Sankhyā : The Indian Journal of Statistics
Special issue in memory of D. Basu
2002, Volume 64, Series A, Pt. 3, pp 509-531

BASU'S THEOREM WITH APPLICATIONS: A PERSONALISTIC REVIEW

By MALAY GHOSH
University of Florida, Florida, USA

SUMMARY. The paper revisits Basu's Theorem, and documents some of its many applications in statistical inference. There is also some discussion on the relationship between the concepts of sufficiency, completeness and ancillarity.

1. Introduction

There are not many simple results in statistics whose popularity has stood the test of time. Basu's Theorem is a clear exception. The result not only remains popular nearly fifty years after its discovery, but the zest and enthusiasm about the theorem that was evidenced in the fifties and sixties has not diminished with the passage of time. Indeed, the theorem makes eminent appearance in advanced texts such as Lehmann (1986) and Lehmann and Casella (1998) as well as intermediate texts such as Mukhopadhyay (2000) and Casella and Berger (2001).

At first sight, Basu's Theorem appears as a purely technical result, and it is possibly the technical aspect of the theorem that has led to its many diverse applications. But the result also has its conceptual merits. As pointed out by Basu (1982) himself, the theorem shows connection between sufficiency, ancillarity and independence, concepts that were previously perceived as unrelated, and has subsequently led to deeper understanding of the interrelationship between these concepts.

I first came across Basu's Theorem in my Ph.D. inference course in Chapel Hill. I was immediately struck by the sheer elegance of the result, though, admittedly at that point, I did not see its full potentiality for applications.

Paper received December 2001.
AMS (2000) *subject classification.* 62B05, 62A01.
Keywords and phrases. Sufficiency, completeness, ancillarity, empirical Bayes, infinite divisibility.

Over the years, while teaching inference courses for graduate students, I could realize more and more the utility of the result. The theorem is not just useful for what it says, it can be used in a wide range of applications such as in distribution theory, hypothesis testing, theory of estimation, calculation of ratios of moments of many complicated statistics, calculation of mean squared errors of empirical Bayes estimators, and even surprisingly, establishing infinite divisibility of certain distributions. The application possibly extends to many other areas of statistics which I have not come across.

What I attempt in this article is a personalistic review of Basu's Theorem and many of its applications. The review is, by no means, comprehensive and has possibly left out many important contributions by different authors. For this, I offer my sincerest apology at the outset.

The outline of the remaining sections is as follows. Section 2 introduces Basu's Theorem and its implications. This section also discusses possible converses of this theorem (Koehn and Thomas, 1975; Lehmann, 1981). Section 3 discusses some immediate simple applications of Basu's Theorem in distribution theory as well as in some moment calculations. Section 4 discusses applications in classical hypothesis testing. Section 5 shows some uses in finding uniformly minimum variance unbiased estimators (UMVUE's) and best equivariant estimators. Section 6 contains applications in the calculation of mean squared errors of empirical Bayes estimators. Finally, Section 7 shows its use in proving some infinite divisibility theorems. Those interested only in applications of the theorem may skip the remainder of Section 2 after Theorem 1, and proceed straight to Section 3.

While writing this review article, I have benefitted considerably by Basu's (1982) own writing on the topic in the Encyclopedia of Statistical Sciences, the illuminating article of Lehmann (1981), and the very elegant contribution of Boos and Hughes-Oliver (1998) which deals with many interesting applications, some of which are repeated in this paper. I enjoyed reading many other interesting contributions, too numerous to mention each one individually.

2. Basu's Theorem and its Role in Statistical Inference

Consider a random variable X on some measurable space $\mathcal{X}$. Typically, $\mathcal{X} = \mathcal{R}^n$, the n-dimensional Euclidean space for some n. The distribution of X is indexed by $\theta \in \Theta$. Let $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ denote the family of distributions of X. A statistic $T \equiv T(X)$ is said to be *sufficient* for θ, (or more appropriately for the family $\mathcal{P}$) if the conditional distribution of

X given T does not depend on θ for every P_θ ($\theta \in \Theta$). A statistic T is *minimal sufficient* if T is sufficient, and no other function of T is sufficient unless it is a one-to-one function of T. A statistic $U \equiv U(X)$ is *ancillary* if the distribution of U does not depend on θ for every P_θ ($\theta \in \Theta$). A statistic $S \equiv S(X)$ is said to be *complete* if for every real-valued function g, $E_\theta[g(S)] = 0$ for all $\theta \in \Theta$ implies $P_\theta[g(S) = 0] = 1$ for all $\theta \in \Theta$. The statistic S is said to be *boundedly complete* if the same implication holds only for all bounded g. Thus, completeness implies bounded completeness, but the converse is not necessarily true. Without sufficiency, the property of completeness, by itself, has very little value. Hence, from now on, we will confine our discussion only to complete or boundedly complete sufficient statistics.

THEOREM 1. (Basu's Theorem). *If T is boundedly complete sufficient for $\mathcal{P}$, and U is ancillary, then T and U are independently distributed (conditionally) on every $\theta \in \Theta$.*

PROOF. The proof is well-known, but is included here for the sake of completeness. To emphasize conditionality on θ, we will write $P(\cdot|\theta)$ instead of $P_\theta(\cdot)$. Since U is ancillary and T is sufficient, writing $p = P(U \leq u) = P(U \leq u|\theta)$ for all $\theta \in \Theta$, and $g(t) = P(U \leq u|T = t) = E[I_{[U \leq u]}|T = t]$, where I is the usual indicator function, by the iterated formula of conditional expectations, $E[g(T)|\theta] = p$ for all $\theta \in \Theta$. The bounded completeness of T now implies that $P[g(T) = p|\theta] = 1$ for all $\theta \in \Theta$. Hence, applying once again the iterated formula of conditional expectations,

$$\begin{aligned} P(U \leq u, T \leq t|\theta) &= E[g(T)I_{[T \leq t]}|\theta] = pP(T \leq t|\theta) \\ &= P(U \leq u)P(T \leq t|\theta) \text{ for all } \theta \in \Theta. \end{aligned}$$

This completes the proof of the theorem.

REMARK 1. The original version of Theorem 1, written purely within a frequentist paradigm, does not include the word "conditionally" in its statement. This was later inserted by Basu (1982), presumably to distinguish the frequentist framework from the Bayesian framework.

REMARK 2. Hogg (1953) had found results similar to Basu's Theorem in some special cases.

Lehmann (1981) in an inspiring article points out that the properties of minimality and completeness of a sufficient statistic are of a rather different nature. Virtually, in every statistical problem, the minimal sufficient statistic exists. On the other hand, existence of a complete sufficient statistic implies the existence of a minimal sufficient statistic (Bahadur, 1957). As pointed out by Lehmann (see also Boos and Hughes-Oliver, 1998), existence of a minimal sufficient statistic, by itself, does not guarantee that there does not exist any function of T which is ancillary. Basu's Theorem tells us that if T is complete in addition to being sufficient, then no ancillary statistic other than the constants can be computed from T. Thus, by Basu's Theorem, completeness of a sufficient statistic T characterizes the success of T in separating the informative part of the data from that part, which by itself, carries no information.

The above point is well-illustrated by the following example (see Lehmann, 1981).

Example 1. Let $X_1, \cdots, X_n$ be independent and identically distributed (iid) with joint probability density function (pdf) belonging to the logistic family

$$\mathcal{P} = \left\{ \prod_{i=1}^{n} f_\theta(x_i) : f_\theta(x) = f(x-\theta),\ x \in \mathcal{R}^1, \theta \in \mathcal{R}^1 \right\}, \tag{2.1}$$

where $f(x) = \exp(x)/[1+\exp(x)]^2$. Let $X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$ denote the ordered X_i's. It is well-known (see e.g. Lehmann and Casella, 1998, p.36) that $T = (X_{(1)}, \cdots, X_{(n)})$ is minimal sufficient for $\mathcal{P}$. But clearly T is not boundedly complete as can be verified directly by showing that the expectation of any bounded function of $X_{(n)} - X_{(1)}$ (with the exception of constants) is a nonzero constant not depending on θ, but that the probability that the function is equal to that constant is not 1 (in fact equal to zero). Alternately, one may note that $X_{(n)} - X_{(1)}$ is ancillary, and is not independent of T. On the other hand, if one considers instead the augmented class of all continuous pdf's $\mathcal{P} = \{\prod_{i=1}^n f(x_i) : f \text{ continuous}\}$, then T is indeed complete (see Lehmann, 1986, pp 143-144), and Basu's Theorem asserts that there does not exist any non-constant function of T which is ancillary.

Hence, as pointed out by Lehmann (1981), for the logistic family, (he produces other examples as well), sufficiency has not been successful in "squeezing out" all the ancillary material, while for the augmented family, success takes place by virtue of Basu's Theorem.

An MLE is always a function of the minimal sufficient statistic T, but need not be sufficient by itself. Often in these cases, an MLE, say M, in conjunction with some ancillary statistic U is one-to-one with T. The bounded completeness of T precludes the existence of such a U (other than the constants) by Basu's Theorem. Thus, it makes sense to consider Fisher's recommendation that inference should be based on the conditional distribution of M given U only when T is not boundedly complete. We add parenthetically here that even otherwise, lack of a unique U may prevent a well-defined inferential procedure (see Basu, 1964).

