NOTES

This section is devoted to brief research and expository articles and other short items.

EXTENSION OF A THEOREM OF BLACKWELL[1]

By E. W. Barankin

University of California, Berkeley

1. Introduction. In [1] (§1) the author has announced, as bearing on the results there, that Blackwell's method [2] of uniformly improving the variance of an unbiased estimate by taking the conditional expectation with respect to a sufficient statistic, is in fact similarly effective on every absolute central moment of order $s \geqq 1$. Our purpose here is to establish this. In addition, the equality condition (null improvement of the moment) is presented in terms of a primitive property of the estimate. The asserted uniform diminution of the s-th moments for a family W of distributions is, as in the case $s = 2$, a twice removed consequence of the fundamental fact for a single distribution that the absolute s-th power of the conditional expectation of a measurable function is almost everywhere (a.e.) not greater than the conditional expectation of the absolute s-th power of the function. This is the substance of the theorem below. The second corollary then states the result for unbiased estimates.

2. Preliminaries. Let Ω be a space of points x; $\mathfrak{F}$, a σ-field of subsets of Ω; and μ, a probability measure on $\mathfrak{F}$. Let t be a function on Ω onto a space Γ of points τ; $\mathfrak{T}^{\Gamma}$ a σ-field of subsets of Γ; and $\mathfrak{T}$—a sub-σ-field of $\mathfrak{F}$—the inverse of $\mathfrak{T}^{\Gamma}$ under t. A set in $\mathfrak{T}^{\Gamma}$ will be denoted by A^{Γ}, where A is its inverse undeı t. Let ν denote the measure on $\mathfrak{T}^{\Gamma}$ defined by $\nu(A^{\Gamma}) = \mu(A)$.

If f is a real-valued,[2] $\mathfrak{F}$-measurable, μ-integrable function on Ω, we denote by $E(f \mid \cdot)$ the conditional expectation of f with respect to t. Corresponding to any particular function h on Γ (as, for example, $E(f \mid \cdot)$) we define the function h^* on Ω by

$$h^*(x) = h(\tau), \qquad t(x) = \tau.$$

The qualification "essentially" prefixing a statement will mean that with the possible exception of a set of points of measure 0, that statement holds true.

The following two simple lemmas enable us to present the conditions for equality, in the results below, in terms of the elementary characteristics of the function f.

[1] This note was prepared under O. N. R. contract.

[2] With no changes in this note, and only minor changes in [1], the results we have set forth concerning unbiased estimation pertain as well to complex-valued functions.

LEMMA 1. *A necessary and sufficient condition that* $\operatorname{sgn} f(x) = \operatorname{sgn} E^*(f \mid x)$ a.e. (μ) *is that* $\operatorname{sgn} f$ *be essentially a function of* t.
The necessity of the condition is clear. To prove sufficiency, let f' be a function on Ω which is a.e. equal to f, and such that $\operatorname{sgn} f'$ is an (unqualified) function of t. Now if $\operatorname{sgn} f'(x) = \operatorname{sgn} E^*(f \mid x)$ does not hold a.e. (μ), then there is a $\mathfrak{T}$-set, A, of positive measure, such that, for example, for $x \in A, f'(x) > 0$ while $E^*(f \mid x) \leqq 0$. We then have the contradiction

$$0 < \int_A f' \, d\mu = \int_A f \, d\mu = \int_A E^*(f \mid \cdot) \, d\mu \leqq 0.$$

LEMMA 2. *A necessary and sufficient condition that* $f(x) = E^*(f \mid x)$ a.e. (μ) *is that* f *be essentially a function of* t.
Again the necessity is obvious. To show sufficiency, let f' be a function on Ω which is a.e. equal to f, and is an (unqualified) function of t. Define h on Γ by

$$h(\tau) = f'(x), \qquad t(x) = \tau.$$

Then $h^* = f'$, and we have

$$\int_A f \, d\mu = \int_A f' \, d\mu = \int_{A\Gamma} h \, d\nu, \qquad A \in \mathfrak{T}.$$

But this implies that $h(\tau) = E(f \mid \tau)$ a.e. (ν), and therefore $f(x) = E^*(f \mid x)$ a.e. (μ), as was to be shown.

3. Results. For a proof of the Hölder inequality that we use in establishing the following theorem, we refer the reader to [3] (p. 233).

THEOREM.[3] *Let* $s \geqq 1$. *Then for almost all* $(\mu)x$,

$$| E^*(f \mid x) |^s \leqq E^*(|f|^s \mid x).^4 \tag{1}$$

Equality holds a.e.

(i) *for* $s = 1$, *if and only if* $\operatorname{sgn} f$ *is essentially a function of* t;
(ii) *for* $s > 1$, *if and only if* f *is essentially a function of* t.

