model which is strongly reproductive in (u, v, w).

While it is thus possible to construct a considerable variety of reproductive exponential models, a general, explicit description of the mathematical form of these models is wanting.

Acknowledgment. We are indebted to A. Holst Andersen for a critical reading of the manuscript.

REFERENCES

BAR-LEV, S. K. (1983). A characterization of certain statistics in exponential models whose distributions depend on a sub-vector of parameters only. *Ann. Statist.* **11** 746–752.

BAR-LEV, S. K. and REISER, B. (1982). An exponential subfamily which admits UMPU tests based on a single test statistic. *Ann. Statist.* **10** 979–989.

BARNDORFF-NIELSEN, O. (1978). *Information and Exponential Families.* Wiley, Chichester.

BARNDORFF-NIELSEN, O. (1983). On a formula for the distribution of the maximum likelihood estimator. *Biometrika* **70** 343–365.

BARNDORFF-NIELSEN, O. and BLÆSILD, P. (1983). Exponential models with affine dual foliations. *Ann. Statist.* **11** 753–769.

NELDER, J. A. (1975). Announcement by the Working Party on Statistical Computing—GLIM (Generalized Linear Interactive Modelling Program). *Appl. Statist.* **24** 259–261.

NELDER, J. A. and WEDDERBURN, R. W. M. (1972). Generalized linear models. *J. Roy. Statist. Soc. Ser. A* **135** 370–384.

PATIL, G. P. (1965). On multivariate generalized power series distributions and its application to the multinomial and negative multinomial. In G. P. Patil (Ed.): *Classical and Contagious Discrete Distributions.* Statistical Publ. Soc., Calcutta, and Pergamon Press, New York. 183–194.

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

AFDELING FOR TEORETISK STATISTIK
MATEMATISK INSTITUT
AARHUS UNIVERSITET
DK-8000 AARHUS C
DENMARK