There are several ways to think of possible converses to Basu's Theorem. One natural question is that if T is boundedly complete sufficient, and U is distributed independently of T for every $\theta \in \Theta$, then is U ancillary? The following simple example presented in Koehn and Thomas (1975) shows that this is not the case.

EXAMPLE 2. Let $X \sim \text{uniform}[\theta, \theta+1)$, where $\theta \in \Theta = \{0, \pm 1, \pm 2, \cdots, \}$. Then X has pdf $f_\theta(x) = I_{[[x]=\theta]}$, where $[x]$ denotes the integer part of x. It is easy to check that $[X]$ is complete sufficient for $\theta \in \Theta$, and is also distributed independently of X, but clearly X is not ancillary!

The above apparently trivial example brings out several interesting issues. First, since $P_\theta([X] = \theta) = 1$ for all $\theta \in \Theta$, $[X]$ is degenerate with probability 1. Indeed, in general, a nontrivial statistic cannot be independent of X, because if this were the case, it would be independent of every function of X, and thus independent of itself! However, this example shows also that if there exists a nonempty proper subset $\mathcal{X}_0$ of $\mathcal{X}$, and a nonempty proper subset Θ_0 of Θ such that

$$P_\theta(\mathcal{X}_0) = \begin{cases} 1 & \text{for } \theta \in \Theta_0, \\ 0 & \text{for } \theta \in \Theta - \Theta_0, \end{cases} \tag{2.2}$$

then the converse to Basu's Theorem may fail to hold. In Example 2, $\mathcal{X} = \mathcal{R}^1$, and Θ is the set of all integers. Taking $\Theta_0 = \{\theta_0\}$ and $\mathcal{X}_0 = [\theta_0, \theta_0 + 1)$, one produces a counterexample to a possible converse to Basu's Theorem.

Koehn and Thomas (1975) called such a set $\mathcal{X}_0$ a *splitting set*. It is the existence of these splitting sets which prevents the converse in question to hold. Indeed, the non-existence of *any* such splitting set assures a converse to Basu's Theorem. The following general result is proved in Koehn and Thomas (1975).

THEOREM 2. *Let T be a sufficient statistic for $\mathcal{P}$. Then a statistic U distributed independently of T is ancillary if and only if there does not exist any splitting set as described in (2.2).*

Basu (1958) gave a sufficient condition for the same converse to his theorem. First he defined two probability measures P_θ and $P_{\theta'}$ to be *overlapping* if they do not have disjoint supports. In Example 2, all probability measures P_θ, $\theta \in \Theta$ are non-overlapping. The family $\mathcal{P}$ is said to be *connected* if for every pair $\{\theta, \theta'\}$, $\theta \in \Theta$, $\theta' \in \Theta$, there exist $\theta_1, \cdots, \theta_k$, each belonging to Θ

such that $\theta_1 = \theta$, $\theta_k = \theta'$ and for each i, P_{θ_i} and $P_{\theta_{i+1}}$ overlap. The following theorem is given in Basu (1958).

Theorem 3. *Let $\mathcal{P} = \{P_\theta,\ \theta \in \Theta\}$ be connected, and T be sufficient for $\mathcal{P}$. Then U is ancillary if T and U are (conditionally) independent for every $\theta \in \Theta$.*

It is easy to see that the non-existence of any splitting set implies connectedness. The converse is not necessarily true, although counterexamples are hard to find. The only counterexample provided in Koehn and Thomas (1975) is rather pathological.

It is only the sufficiency and *not* the completeness of T which plays a role in Theorems 2 and 3. An alternative way to think about a possible converse to Basu's Theorem is whether the independence of *all* ancillary statistics with a sufficient statistic T implies that T is boundedly complete. The answer is again NO as Lehmann (1981) produces the following counter example.

Example 3. Let X be a discrete random variable assuming values x with probabilities $p(x)$ as given below:

x	-5	-4	-3	-2	-1	1	2	3	4	5
$p(x)$	$\alpha' p^2 q$	$\alpha' p q^2$	$\frac{1}{2}p^3$	$\frac{1}{2}q^3$	$\gamma' pq$	γpq	$\frac{1}{2}q^3$	$\frac{1}{2}p^3$	$\alpha p q^2$	$\alpha p^2 q$

Here $0 < p = 1 - q < 1$, is the unknown parameter, and α, α', γ, γ' are known positive constants. Also, $\alpha + \gamma = \alpha' + \gamma' = 3/2$. In this example, $|X|$ is minimal sufficient, $P(X > 0) = 1/2$ so that $U = I_{[X>0]}$ is ancillary. However, if $\alpha \neq \alpha'$, then U is not distributed independently of T.

Lehmann (1981) has pointed out very succinctly that this converse to Basu's Theorem fails to hold because ancillarity is a property of the whole distribution of a statistic, while completeness is a property dealing only with expectations. He showed also that correct versions of the converse could be obtained either by replacing ancillarity with the corresponding first order property or completeness with a condition reflecting the whole distribution.

To this end, define a statistic $V \equiv V(X)$ to be first order ancillary if $E_\theta(V)$ does not depend on $\theta \in \Theta$. Then one has a necessary and sufficient condition for Basu's Theorem.

Theorem 4. *A necessary and sufficient condition for a sufficient statistic T to be boundedly complete is that every bounded first order ancillary V is uncorrelated (conditionally) with every bounded real-valued function of T for every $\theta \in \Theta$.*

An alternative approach to obtain a converse is to modify instead the definition of completeness. Quite generally, a sufficient statistic T is said to be $\mathcal{G}$-complete ($\mathcal{G}$ is a class of functions) if for every $g \in \mathcal{G}$, $E_\theta[g(T)] = 0$ for all $\theta \in \Theta$ implies that $P_\theta[g(T) = 0] = 1$ for all $\theta \in \Theta$. Suppose, in particular, $\mathcal{G} = \mathcal{G}_0$, where $\mathcal{G}_0$ is the class of all two- valued functions. Then Lehmann (1981) proved the following theorem.

Theorem 5. *Suppose T is sufficient and every ancillary statistic U is distributed independently of T. Then T is $\mathcal{G}_0$-complete.*

Basu's Theorem implies the independence of T and U when T is boundedly complete and sufficient, while U is ancillary. This, in turn, implies the $\mathcal{G}_0$-completeness of T. However, the same example 3 shows that neither of the reverse implications is true. However, if instead of $\mathcal{G}_0$, one considers $\mathcal{G}_1$ which are conditional expectations of all two-valued functions with respect to a sufficient statistic T, then the following theorem holds.

Theorem 6. *(Lehmann, 1981). A necessary and sufficient condition for a sufficient statistic T to be $\mathcal{G}_1$-complete is that every ancillary statistic U is independent of T (conditionally) for every $\theta \in \Theta$.*

Theorems 4-6, provide conditions under which a sufficient statistic T has some form of completeness (not necessarily bounded completeness) if it is independent of *every* ancillary U. However, Theorem 2 says that ancillarity of U does not follow even if it is independent of a complete sufficient statistic. As shown in Example 2, $[X]$ is complete sufficient, and hence, by Basu's Theorem, is independent of every ancillary U, but $[X]$ is independent of X, and X is not ancillary.

The theorems of this section are all stated in a frequentist framework. Basu (1982) points out that if instead one takes a Bayesian point of view and regards θ as a random variable with some prior distribution, then the notions of sufficiency and ancillarity are directly related to the notion of *conditional independence* (cf. Dawid, 1979).

To see this for a model $\mathcal{P}$ and for each prior ξ, consider the joint probability distribution of the pair (X, θ). In this framework, a statistic U is said to be ancillary if its conditional distribution given θ does not depend on θ. Equivalently, U is ancillary if for each joint distribution Q_ξ of (X, θ), the two random variables U and θ are independent. A statistic T is sufficient if the conditional distribution of X given θ and T depends only on T. Thus, an alternative definition of sufficiency of T is that T is sufficient if for each Q_ξ, X and θ are conditionally independent given T. This may be connected directly to the notion of *Bayes sufficiency* due to Kolmogorov (1942). To see this we may note that if T is sufficient in the usual sense, the conditional distribution of θ given x depends only on $T(x)$. Conversely, if the prior pdf of θ is positive for all $\theta \in \Theta$, and the conditional distribution of θ given x depends only on $T(x)$, then T is sufficient for θ in the usual sense. Couched in a Bayesian formulation, Basu's Theorem can be paraphrased as follows.

THEOREM 1A. *(Basu, 1982). Suppose that for each Q_ξ, the variables U and θ are conditionally independent given T. Then U and T are conditionally independent given θ provided that T is boundedly complete in the sense described earlier.*

3. Some Simple Applications of Basu's Theorem

We provide in this section some simple examples where Basu's Theorem leads to important conclusions.

EXAMPLE 4. We begin with the basic result of independence of the sample mean vector and the sample variance-covariance matrix based on a random sample from a multivariate normal distribution.