PROOF: Consider first the case $s = 1$. Let

$$S = \{x \in \Omega \mid E^*(f \mid x) > 0\},$$
$$S' = \Omega - S.$$

Then, for any $A \in \mathfrak{T}$,

$$\int_A | E^*(f \mid \cdot) | \, d\mu = \int_{SA} E^*(f \mid \cdot) \, d\mu - \int_{S'A} E^*(f \mid \cdot) \, d\mu$$
$$= \int_{SA} f \, d\mu - \int_{S'A} f \, d\mu \leqq \int_A |f| \, d\mu = \int_A E^*(|f| \mid \cdot) \, d\mu.$$

[3] The proof we present here was suggested by the referee, and is much shorter than our own.

[4] For $s = 1$ this inequality was used by Doob in "Regularity properties of certain families of chance variables", *Trans. Amer. Math. Soc.*, Vol. 47 (1940), pp. 455–486 (Theorem 0.2).

Since A is arbitrary, we have the result (1) with $s = 1$. It is clear that the equality sign holds a.e. (μ) if and only if, except possibly for a set of measure 0, f is positive on S and non-positive on S'; that is, if and only if $\operatorname{sgn} f(x) = \operatorname{sgn} E^*(f \mid x)$ a.e. (μ). Applying Lemma 1, we have the equality condition as stated in the theorem.

Now let $s > 1$. To establish (1) it will suffice, by virtue of what has already been proved for $s = 1$, to consider $f \geqq 0$ a.e. (μ). We may then argue as follows. Unless (1) holds a.e., there is a $\mathfrak{T}$-set, R, of positive measure, and numbers $a > b \geqq 0$ such that for $x \in R$,

$$[E^*(f \mid x)]^s \geqq a,$$

and

$$E^*(f^s \mid x) \leqq b.$$

But then, with an application of the Hölder inequality we meet a contradiction. For,

$$\begin{aligned} a[\mu(R)]^s &\leqq \left\{\int_R E^*(f \mid \cdot)\, d\mu\right\}^s = \left\{\int_R f\, d\mu\right\}^s \\ &\leqq \int_R f^s\, d\mu \cdot [\mu(R)]^{s-1} = \int_R E^*(f^s \mid \cdot)\, d\mu \cdot [\mu(R)]^{s-1} \\ &\leqq b[\mu(R)]^s, \end{aligned}$$

which contradicts $a > b$. Thus, (1) is proved in general.

If $f(x) = E^*(f \mid x)$ a.e. (μ), it is readily proved by a direct argument that then equality holds in (1) a.e. (μ). Conversely, suppose equality in (1) holds a.e. Then we have, in fact, a.e.,

$$(2) \qquad \mid E^*(f \mid x) \mid = E^*(\mid f \mid \mid x),$$

and

$$(3) \qquad [E^*(\mid f \mid \mid x)]^s = E^*(\mid f \mid^s \mid x).$$

For brevity, denote the function $E^*(\mid f \mid \mid \cdot)$ by v. Since f vanishes at almost all points where v vanishes, we may write $\mid f \mid = w \cdot v$, where

$$w(x) = \begin{cases} 1, & v(x) = 0, \\ \mid f(x) \mid / v(x), & v(x) > 0. \end{cases}$$

(If v vanishes almost everywhere, we are through.) For any $\mathfrak{T}$-measurable, real-valued function, u, on Ω, we have

$$(4) \qquad \int_\Omega u \cdot v\, d\mu = \int_\Omega u \cdot v \cdot w\, d\mu,$$

when either of these integrals exists (cf. [4], p. 50, eq. (15)). Similarly, and taking account of the equality assumption (3) we have

$$\int_{\Omega} u \cdot v^s \, d\mu = \int_{\Omega} u \cdot v^s \cdot w^s \, d\mu. \tag{5}$$

In particular, consider the two functions

$$u_1(x) = \begin{cases} 1/v(x), & v(x) > 0, \\ 0, & v(x) = 0, \end{cases}$$

and

$$u_2(x) = \begin{cases} 1/[v(x)]^s, & v(x) > 0, \\ 0, & v(x) = 0. \end{cases}$$

If

$$S_0 = \{x \in \Omega \mid v(x) > 0\},$$

it is seen that u_1 taken in conjunction with (4), and u_2 taken in conjunction with (5), bring out

$$\int_{S_0} w \, d\mu = \int_{S_0} w^s \, d\mu = \mu(S_0).$$

From this it follows (e.g., by the equality condition attending the Hölder inequality) that $w(x) = 1$ a.e. in S_0. Hence $|f(x)| = v(x)$ a.e. in Ω. Therefore, by (2), $|f(x)| = |E^*(f \mid x)|$ a.e. But (2) also implies, as already shown, $\operatorname{sgn} f(x) = \operatorname{sgn} E^*(f \mid x)$ a.e. Thus, finally, we have $f(x) = E^*(f \mid x)$ a.e. Now apply Lemma 2, and the proof of the theorem is complete.