Let $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n$ $(n \geq p+1)$ be iid $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu} \in \mathcal{R}^p$, and $\boldsymbol{\Sigma}$ is positive definite (p.d.). We define $\bar{\boldsymbol{X}} = n^{-1} \sum_{i=1}^n \boldsymbol{X}_i$ and $\boldsymbol{S} = \sum_{i=1}^n (\boldsymbol{X}_i - \bar{\boldsymbol{X}})(\boldsymbol{X}_i - \bar{\boldsymbol{X}})^T$. Now for every *fixed* p.d. $\boldsymbol{\Sigma}$, $\bar{\boldsymbol{X}}$ is complete sufficient for $\boldsymbol{\mu}$, while $\boldsymbol{S}$ is ancillary. This proves the independence of $\bar{\boldsymbol{X}}$ and $\boldsymbol{S}$ for each $\boldsymbol{\mu} \in \mathcal{R}^p$ and every fixed p.d. $\boldsymbol{\Sigma}$. Equivalently, $\bar{\boldsymbol{X}}$ and $\boldsymbol{S}$ are independently distributed for all $\boldsymbol{\mu} \in \mathcal{R}^p$ and all p.d. $\boldsymbol{\Sigma}$. This implies in particular that $\bar{\boldsymbol{X}}$ and $\boldsymbol{S}$ are independently distributed when $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n$ $(n \geq p+1)$ are iid $N_p(\boldsymbol{0}, \boldsymbol{I}_p)$.

EXAMPLE 5. This example appears in Hogg and Craig (1956). Let $X_1, \cdots, X_n$ be iid $N(\mu, \sigma^2)$, $\mu \in \mathcal{R}^1$, and $\sigma > 0$. Let $\boldsymbol{X} = (X_1, \cdots, X_n)^T$. Then a well-known result is that a quadratic form $\boldsymbol{X}^T \boldsymbol{A} \boldsymbol{X}$ is distributed independently of $\bar{X} = n^{-1} \sum_{i=1}^n X_i$ if and only if $\boldsymbol{A} \boldsymbol{1}_n = \boldsymbol{0}$, where $\boldsymbol{1}_n$ is an n-component column vector with each element equal to 1.

To derive the result from Theorems 1 and 2, first we observe that there does not exist *any* splitting set in this example. Also, for every *fixed* σ^2, $\bar{X}$ is complete sufficient for μ. Hence, for every fixed σ^2, we need to find conditions under which $\boldsymbol{X}^T \boldsymbol{A} \boldsymbol{X}$ is ancillary. To this end, let $\boldsymbol{Y} = \boldsymbol{X} - \mu \boldsymbol{1}_n$. Then

$$\boldsymbol{X}^T \boldsymbol{A} \boldsymbol{X} = \boldsymbol{Y}^T \boldsymbol{A} \boldsymbol{Y} + 2\mu (\boldsymbol{A} \boldsymbol{1}_n)^T \boldsymbol{Y} + \mu^2 (\boldsymbol{A} \boldsymbol{1}_n)^T \boldsymbol{1}_n.$$

Hence, since $\boldsymbol{Y}$ is ancillary for every fixed σ^2, $\boldsymbol{X}^T \boldsymbol{A} \boldsymbol{X}$ is ancillary if and only if $\boldsymbol{A} \boldsymbol{1}_n = 0$.

Example 6. This example appears in Basu (1982). Let $\{X_1, X_2, \cdots\}$ be a sequence of independent gamma variables with shape parameters $\alpha_1, \alpha_2, \cdots$, i.e. X_i has pdf $f_{\alpha_i}(x_i) = [\exp(-x_i)x_i^{\alpha_i-1}/\Gamma(\alpha_i)]I_{[x_i>0]}$. Let $T_i = \sum_{j=1}^i X_j$ and $Y_i = T_i/T_{i+1}$ $(i = 1, 2, \cdots)$. We want to show that for every $n \geq 2$, $Y_1, Y_2, \cdots, Y_{n-1}$ and T_n are mutually independent.

Note that for a fixed n, if $X_1, X_2, \cdots, X_n$ are independent gammas with a common scale parameter σ, and shape parameters $\alpha_1, \alpha_2, \cdots, \alpha_n$, then for *fixed* $\alpha_1, \alpha_2, \cdots, \alpha_n$, T_n is complete sufficient for σ, while $(Y_1, \cdots, Y_{n-1})$ is ancillary. This establishes the independence of $(Y_1, \cdots, Y_{n-1})$ with T_n. Next since $(Y_1, \cdots, Y_{n-2})$ is independent of (T_{n-1}, X_n) and $Y_{n-1} = T_{n-1}/(T_{n-1} + X_n)$, we get independence of $(Y_1, \cdots, Y_{n-2})$ with (Y_{n-1}, T_n). Proceeding inductively, we establish mutual independence of $Y_1, Y_2, \cdots, Y_{n-1}$ and T_n for every $n \geq 2$ when $\alpha_1, \alpha_2, \cdots, \alpha_n$ are fixed. Hence the result holds for all $\alpha_1, \alpha_2, \cdots, \alpha_n$ and every $\sigma(> 0)$, and in particular when $\sigma = 1$.

Example 7. Let $\boldsymbol{X}_i = (X_{1i}, X_{2i})^T$ be n iid random variables, each having a bivariate normal distribution with means $\mu_1(\in \mathcal{R}^1)$ and $\mu_2(\in \mathcal{R}^1)$, variances $\sigma_1^2(> 0)$ and $\sigma_2^2(> 0)$, and correlation $\rho \in (-1, 1)$. Let $\bar{X}_j = n^{-1}\sum_{i=1}^n X_{ji}$, $S_j^2 = \sum_{i=1}^n (X_{ji} - \bar{X}_j)^2$ $(j = 1, 2)$ and $R = \sum_{i=1}^n (X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2)/(S_1S_2)$. Under the null hypothesis $H_0 : \rho = 0$, $(\bar{X}_1, \bar{X}_2, S_1^2, S_2^2)$ is complete sufficient for $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)$, while R is ancillary. Thus $(\bar{X}_1, \bar{X}_2, S_1^2, S_2^2)$ is distributed independently of R when $\rho = 0$. Due to the mutual independence of $\bar{X}_1$, $\bar{X}_2$, S_1^2 and S_2^2 when $\rho = 0$, one gets now the mutual independence of $\bar{X}_1$, $\bar{X}_2$, S_1^2, S_2^2 and R when $\rho = 0$, and the joint pdf of these five statistics is now the product of the marginals. Now to derive the joint pdf $q_{\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho}(\bar{x}_1, \bar{x}_2, s_1^2, s_2^2, r)$ of these five statistics for an arbitrary $\rho \in (-1, 1)$, by the Factorization Theorem of sufficiency, one gets

$$q_{\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho}(\bar{x}_1, \bar{x}_2, s_1^2, s_2^2, r) = q_{0,0,1,1,0}(\bar{x}_1, \bar{x}_2, s_1^2, s_2^2, r)\frac{L(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)}{L(0, 0, 1, 1, 0)},$$

where $L(\cdot)$ denotes the likelihood function under the specified values of the parameters.

Example 8. This example, taken from Boos and Hughes-Oliver (BH) (1998), is referred to as the *Monte Carlo Swindle*. The latter refers to a simulation technique that ensures statistical accuracy with a smaller number of replications at a level which one would normally expect from a much larger number of replications. Johnstone and Velleman (1985) provide many such examples. One of their examples taken by BH shows that if M denotes a sample median in a random sample of size n from a $N(\mu, \sigma^2)$ distribution, then the Monte Carlo estimate of $V(M)$ requires a much smaller sample size to attain a prescribed accuracy, if instead one finds the Monte Carlo estimate of $V(M-\bar{X})$ and adds the usual estimate of σ^2/n to the same.

We do not provide the detailed arguments of BH to demonstrate this. We point out only the basic identity $V(M)=V(M-\bar{X})+V(\bar{X})$ as used by these authors. As noticed by BH, this is a simple consequence of Basu's Theorem. As mentioned in Example 2, for fixed σ^2, $\bar{X}$ is complete sufficient for μ, while $M-\bar{X}=\text{med}(X_1-\mu,\cdots,X_n-\mu)-(\bar{X}-\mu)$ is ancillary. Hence, by Basu's Theorem,

$$V(M)=V(M-\bar{X}+\bar{X})=V(M-\bar{X})+V(\bar{X})=V(M-\bar{X})+\sigma^2/n.$$

One interesting application of Basu's Theorem involves evaluation of moments of ratios of many complicated statistics. The basic result used in these calculations is that if X/Y and Y are independently distributed, then $E[X/Y]^k=E(X^k)/E(Y^k)$ provided the appropriate moments exist.

Example 9. We begin with the evaluation of $E[\bar{X}/X_{(n)}]$, where $\bar{X}$ is the sample average, and $X_{(n)}$ is the largest order statistic in a random sample of size n from the uniform$(0,\theta)$, $\theta\in(0,\infty)$ distribution. Here $X_{(n)}$ is complete sufficient for θ, while $\bar{X}/X_{(n)}$ is ancillary. Thus $\bar{X}/X_{(n)}$ is distributed independently of $X_{(n)}$. Hence, applying the basic formula for ratio of moments,

$$E[\bar{X}/X_{(n)}]=E(\bar{X})/E(X_{(n)})=\frac{\theta/2}{n\theta/(n+1)}=(n+1)/(2n).$$

Example 10. A second example in a similar vein is evaluation of the moments of the studentized range in a random sample of size n from the $N(\mu,\sigma^2)$ distribution. David (1981, p. 89) discusses this example. Let $X_1,\cdots,X_n$ $(n\geq 2)$ denote the random sample, $\bar{X}$ the sample mean, $S^2=(n-1)^{-1}\sum_{i=1}^n(X_i-\bar{X})^2$ the sample variance, and $X_{(1)}$ and $X_{(n)}$ the smallest and largest order statistics. Then the studentized range is defined by $U=(X_{(n)}-X_{(1)})/S$. Applying the basic result, $E(U^k)=E(X_{(n)}-X_{(1)})^k/E(S^k)$, $(k>0)$, the calculations simplify considerably, since one then needs to know only the moments of the sample range and the sample standard deviation. A more familiar example in the same setting is evaluation of the moments of Student's t-statistic $T_n=n^{\frac{1}{2}}\bar{X}/S$ when $\mu=0$.

The final example of this section, quite in the spirit of Examples 9 and 10, shows how Basu's Theorem simplifies the derivation of the distribution of Wilks's Λ-statistic.

EXAMPLE 11. Let $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ be two independent p- dimensional Wishart random variables with common scale parameter $\boldsymbol{\Sigma}$, and degrees of freedom m_1 and m_2, where $\max(m_1, m_2) \geq p$. Without loss of generality, let $m_1 \geq m_2$. Wilk's Λ-statistic is defined by $\Lambda = |\boldsymbol{S}_1|/|\boldsymbol{S}_1 + \boldsymbol{S}_2|$, and is very useful in multivariate regression and multivariate analysis of variance problems. Writing

$$\Lambda = |\boldsymbol{\Sigma}^{-1/2}\boldsymbol{S}_1\boldsymbol{\Sigma}^{-1/2}|/|\boldsymbol{\Sigma}^{-1/2}(\boldsymbol{S}_1 + \boldsymbol{S}_2)\boldsymbol{\Sigma}^{-1/2}|,$$

and the representation of $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ respectively as $\boldsymbol{S}_1 = \sum_{i=1}^{m_1} \boldsymbol{Y}_i\boldsymbol{Y}_i^T$ and $\boldsymbol{S}_2 = \sum_{j=1}^{m_2} \boldsymbol{Z}_j\boldsymbol{Z}_j^T$, where the $\boldsymbol{Y}_i$'s and $\boldsymbol{Z}_j$'s are iid $N_p(\boldsymbol{0}, \boldsymbol{\Sigma})$ random variables, it is immediate that Λ is ancillary, that is, its distribution does not depend on $\boldsymbol{\Sigma}$. Also, it is evident that $\boldsymbol{S}_1 + \boldsymbol{S}_2$ is complete sufficient for $\boldsymbol{\Sigma}$. Hence Λ is distributed independently of $\boldsymbol{S}_1 + \boldsymbol{S}_2$, and hence, it is independent of $|\boldsymbol{S}_1 + \boldsymbol{S}_2|$. Thus, as before,

$$E(\Lambda^k) = E(|\boldsymbol{S}_1|^k)/E(|\boldsymbol{S}_1 + \boldsymbol{S}_2|^k), \tag{3.1}$$

for every positive integer k. Since the distribution of Λ does not depend on $\boldsymbol{\Sigma}$, one computes the moments of the right side of (3.1) under $\boldsymbol{\Sigma} = \boldsymbol{I}_p$. When $\boldsymbol{\Sigma} = \boldsymbol{I}_p$, $|\boldsymbol{S}_1|$ is distributed as $\prod_{i=1}^p W_{i1}$, where W_{i1}'s are independent,and $W_{i1} \sim \chi^2_{m_1-i+1}$. Similarly, $|\boldsymbol{S}_1 + \boldsymbol{S}_2|$ is distributed as $\prod_{i=1}^p W_{i2}$, where the W_{i2}'s are independent and $W_{i2} \sim \chi^2_{m_1+m_2-i+1}$. Now, after some simplifications, it follows that $E(\Lambda^k)$ equals the kth moment of the product of p independent Beta variables for every positive integer k. This leads to the well-known result that Λ is distributed as the product of p independent Beta variables.

4. Basu's Theorem in Hypothesis Testing

Hogg and Craig (1956) have provided several interesting applications of Basu's Theorem. Among these, there are some hypothesis testing examples where Basu's Theorem aids in the derivation of the exact distribution of $-2\log_e\lambda$ under the null hypothesis H_0, λ being the generalized likelihood ratio test (GLRT) statistic. One common feature in all these problems is that the supports of all the distributions depend on parameters. We discuss one of these examples in its full generality.

EXAMPLE 12. Let X_{ij} $(j = 1, \cdots, n_i; i = 1, \cdots, k)$ $(k \geq 2)$ be mutually independent, X_{ij} $(j = 1. \cdots, n_i)$ being iid with common pdf

$$f_{\theta_i}(x_i) = [h(x_i)/H(\theta_i)]I_{[0 \leq x_i \leq \theta_i]},\ i = 1, \cdots, k, \tag{4.1}$$

where $H(u) = \int_0^u h(x)dx$, and $h(x) > 0$ for all $x > 0$. We want to test $H_0 : \theta_1 = \cdots = \theta_k$ against the alternative H_1: not all θ_i are equal. We write $\boldsymbol{X}_i = (X_{i1}, \cdots, X_{in_i})^T$, $i = 1, \cdots, k$ and $\boldsymbol{X} = (\boldsymbol{X}_1^T, \cdots, \boldsymbol{X}_k^T)^T$. Also, let $T_i \equiv T_i(\boldsymbol{X}_i) = \max(X_{i1}, \cdots, X_{in_i})$, $i = 1, \cdots, k$, and $T = \max(T_1, \cdots, T_k)$. The unrestricted MLE's of $\theta_1, \cdots, \theta_k$ are $T_1, \cdots, T_k$. Also, under H_0, the MLE of the common θ_i is T. Then the GLRT statistic for testing H_0 against H_1 simplifies to $\lambda(\boldsymbol{X}) = \prod_{i=1}^k H^{n_i}(T_i)/H^n(T)$, where $n = \sum_{i=1}^k n_i$. Hence,

$$-2\log_e \lambda(\boldsymbol{X}) = \sum_{i=1}^k [-2\log_e\{H^{n_i}(T_i)/H^{n_i}(\theta)\}] - [-2\log_e\{H^n(T)/H^n(\theta)\}], \tag{4.2}$$

where θ denotes the common unknown value of the θ_i's under H_0. It follows from (4.1) that $T_1, \cdots, T_k$ are independent with distribution functions $H^{n_i}(t_i)/H^{n_i}(\theta_i)$, $(i = 1, \cdots, k)$. Hence, under H_0, $H^{n_i}(T_i)/H^{n_i}(\theta)$ are iid uniform(0,1). Accordingly, under H_0,

$$\sum_{i=1}^k [-2\log_e\{H^{n_i}(T_i)/H^{n_i}(\theta)\}] \sim \chi^2_{2k}. \tag{4.3}$$

Also, under H_0, the distribution function of T is $H^n(t)/H^n(\theta)$, and hence, $H^n(T)/H^n(\theta)$ is uniform(0,1) under H_0. Thus, under H_0,

$$-2\log_e[H^n(T)/H^n(\theta)] \sim \chi^2_2. \tag{4.4}$$

So far, we have not used Basu's Theorem. In order to use it, first we observe that under H_0, T is complete sufficient for θ, while λ is ancillary. Hence, under H_0, T is distributed independently of $-2\log_e \lambda$. Also, from (4.2),

$$\sum_{i=1}^k [-2\log_e\{H^{n_i}(T_i)/H^{n_i}(\theta)\}] = [-2\log_e \lambda] + [-2\log_e\{H^n(T)/H^n(\theta)\}]. \tag{4.5}$$

The two components in the right side of (4.5) are independent. Now by (4.3), (4.4) and the result that if W_1 and W_2 are independent with $W_1 \sim \chi^2_m$ and $W_1 + W_2 \sim \chi^2_{m+n}$, then $W_2 \sim \chi^2_n$, one finds that $-2\log_e \lambda \sim \chi^2_{2k-2}$ under H_0.

The above result should be contrasted to the regular case (when the support of the distribution does not depend on parameters) where under some regularity conditions, $-2\log_e \lambda$ is known to have an asymptotic chi-squared distribution. In a similar scenario with n observations and k unknown parameters in general, and 1 under the null, the associated degrees of freedom in the regular case would have been $(n-1)-(n-k)=k-1$ instead of $2(k-1)$.

Basu's theorem also helps finding explicitly many uniformly most powerful unbiased (UMPU) tests. Although the general interest in such tests is on the decline, it is interesting to see how Basu's Theorem is applicable in this context. A very general result to this effect is given in Lehmann (1986, p.191). We illustrate this with a simple example.