COROLLARY 1. *Let $s \geq 1$, and let g_0 denote the expectation of f. Then*

$$\int_{\Omega} | E^*(f \mid \cdot) - g_0 |^s \, d\mu \leqq \int_{\Omega} | f - g_0 |^s \, d\mu. \tag{6}$$

Equality holds

(i) *for $s = 1$, if and only if* sgn $[f - g_0]$ *is essentially a function of t;*

(ii) *for $s > 1$, if and only if f is essentially a function of t.*

This result expresses the domination over the s-th absolute central moment of the conditional expectation of f by the corresponding moment of f itself. It follows almost immediately from the theorem when we write (6) in the form

$$\int_{\Omega} | E^*(f - g_0 \mid \cdot) |^s \, d\mu \leqq \int_{\Omega} E^*(| f - g_0 |^s \mid \cdot) \, d\mu. \tag{7}$$

Thus, from the theorem we know that the integrand of the left-hand side of (7) is a.e. $\leqq$ the integrand on the right. Hence (7) holds. Equality in (7) holds then if and only if the integrands are a.e. equal. The theorem therefore directly provides the equality conditions as stated.

Let $W = \{\mu_\theta, \theta \in \Theta\}$ be a family of probability measures on $\mathfrak{F}$; and t, a sufficient statistic for W (cf. [5], p. 232, §5). Let f be an unbiased estimate of the function g^2 on Θ. For each $\mu_\theta \epsilon W$, the conditional expectation, $E_\theta(f \mid \cdot)$, of f with respect to t is defined. Since conditional expectations are fully determined by conditional probabilities (although, in general, not as usual integrals. Cf. [4], pp. 48, 49; also [5], p. 230) it follows from the sufficiency of t that there exists a function $E(f \mid \cdot)$, on Γ, with $E_\theta(f \mid \tau) = E(f \mid \tau)$ a.e. (ν_θ) for each $\theta \epsilon \Theta \cdot E^*(f \mid \cdot)$ is again an unbiased estimate of g, and we have

COROLLARY 2. *Let t be a sufficient statistic for the family $W = \{\mu_\theta, \theta \epsilon \Theta\}$; and f, an unbiased estimate of g. For $s \geqq 1$, and each $\theta \epsilon \Theta$,*

$$\int_\Omega | E^*(f \mid \cdot) - g(\theta) |^s d\mu_\theta \leqq \int_\Omega | f - g(\theta) |^s d\mu_\theta .$$

Equality holds

(i) *for $s = 1$, if and only if* sgn $[f - g(\theta)]$ *is essentially (μ_θ) a function of t;*

(ii) *for $s > 1$, if and only if f is essentially (μ_θ) a function of t.*

REFERENCES

[1] E. W. BARANKIN, "Locally best unbiased estimates", *Annals of Math. Stat.*, Vol. 20 (1949), pp. 477–501.

[2] D. BLACKWELL, "Conditional expectation and unbiased sequential estimation", *Annals of Math. Stat.*, Vol. 18 (1947), pp. 105–110.

[3] L. M. GRAVES, *The Theory of Functions of Real Variables*, McGraw-Hill, 1946.

[4] A. KOLMOGOROFF, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Ergebnisse der Mathematik, Vol. 2 (1933).

[5] P. R. HALMOS AND L. J. SAVAGE, "Application of the Radon-Nikodym Theorem to the theory of sufficient statistics", *Annals of Math. Stat.*, Vol. 20 (1949), pp. 225–241.

NOTE ON CONSISTENT ESTIMATES OF THE LINEAR STRUCTURAL RELATION BETWEEN TWO VARIABLES[1]

BY ELIZABETH L. SCOTT

University of California, Berkeley

1. Introduction. The purpose of this note is to present another case in which the structural linear relation between two observable random variables may be consistently estimated. Of the recent papers on this subject I wish to mention the paper by Wald [1], which contains a history of the work done on the problem, and the more recent paper by Housner and Brennan [2]. Also relevant is the important result due to Reiersøl [3], [4].

2. Statement of problem. Assume that the two observable random variables x and y have the structure

[1] Paper prepared with partial support of the Office of Naval Research.

The results summarized were presented in a discussion held at the Cleveland Meeting of the Institute of Mathematical Statistics, December, 1948.