EXAMPLE 13. Let $X_1,\cdots,X_n$ $(n\ge 2)$ be iid $N(\mu,\sigma^2)$, where both $\mu\in\mathcal{R}^1$ and $\sigma>0$ are unknown. Define $T_1=\sum_{i=1}^n X_i$ and $T_2=\sum_{i=1}^n X_i^2$. It is shown in Lehmann (1986) that the UMPU test for $H_0:\mu=0$ against $H_1:\mu\neq 0$ is given by

$$\phi(T_1,T_2)=\begin{cases}1 & \text{if } T_1<c_1(T_2) \text{ or } T_1>c_2(T_2),\\ 0 & \text{otherwise,}\end{cases} \tag{4.6}$$

where one requires (i) $E_{0,\sigma^2}[\phi(T_1,T_2)|T_2]=\alpha$ a.e. and (ii) $\text{Cov}_{0,\sigma^2}[T_1,\phi(T_1,T_2)|T_2]=0$ a.e. Now (i) can be alternately expressed as $(i)'$ $P_{0,\sigma^2}(c_1(T_2)\le T_1\le c_2(T_2)|T_2)=1-\alpha$ a.e. Now, we observe that T_2 is complete sufficient under H_0, while $T_1/T_2^{1/2}$ is ancillary. This leads to the independence of $T_1/T_2^{1/2}$ and T_2, and thus the conditional distribution of the former given the latter is the same as its unconditional distribution. Writing $U=T_1/T_2^{1/2}$, the requirements (i) and (ii) now simplify to (iii) $P(d_1\le U\le d_2)=1-\alpha$ and (iv) $\text{Cov}[U,I_{[d_1\le U\le d_2]}]=0$ for some constants d_1 and d_2 not depending on the data. Since $E(U)=0$ under H_0, (iv) simplifies further to $\int_{d_1}^{d_2}uh(u)du=0$, where $h(u)$ is the marginal pdf of U. This equation along with (iii) $\int_{d_1}^{d_2}h(u)du=1-\alpha$, finds the constants d_1 and d_2 explicitly. Thus Basu's Theorem simplifies calculations from the conditional distribution of T_1 given T_2 to that of the marginal distribution of U which is ancillary under H_0. The calculations simplify further in this case, since marginally U is one-to-one with a Student's t-statistic with $n-1$ degrees of freedom.

Hogg (1961) discussed another application of Basu's Theorem in performing the test of a composite hypothesis on the basis of several mutually independent statistics. As one may anticipate, in actual examples, the independence of these statistics is established by Basu's Theorem. Hogg's general formulation of the problem is as follows.

Let $\Theta=\Theta_1$ be the parameter space of θ, where θ is a real- or a vector-valued parameter. We want to test $H_0:\theta\in\Theta_*$ against the alternatives $H_1:\theta\in\Theta_1-\Theta_*$, where Θ_* is some subset of Θ_1. Suppose there exists a nested sequence of subsets of the parameter space given by $\Theta_1\supset\Theta_2\supset\cdots\supset\Theta_k=\Theta_*$, and one is interested in testing a sequence of hypotheses $H_0^i : \theta \in \Theta_i$ against $\theta \in \Theta_{i-1} - \Theta_i$, $i = 2, \cdots, k$. Obviously, one does not test H_0^i unless one accepts H_0^{i-1}. Further, since $H_0 \equiv \cap_{i=2}^k H_0^i$, H_0 is accepted if and only if all of $H_0^2, \cdots, H_0^k$ are accepted. If now λ_i denotes the GLRT statistic for H_0^i : $\theta \in \Theta_i$ against $\theta \in \Theta_{i-1} - \Theta_i$, $i = 2, \cdots, k$, and λ denotes the GLRT statistic for H_0 against H_1, then $\lambda = \prod_{i=2}^k \lambda_i$. Hence, $-2\log\lambda = \sum_{i=2}^k (-2\log\lambda_i)$. If the λ_i are mutually independent, one can, in principle, get the distribution of $-2\log\lambda$ by applying some convolution result. Also (though not necessary), the asymptotic chi-squaredness of $-2\log\lambda$ follows from the chi-squaredness of the individual λ_i's. In order to establish the independence of the λ_i's in actual examples, suppose for each i, when $\theta \in \Theta_i$, there exist a complete sufficient statistic, while λ_i is ancillary. Then by Basu's Theorem, λ_i is distributed independently of this complete sufficient statistic under H_0^i. If now $\lambda_{i+1}, \cdots, \lambda_k$ are all functions of this complete sufficient statistic, then under H_0^i, λ_i is distributed independently of $(\lambda_{i+1}, \cdots, \lambda_k)$, $i = 2, \cdots, k-1$. Since $H_0 \equiv \bigcap_{i=2}^{k} H_0^i$, the mutual independence of $\lambda_2, \cdots, \lambda_k$ follows under H_0.

The following simple example illustrates this idea. For other interesting applications, we refer to Hogg (1961).

Example 14. Let $X_1, \cdots, X_k$ be independently distributed with $X_i \sim N(0, \sigma_i^2)$, $i = 1, \cdots, k$. Let $\Theta = \Theta_1 = \{(\sigma_1^2, \cdots, \sigma_k^2) : \sigma_i^2 > 0 \text{ for all } i\}$. Also, let $\Theta_i = \{(\sigma_1^2, \cdots, \sigma_k^2) : \sigma_1^2 = \cdots = \sigma_i^2\}$, $i = 2, \cdots, k$. Then if λ denotes the GLRT statistic for testing $H_0 : \sigma_1^2 = \cdots = \sigma_k^2$ against all possible alternatives, then $\lambda = \prod_{i=2}^k \lambda_i$, where λ_i is the GLRT statistic for testing $H_0^i : (\sigma_1^2, \cdots, \sigma_k^2) \in \Theta_i$ against the alternative $H_1^i : (\sigma_1^2, \cdots, \sigma_k^2) \in \Theta_{i-1} - \Theta_i$, $i = 2, \cdots, k$. Let $S_i = X_i^2$, $i = 1, \cdots, k$. It can be shown that λ_i is a function of the ratio $\sum_{j=1}^{i-1} S_j / \sum_{j=1}^{i} S_j$, for each $i = 2, \cdots, k$, and $(\sum_{j=1}^{i} S_j, S_{i+1}, \cdots, S_k)$ is complete sufficient under H_0^i for each $i = 2, \cdots, k$. Thus, under H_0^i, λ_i is ancillary, and is distributed independently of $\lambda_{i+1}, \cdots, \lambda_k$. Since $H_0 \equiv \cap_{i=2}^k H_0^i$, one gets the mutual independence of $\lambda_2, \cdots, \lambda_k$ under H_0.

Remark 4. Hogg and Randles (1971) have proved the independence of certain statistics by applying Basu's Theorem in a nonparametric framework. To see this in a simple case, let $X_1, \cdots, X_n$ be iid with an unknown continuous distribution function F, and let $X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$ denote the ordered X_i's. Then it is well-known (see for example Fraser, 1953) that $(X_{(1)}, \cdots, X_{(n)})$ is complete sufficient. On the other hand, if $R_1, \cdots, R_n$ denote the ranks of $X_1, \cdots, X_n$, then the joint distribution of these ranks is permutation invariant, and does not depend on F. Hence, $(R_1, \ldots, R_n)$ is distributed independently of $(X_{(1)}, \ldots, X_{(n)})$ and thus any linear rank statistic (which is a function of $R_1, \ldots, R_n$) is distributed independently of any function of $(X_{(1)}, ..., X_{(n)})$. This simple fact can be used to establish the independence of many statistics emerging naturally in nonparametric problems.

5. Basu's Theorem in Estimation

5.1 *Best Equivariant Estimators.* Basu's Theorem is applicable in certain estimation problems as well. We begin with its application in the derivation of best equivariant estimators. For simplicity, we illustrate this only for location family of distributions. Specifically, let $X_1, \cdots, X_n$ have joint pdf $f_\theta(x_1, \cdots, x_n) = f(x_1 - \theta, \cdots, x_n - \theta)$. Consider the group $\mathcal{G}$ of transformations such that for any $g_c \in \mathcal{G}$, c real, $g_c(x_1, \cdots, x_n) = (x_1 + c, \cdots, x_n + c)$. Then the maximal invariant is given by $Y = (Y_1, \cdots, Y_{n-1})^T$, where $Y_i = X_i - X_n$, $i = 1, \cdots, n-1$, and an estimator δ of θ is said to be *location-equivariant* if $\delta(X_1 + c, \cdots, X_n + c) = \delta(X_1, \cdots, X_n) + c$ for every real c. If the loss is squared error, then the best location-equivariant estimator of θ is the Pitman estimator given by $\hat{\theta}_P = \int \theta f(X_1 - \theta, \cdots, X_n - \theta) d\theta / \int f(X_1 - \theta, \cdots, X_n - \theta) d\theta$. For a more general loss, this need not be the case. Even in such cases, it is often possible to find the best location-equivariant estimator, and Basu's Theorem facilitates the calculations. To this end, we begin with a general result given in Lehmann and Casella (1998, pp. 151-152).

THEOREM 7. *Consider the location family of distributions and the group $\mathcal{G}$ of transformations as above. Suppose the loss function $L(\theta, a)$ is of the form $L(\theta, a) = \rho(a - \theta)$, and that there exists a location-equivariant estimator δ_0 of θ with finite risk. Assume for each y, there exists $v^*(y)$ which minimizes $E_0[\rho(\delta_0(X) - v(y))|Y = y]$. Then the best (minimum risk) location-equivarint estimator δ^* of θ exists, and is given by $\delta^*(X) = \delta_0(X) - v^*(Y)$.*

In many applications, there exists a complete sufficient statistic T for θ, and one can find a δ_0 which is a function of T, while Y is ancillary. Then the task of finding a minimum risk equivariant estimator greatly simplifies, because one can then work with the unconditional distribution of T so that $v(y)$ becomes a constant. We illustrate this with two examples.

EXAMPLE 15. Let $X_1, \cdots, X_n$ be iid $N(\theta, 1)$, where $\theta \in \mathcal{R}^1$. Then the location-equivariant estimator $\delta_0(X_1, \cdots, X_n) = \bar{X}$ is complete sufficient, and the maximal invariant Y as defined earlier is ancillary. Hence, by Basu's Theorem, $\bar{X}$ is independent of Y. Now for any general loss $L(\theta, a) = W(|a - \theta|)$, where W is monotonically non-decreasing in its argument, the minimum risk location-equivariant estimator of θ is $\bar{X} - E_0(\bar{X}) = \bar{X} - 0 = \bar{X}$.

EXAMPLE 16. Let $X_1, \cdots, X_n$ be iid with common pdf $f_\theta(x) = \exp[-(x-\theta)]I_{[x \geq \theta]}$, where $\theta \in \mathcal{R}^1$. Here $X_{(1)} = \min(X_1, \cdots, X_n)$ is location-equivariant as well as complete sufficient for θ, and Y is ancillary. Then, under squared error loss, the best location-equivariant estimator of θ is $X_{(1)} - E_0(X_{(1)}) = X_{(1)} - \frac{1}{n}$, while under absolute error loss, the best equivariant estimator of θ is $X_{(1)} - \text{med}_0(X_{(1)}) = X_{(1)} - \frac{\log_e 2}{n}$.

5.2 *Uniformly minimum variance unbiased estimators.* Next we consider applications of Basu's Theorem in finding uniformly minimum variance unbiased estimators (UMVUE's) of distribution functions. We present below a general theorem due to Sathe and Varde (1969).

THEOREM 8. *Let Z be a real-valued random variable with distribution function $F_\theta(z)$, where $\theta \in \Theta$. Let T be complete sufficient for θ, and let $V(z,t)$ be a function of z and t such that*
(i) $V(z,t)$ is strictly increasing in z for fixed t;
(ii) $V(Z,T) = U$ (say) is ancillary, and has distribution function $H(u)$.
Then the UMVUE of $F_\theta(z)$ is $H(V(z,T))$.

PROOF. Since $I_{[Z \leq z]}$ is an unbiased estimator of $F_\theta(z)$, by the Rao-Blackwell-Lehmann-Scheffe Theorem, the UMVUE of $F_\theta(z)$ is

$$\begin{aligned} E[I_{[Z\leq z]}|T=t] &= P(Z \leq z|T=t) \\ &= P(V(Z,T) \leq V(z,T)|T=t) \text{ (by (i))} \\ &= P(U \leq V(z,t)|T=t) \\ &= P(U \leq V(z,t)) \text{ (by Basu's Theorem)} \\ &= H(V(z,t)) \text{ (by (ii))} \end{aligned}$$

We illustrate the theorem with two examples.

EXAMPLE 17. Suppose $X_1, \cdots, X_n$ $(n \geq 3)$ are iid $N(\mu, \sigma^2)$, where $\mu \in \mathcal{R}^1$ and $\sigma > 0$. Kolmogorov (1950) found the UMVUE of $P_{\mu,\sigma^2}(X_1 \leq x)$. Writing $\bar{X} = n^{-1}\sum_{i=1}^n X_i$ and $S^2 = \sum_{i=1}^n (X_i - \bar{X})^2$, $T = (\bar{X}, S)$ is complete sufficient, and $U = (X_1 - \bar{X})/S$ is ancillary. Hence, by Theorem 8, the UMVUE of $F_\theta(x)$ is given by $H((x - \bar{X})/S)$, where H is the marginal distribution function of U. The marginal pdf of U in this case is given in Rao (1973, p. 323).

EXAMPLE 18. Let $X_1, \cdots, X_n$ $(n \geq 2)$ be iid with common Weibull pdf

$$f_\theta(x) = \exp(-x^p/\theta)(p/\theta)x^{p-1};\ 0 < x < \infty, 0 < \theta < \infty,$$

$p(>0)$ being known. In this case, $T = \sum_{i=1}^{n} X_i^p$ is complete sufficient for θ, while $U = X_1^p/T$ is ancillary. Also, since $X_1^p, \cdots, X_n^p$ are iid exponential with scale parameter θ, $U \sim \text{Beta}(1, n-1)$. Hence, the UMVUE of $P_\theta(X_1 \leq x) = P_\theta(X_1^p \leq x^p)$ is given by

$$k(T) = \begin{cases} 1 - x^{np}/T^n & \text{if } T > x^p, \\ \\ 1 & \text{if } T \leq x^p. \end{cases}$$

Eaton and Morris (1970) proposed an alternative to Theorem 8 in order to enhance the scope of applications, including multivariate generalizations. They noted that if an unbiased estimator of a parameter could be expressed as a function of the complete sufficient statistic and an ancillary statistic, then the UMVUE of the parameter could be expressed as the expectation of the same function, expectation being taken over the marginal distribution of the ancillary statistic. The complete sufficient statistic does not play any role in the final calculation, since by Basu's Theorem, it is independent of the ancillary statistic. We now state the relevant theorem of Eaton and Morris along with its proof.

THEOREM 9. *Let $h(X)$ be an unbiased estimator of a parameter $\gamma(\theta)$. Also, let $h(X) = W(T, U)$, where T is complete sufficient for θ and U is ancillary. Then the UMVUE of $\gamma(\theta)$ is given by $\gamma^*(T)$, where $\gamma^*(T) = E_U[W(T, U)]$, E_U denoting expectation over the marginal distribution of U.*

PROOF. From the Rao-Blackwell-Lehmann-Scheffe Theorem, the UMVUE of $\gamma(\theta)$ is given by

$$\begin{aligned} E[h(X)|T=t] &= E[W(T,U)|T=t] = E[W(t,U)|T=t] \\ &= E_U[W(t,U)] = \gamma^*(t). \end{aligned}$$

REMARK 5. Eaton and Morris (1970) did not require existence of a function $V(z, t)$ which was non-decreasing in z for fixed t. Lack of monotonicity enhances the scope of application. However, Eaton and Morris required instead an unbiased estimator to be a function of the complete sufficient statistic and an ancillary statistic. This requirement is met in many applications including the examples of Sathe and Varde as given above. In order to find U and W, Eaton and Morris brought in very successfully the notion of group invariance which led to many interesting applications including a multivariate generalization of Kolmogorov's result.

6. Basu's Theorem in Empirical Bayes Analysis

Empirical Bayes (EB) analysis has, of late, become very popular in statistics, especially when the problem is simultaneous estimation of several parameters. As described by Berger (1985), an EB scenario is one in which known relationships among the coordinates of a parameter vector, say, $\theta = (\theta_1, \cdots, \theta_k)^T$ allow use of the data to estimate some features of the prior distribution. For example, one may have reasons to believe that the θ_i are iid from a prior Π_0, where Π_0 is structurally known except possibly for some unknown parameter (possibly vector-valued) λ. A parametric EB procedure is one where λ is estimated from the marginal distribution of the observations.

Often in an EB analysis, one is interested in finding Bayes risks of the EB estimators. Basu's Theorem helps considerably in many such calculations as we demonstrate below.

Example 19. We consider an EB framework as proposed in Morris (1983a, 1983b). Let $X_i|\theta_i$ be independent $N(\theta_i, V)$, where $V(>0)$ is assumed known. Let θ_i be independent $N(\boldsymbol{z}_i^T\boldsymbol{b}, A)$, $i = 1, \cdots, k$. The p-component ($p < k$) design vectors $\boldsymbol{z}_i$ are assumed to be known, and let $\boldsymbol{Z}^T = (\boldsymbol{z}_1, \cdots, \boldsymbol{z}_k)$. We assume rank$(\boldsymbol{Z}) = p$. Based on the above likelihood and the prior, the posteriors of the θ_i are independent $N((1-B)X_i+B\boldsymbol{z}_i^T\boldsymbol{b}, V(1-B))$, where $B = V/(V+A)$. Accordingly, the posterior means, the Bayes estimators of the θ_i are given by

$$\hat{\theta}_i^{BA} = (1-B)X_i + B\boldsymbol{z}_i^T\boldsymbol{b},\ i = 1, \cdots, k. \tag{6.1}$$

In an EB set up, $\boldsymbol{b}$ and A are unknown, and need to be estimated from the marginal distributions of the X_i's. Marginally, the X_i's are independent with $X_i \sim N(\boldsymbol{z}_i^t\boldsymbol{b}, V + A)$. Then, writing $\boldsymbol{X} = (X_1, \cdots, X_k)^T$, based on the marginal distribution of $\boldsymbol{X}$, the complete sufficient statistic for $(\boldsymbol{b}, A)$ is $(\hat{\boldsymbol{b}}, S^2)$, where $\hat{\boldsymbol{b}} = (\boldsymbol{Z}^T\boldsymbol{Z})^{-1}\boldsymbol{Z}^T\boldsymbol{X}$ is the least squares estimator or the MLE of $\boldsymbol{b}$, and $S^2 = \sum_{i=1}^k (X_i - \boldsymbol{z}_i^T\hat{\boldsymbol{b}})^2$. Also, based on the marginal of $\boldsymbol{X}$, $\hat{\boldsymbol{b}}$ and S^2 are independently distributed with $\hat{\boldsymbol{b}} \sim N(\boldsymbol{b}, (V+A)(\boldsymbol{Z}^T\boldsymbol{Z})^{-1})$, and $S^2 \sim (V+A)\chi^2_{k-p}$. Accordingly $\boldsymbol{b}$ is estimated by $\hat{\boldsymbol{b}}$. The MLE of B is given by $\min(kV/S^2, 1)$, while its UMVUE is given by $V(k-p-2)/S^2$, where we must assume $k > p + 2$ for the latter to be meaningful. If instead, one assigns the prior $\Pi(\boldsymbol{b}, A) \propto 1$ as in Morris (1983a, 1983b), then the HB estimator of θ_i is given by $\hat{\theta}_i^{HB} = (1 - B^*(S^2))X_i + B^*(S^2)z_i^T\hat{\boldsymbol{b}}$, where $B^*(S^2) = \int_0^1 B^{\frac{1}{2}(k-p-2)}\exp(-\frac{1}{2V}BS^2)dB / \int_0^1 B^{\frac{1}{2}(k-p-4)}\exp(-\frac{1}{2V}BS^2)dB$. Thus a general EB estimator of θ_i is of the form

$$\hat{\theta}_i = [1 - \hat{B}(S^2)]X_i + \hat{B}(S^2)\boldsymbol{z}_i^T\hat{\boldsymbol{b}}. \tag{6.2}$$

We will now demonstrate an application of Basu's Theorem in finding the mean squared error (MSE) matrix $E[(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T]$, where $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \cdots, \hat{\theta}_k)^T$, and expectation is taken over the joint distribution of $\boldsymbol{X}$ and $\boldsymbol{\theta}$. The following theorem provides a general expression for the MSE matrix.

THEOREM 10. *With the notations of this section,*

$$\begin{aligned} E[(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T] &= V(1-B)\boldsymbol{I}_k + VB\boldsymbol{Z}(\boldsymbol{Z}^T\boldsymbol{Z})^{-1}\boldsymbol{Z}^T \\ &+ E[(\hat{B}(S^2) - B)^2 S^2](k-p)^{-1}(\boldsymbol{I}_k - \boldsymbol{Z}(\boldsymbol{Z}^T\boldsymbol{Z})^{-1}\boldsymbol{Z}^T). \end{aligned}$$

PROOF. Write $\hat{\boldsymbol{\theta}}^{BA} = (1-B)\boldsymbol{X} + B\boldsymbol{Z}\boldsymbol{b}$. Then

$$\begin{aligned} E[(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T] &= E[(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}^{BA} + \hat{\boldsymbol{\theta}}^{BA} - \hat{\boldsymbol{\theta}})(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}^{BA} + \hat{\boldsymbol{\theta}}^{BA} - \hat{\boldsymbol{\theta}})^T] \\ &= E[(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}^{BA})(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}^{BA})^T] + E[(\hat{\boldsymbol{\theta}}^{BA} - \hat{\boldsymbol{\theta}})(\hat{\boldsymbol{\theta}}^{BA} - \hat{\boldsymbol{\theta}})^T], \end{aligned} \tag{6.3}$$

since

$$E[(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}^{BA})(\hat{\boldsymbol{\theta}}^{BA} - \hat{\boldsymbol{\theta}})^T] = E[E(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}^{BA}|\boldsymbol{X})(\hat{\boldsymbol{\theta}}^{BA} - \hat{\boldsymbol{\theta}})^T] = \boldsymbol{0}.$$

Now

$$\begin{aligned} E[(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}^{BA})(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}^{BA})^T] &= E[(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}^{BA})(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}^{BA})^T|\boldsymbol{X}] \\ &= E[Var(\boldsymbol{\theta}|\boldsymbol{X})] = E[V(1-B)\boldsymbol{I}_k] = V(1-B)\boldsymbol{I}_k. \end{aligned} \tag{6.4}$$

Next after a little algebra, we get

$$\hat{\boldsymbol{\theta}}^{BA} - \hat{\boldsymbol{\theta}} = (\hat{B}(S^2) - B)(\boldsymbol{X} - \boldsymbol{Z}\hat{\boldsymbol{b}}) + B\boldsymbol{Z}(\hat{\boldsymbol{b}} - \boldsymbol{b}).$$

Now by the independence of $\hat{\boldsymbol{b}}$ with $\boldsymbol{X} - \boldsymbol{Z}\hat{\boldsymbol{b}}$, noting $S^2 = ||\boldsymbol{X} - \boldsymbol{Z}\hat{\boldsymbol{b}}||^2$, where $||\cdot||$ denotes the Euclidean norm, and $Var(\hat{\boldsymbol{b}}) = VB^{-1}(\boldsymbol{Z}^T\boldsymbol{Z})^{-1}$, one gets

$$\begin{aligned} E(\hat{\boldsymbol{\theta}}^{BA} - \hat{\boldsymbol{\theta}})(\hat{\boldsymbol{\theta}}^{BA} - \hat{\boldsymbol{\theta}})^T &= E[(\hat{B}(S^2) - B)^2(\boldsymbol{X} - \boldsymbol{Z}\hat{\boldsymbol{b}})(\boldsymbol{X} - \boldsymbol{Z}\hat{\boldsymbol{b}})^T] \\ &+ VB\boldsymbol{Z}(\boldsymbol{Z}^T\boldsymbol{Z})^{-1}\boldsymbol{Z}^T. \end{aligned} \tag{6.5}$$

Next we observe that

$$\frac{(\boldsymbol{X} - \boldsymbol{Z}\hat{\boldsymbol{b}})(\boldsymbol{X} - \boldsymbol{Z}\hat{\boldsymbol{b}})^T}{S^2} = \frac{\left[(\boldsymbol{X} - \boldsymbol{Z}\boldsymbol{b}) - \boldsymbol{Z}\left(\hat{\boldsymbol{b}} - \boldsymbol{b}\right)\right]\left[(\boldsymbol{X} - \boldsymbol{Z}\boldsymbol{b}) - \boldsymbol{Z}\left(\hat{\boldsymbol{b}} - \boldsymbol{b}\right)\right]^T (V+A)^{-1}}{\|(\boldsymbol{X} - \boldsymbol{Z}\boldsymbol{b}) - \boldsymbol{Z}\left(\hat{\boldsymbol{b}} - \boldsymbol{b}\right)\|^2 (V+A)^{-1}}$$

is ancillary, and by Basu's Theorem, is independent of S^2, which is a function of the complete sufficient statistic $(\hat{\boldsymbol{b}}, S^2)$. Accordingly,

$$\begin{aligned} &E[(\hat{B}(S)-B)^2(\boldsymbol{X}-\boldsymbol{Z}\hat{\boldsymbol{b}})(\boldsymbol{X}-\boldsymbol{Z}\hat{\boldsymbol{b}})^T] \\ &\quad = E[(\hat{B}(S)-B)^2 S^2]E[(\boldsymbol{X}-\boldsymbol{Z}\hat{\boldsymbol{b}})(\boldsymbol{X}-\boldsymbol{Z}\hat{\boldsymbol{b}})^T/S^2], \end{aligned} \tag{6.6}$$

and then by the formula for moments of ratios,

$$\begin{aligned} E[(\boldsymbol{X}-\boldsymbol{Z}\hat{\boldsymbol{b}})(\boldsymbol{X}-\boldsymbol{Z}\hat{\boldsymbol{b}})^T/S^2] &= E[(\boldsymbol{X}-\boldsymbol{Z}\hat{\boldsymbol{b}})(\boldsymbol{X}-\boldsymbol{Z}\hat{\boldsymbol{b}})^T]/E(S^2) \\ &= (k-p)^{-1}[\boldsymbol{I}_k - \boldsymbol{Z}(\boldsymbol{Z}^T\boldsymbol{Z})^{-1}\boldsymbol{Z}^T]. \end{aligned} \tag{6.7}$$

The theorem follows now from (6.3)-(6.7).

Remark 6. With the choice $\hat{B}(S^2) = V(k-p-2)/S^2$ $(k > p+2)$, Lindley's modification of the James-Stein shrinker, the expression $E[(\hat{B}(S^2) - B)^2 S^2]$ simplifies further, and is given by $E[(\hat{B}(S^2)-B)^2S^2] = 2VB$, where we use $S^2 \sim VB^{-1}\chi^2_{k-p}$.

Remark 7. Under the same framework, suppose one is interested in finding $E[(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})^T|U]$, where $U = ||\boldsymbol{X}-\boldsymbol{Z}\hat{\boldsymbol{b}}||^2$ is an ancillary statistic based on the marginal distribution of $\boldsymbol{X}$. Examples of this type are considered by Hill (1989). As before, Basu's Theorem is still applicable. The only change that occurs in Theorem 10 is that the expression $(k-p)^{-1}(\boldsymbol{I}_k - \boldsymbol{Z}(\boldsymbol{Z}^T\boldsymbol{Z})^{-1}\boldsymbol{Z}^T)$, the unconditional expectation of U is now replaced by U itself. Datta, Ghosh, Smith and Lahiri (2002) have considered EB confidence intervals for the different components of $\boldsymbol{\theta}$, where the initial calculations also need Basu's Theorem.

7. Basu's Theorem and Infinite Divisibility

A very novel application of Basu's Theorem appears recently in Dasgupta (2002) in proving the infinite divisibility of certain statistics. In addition to Basu's Theorem, this application requires a version of the Goldie-Steutel law which we now describe below as a lemma.

Lemma 1. *Let U be exponentially distributed with scale parameter 1, and let W be distributed independently of U. Then UW is infinitely divisible.*

A second lemma is needed for proving Dasgupta's main result.

Lemma 2. *Let V be exponentially distributed with scale parameter 1, and let $\alpha > 0$. Then V^α admits the representation $V^\alpha = UW$, where U is* *as in Lemma 1, and W is distributed independently of U. Thus, by Lemma 1,* V^α *is infinitely divisible.*

The two main results of Dasgupta are now as follows.

THEOREM 11. *Let* Z_1 *and* Z_2 *be iid* $N(0,1)$ *and* $(Z_3, \cdots, Z_m)^T$ *be distributed independently of* $(Z_1, Z_2)^T$*. Let g be any homogeneous function in two variables, that is,* $g(cx_1, cx_2) = c^2 g(x_1, x_2)$ *for all real* x_1 *and* x_2 *and for all* $c > 0$*. Then for any positive integer k, and any arbitrary random variable h,* $g^k(Z_1, Z_2)h(Z_3, \cdots, Z_m)$ *is infinitely divisible.*

The proof makes an application of Basu's Theorem in its first step by writing $g(Z_1, Z_2) = (Z_1^2 + Z_2^2)g(Z_1, Z_2)/(Z_1^2 + Z_2^2)$, noting the ancillarity of $g(Z_1, Z_2)/(Z_1^2 + Z_2^2)$ and the complete sufficiency of $Z_1^2 + Z_2^2$ for the augmented $N(0, \sigma^2)$, $(\sigma^2 > 0)$ family of distributions. This establishes the independence of $Z_1^2 + Z_2^2$ and $g(Z_1, Z_2)/(Z_1^2 + Z_2^2)$. Hence, $g(Z_1, Z_2)$ can be expressed as UW, where U is exponentially distributed with scale parameter 1, and W is distributed independently of U. An application of Lemma 2 now establishes the result.

The next result of Dasgupta provides a representation of functions of normal variables as the product of two random variables, where one is infinitely divisible, while the other is not, and the two are independently distributed.

THEOREM 12. *Let* $X_1, X_2, \cdots, X_n$ *be iid* $N(0,1)$*, and let* $h_i(x_1, \cdots, x_n)$ $(1 \le i \le n)$ *be scale-invariant functions in that* $h_i(cx_1, \cdots, cx_n) = h_i(x_1, \cdots, x_n)$ *for all* $c > 0$*. Suppose g is any continuous homogeneous function in the* $n-$ *space, that is,* $g(cx_1, \cdots, cx_n) = c^n g(x_1, \cdots, x_n)$ *for all* $c > 0$*. Let* $||\boldsymbol{X}||$ *denote the Euclidean norm of the vector* $(X_1, \cdots, X_n)^T$*. Then the random variable*

$$k(X_1, \cdots, X_n) = g(||\boldsymbol{X}||h_1(X_1, \cdots, X_n), \cdots, ||X||h_n(X_1, \cdots, X_n))$$

admits the representation $k(X_1, \cdots, X_n) = YZ$*, where Y is infinitely divisible, Z is not, both are non-degenerate, and Y and Z are independent.*

Dasgupta provides an interesting application of the above theorem.

EXAMPLE 20. Let $X_1, X_2, \cdots, X_n$ be iid $N(0,1)$. Then each one of the four variables (i) $\prod_{i=1}^n X_i$, (ii) $\sum_{i=1}^n X_i^n$, (iii) $\prod_{i=1}^n X_i^2/(\sum_{i=1}^n X_i^2)^{n/2}$ and (iv) $\sum_{i=1}^n X_i^{2n}/(\sum_{i=1}^n X_i^2)^{n/2}$ has the factorization as given in Theorem 12.

Acknowledgements. It is a real pleasure to acknowledge Gauri Sankar Datta for his meticulous reading of the manuscript, and correcting many typos and ambiguities. This project would not have materialized without the inspiration and continuous encouragement of Anirban Dasgupta who made also some useful remarks on an earlier draft of the paper. In addition, the paper has benefitted from the comments of Erich Lehmann, Herbert A. David, and in particular, Robert V. Hogg who pointed out to me several important references.

References

Bahadur, R.R. (1957). On unbiased estimates of uniformly minimum variance. *Sankhyā*, **18**, 211-224.

Basu, D. (1955). On statistics independent of a complete sufficient statistic. *Sankhyā*, **15**, 377-380.

Basu, D. (1958). On statistics independent of a sufficient statistic. *Sankhyā*, **20**, 223-226.

Basu, D. (1964). Recovery of ancillary information. *Sankhyā Ser. A*, **26**, 3-16.

Basu, D. (1982). Basu theorems. In *Encyclopedia of Statistical Sciences,* **VI**, S. Kotz and N.L. Johnson, eds., Wiley, New York, 193-196.

Berger, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis*, 2nd Edition, Springer Verlag, New York.

Boos, D.D. and Hughes-Oliver, J.M. (1998). Applications of Basu's Theorem. *Amer. Statist.*, **52**, 218-221.

Casella, G. and Berger, R. (2001). *Statistical Inference*, 2nd Edition, Wadsworth, Pacific Groves, California.

Dasgupta, A. (2002). Basu's theorem, Poincaré inequalities and infinite divisibility. Preprint.

Datta, G.S., Ghosh, M., Smith, D.D. and Lahiri, P. (2002). On an asymptotic theory of conditional and unconditional coverage probabilities of empirical Bayes confidence intervals. *Scand. J. Statist.*, **29**, 139-152.

David, H.A. (1981). *Order Statistics*, 2nd Edition, Wiley, New York.

Dawid, A.P. (1979). Conditional independence in statistical theory. *J.R. Statist. Soc. Ser. B*, **41**, 1-15.

Eaton, M.L. and Morris, C.N. (1970). The application of invariance to unbiased estimation. *Ann. Math. Statist.*, **41**, 1708-1716.

Fraser, D.A.S. (1957). *Nonparametric Methods in Statistics.* Wiley, New York.

Ghosh, M. (1992). Hierarchical and empirical Bayes multivariate estimation. *Current Issues in Statistical Inference: Essays in Honor of D. Basu*, M. Ghosh and P.K. Pathak, eds., Institute of Mathematical Statistics Lecture Notes and Monograph Series, **17**, 151-177.

Hill, J.R. (1990). A general framework for model-based statistics. *Biometrika*, **77**, 115-126.

Hogg, R.V. (1953). Testing the equality of means of rectangular populations. (Abstract). *Ann. Math. Statist.*, **24**, 691.

Hogg, R.V. (1961). On the resolution of statistical hypotheses. *Ann. Math. Statist.*, **32**, 978-989.

HOGG, R.V. and CRAIG, A.T. (1956). Sufficient statistics in elementary distribution theory. *Sankhyā*, **16**, 209-216.

HOGG, R.V. and RANDLES, R.H. (1971). Certain uncorrelated and independent rank statistics. *J. Amer. Statist. Assoc.*, **66**, 569-574.

JOHNSTONE, I.M. and VELLEMAN, P.F. (1985). Efficient scores, variance decomposition, and Monte carlo swindles. *J. Amer. Statist. Assoc.*, **80**, 851-862.

KOEHN, U. and THOMAS, D.L. (1975). On statistics independent of a sufficient statistic: Basu's Lemma. *Amer. Statist.*, **29**, 40-42.

KOLMOGOROV, A.N. (1942). Sur l'estimation statisque des parametres de la loi de Gauss. *Bull. Acad. Sci. URSS Ser. Math.*, **6**, 3-32.

KOLMOGOROV, A.N. (1950). Unbiased estimates. *Izvestiaa Akad. Nauk SSSR Series Math.*, **14**, 303-326.

LEHMANN, E.L. (1981). An interpretation of completeness and Basu's Theorem. *J. Amer. Statist. Assoc.*, **76**, 335-340.

LEHMANN, E.L. (1986). *Testing Statistical Hypotheses*, 2nd Edition. Wiley, New York.

LEHMANN, E.L. and CASELLA, G. (1998). *Theory of Point Estimation*, 2nd Edition. Springer-Verlag, New York.

MORRIS, C.N. (1983a). Parametric empirical Bayes confidence intervals. In *Scientific Inference, Data Analysis and Robustness*, G.E.P. Box, T. Leonard and C.F.J. Wu, eds., Academic Press, New York, 25-50.

MORRIS, C.N. (1983b) (with discussion). Parametric empirical Bayes inference: theory and applications. *J. Amer. Statist. Assoc.*, **78**, 47-65.

MUKHOPADHYAY, N. (2000). *Probability and Statistical Inference.* Marcel Dekker, New York.

RAO, C.R. (1973). *Linear Statistical Inference and its Applications*, 2nd Edition. Wiley, New York.

SATHE, Y.S. and VARDE, S.R. (1969). On minimum variance unbiased estimation of reliability. *Ann. Math. Statist.*, **40**, 710-714.

MALAY GHOSH
DEPARTMENT OF STATISTICS
UNIVERSITY OF FLORIDA
GAINESVILLE, FL 32611-8545
USA
E-mail: ghoshm@stat.ufl.edu