NOTES

CONDITIONAL EXPECTATION AND THE EFFICIENCY OF ESTIMATES

By Paul G. Hoel

University of California, Los Angeles

1. Summary. A probability density function, $f(x; \theta)$, is considered for which there exists a sufficient statistic. It is shown, under certain regularity conditions on the family of distributions and on the class of estimates, that if there exists an unbiased sufficient estimate of θ, it will be unique. This result is used to show that when the regularity conditions are satisfied, the method of Blackwell for improving an unbiased estimate of θ merely yields a natural estimate.

2. Distribution of a sufficient statistic. For the purpose of proving the uniqueness of an unbiased sufficient estimate of θ, it is helpful to find the form of the probability density function of a particular sufficient statistic from a knowledge of the form of $f(x; \theta)$. It has been shown [1], [2] under different sets of assumptions that when a sufficient statistic exists, $f(x; \theta)$ must possess the functional form

$$f(x; \theta) = \exp [g(\theta) + h(\theta)r(x) + s(x)], \tag{1}$$

provided the range of x does not depend on θ.

Assumption 1. *It will be assumed that $f(x; \theta)$ has the form given by* (1).

Koopman [1] proves (1) under the assumption that $f(x; \theta)$ is analytic. Pitman [2] assumes only that $\partial^2 f(x; \theta)/\partial x \partial \theta$ exists but adds a differentiability condition on the density function of the sufficient statistic that is assumed to exist.

Now consider the distribution of a particular statistic. From (1), the probability density function for a random sample is

$$f(x_1, \cdots, x_n; \theta) = \exp [ng(\theta) + h(\theta) \sum_{i=1}^{n} r(x_i) + \sum_{i=1}^{n} s(x_i)]. \tag{2}$$

The particular statistic to be considered here is $z = \Sigma r(x_i)$. From a lemma of Lehmann and Scheffé [3], mild regularity conditions will insure the existence of a transformation to new variables $z, t_2, \cdots, t_n$ such that the density function of the new variables may be expressed in the form

$$F(z, t_2, \cdots, t_n; \theta) = f(x_1, \cdots, x_n; \theta)/| J |, \tag{3}$$

where J is the Jacobian of the transformation and where the x's are replaced by their expressions in terms of the new variables. The essential regularity conditions here are that $r'(x) \neq 0$, except possibly on a set of measure zero, and that $r'(x)$ is continuous.

ASSUMPTION 2. *It will be assumed that the lemma conditions are satisfied.*

If the assumptions made in [1] had been employed, these restrictions on $r(x)$ would have been satisfied because then $r(x)$ would be analytic.

If (2) is substituted in (3), and then (3) is integrated over the range of the variables $t_2, \cdots, t_n$, the density function of z will be obtained in the form

$$p(z;\theta) = L(z)\exp[ng(\theta) + h(\theta)z], \tag{4}$$

because $\exp[\Sigma s(x_i)]/|J|$ does not involve θ and thus its integral over the range of the t's will be a function of z only.

It is easily seen that z is a sufficient statistic because it suffices to show that $f(x_1, \cdots, x_n; \theta)/p(z;\theta)$ is independent of θ. From (2) and (4) it is clear that this ratio is independent of θ.

3. Relationship of sufficient estimates. If an estimate of θ is understood to be a single-valued function of the random variables $x_1, \cdots, x_n$, and if certain derivatives exist, then any sufficient estimate will be a function of z. For, let y be any sufficient estimate. Then

$$f(x_1, \cdots, x_n; \theta) = G(y;\theta)H(x_1, \cdots, x_n), \tag{5}$$

where $G(y;\theta)$ is the density function of y. If $\partial \log f/\partial\theta$ is calculated for both (2) and (5) and if the results are equated, it will follow that

$$\frac{\partial \log G(y;\theta)}{\partial\theta} = ng'(\theta) + h'(\theta)z.$$

This result shows that z is a single-valued function of y, when the derivatives exist. Conversely, since only single-valued functions of the variables $x_1, \cdots, x_n$ are considered as estimates, y will be a single-valued function of z. If the relationship is $z = T(y)$ and $y = T^{-1}(z)$ is multiple-valued, y will be defined only on one branch.

ASSUMPTION 3. *It will be assumed that $g'(\theta)$, $h'(\theta)$, and $\partial G(y;\theta)/\partial\theta$ exist for some value of θ.*

The restriction that $g'(\theta)$ and $h'(\theta)$ should exist would be satisfied if the assumptions made in either [1] or [2] to arrive at formula (1) had been made instead of Assumption 1. The restriction that $\partial G(y;\theta)/\partial\theta$ should exist is a weak restriction on the class of sufficient estimates, y, being considered.

4. Uniqueness. Suppose there exists an unbiased sufficient estimate of θ. From the preceding section it will be a function of z, say $w(z)$. Then, from (4),

$$\int_{-\infty}^{\infty} w(z)L(z)e^{ng(\theta)+h(\theta)z}dz = \theta. \tag{6}$$

Now suppose that there were two such estimates, say $w_1(z)$ and $w_2(z)$. Then, letting $\alpha = h(\theta)$, (6) would yield the following relationship:

$$\int_{-\infty}^{\infty} w_1(z)L(z)e^{\alpha z}dz = \int_{-\infty}^{\infty} w_2(z)L(z)e^{\alpha z}dz. \tag{7}$$

But from the theory of Laplace transforms ([4], p. 244), it follows that (7) implies that $w_1(z) = w_2(z)$ except on sets of measure zero, provided that $w_1(z)L(z)$ and $w_2(z)L(z)$ are integrable in every finite interval and provided that (7) holds for some interval of values of α. It is easily seen that the existence of (6) for all admissible values of θ insures the integrability condition. If $f(x; \theta)$ is defined for an interval of values of θ, (6) will hold for such an interval. Since, from Assumption 3, $\alpha = h(\theta)$ is continuous and from (1) is obviously not constant, α must exist for an interval of values also, and hence (7) will hold for such an interval.

ASSUMPTION 4. *It will be assumed that $f(x; \theta)$ is defined for some interval of values of θ.*

The discussions of the preceding sections may be summarized in the following theorem.

THEOREM. *If Assumptions 1, 2, 3, and 4 are satisfied and if there exists an unbiased sufficient estimate of θ, it will be the unique such estimate.*

5. Efficiency. Let t be any unbiased estimate of θ and let u be any sufficient estimate of θ. Then Blackwell [5] has shown that

$$v = E[t \mid u], \tag{8}$$

which is the conditional expected value of t for u fixed, determines an unbiased estimate of θ whose variance cannot exceed that of t. If t is not a function of u, the variance of v will be less than that of t.

This device for improving the efficiency of an unbiased estimate appears promising. However, since v is a function of u and thus is a sufficient unbiased estimate of θ, and since it has been shown that, subject to mild regularity conditions, any unbiased sufficient estimate is unique, this device will merely yield the unique unbiased sufficient estimate. Since the statistic z may be found by inspecting $f(x; \theta)$, it suffices to find the function $w(z)$ which is unbiased in order to obtain the desired estimate, when such an estimate exists. From (8), the existence of an unbiased estimate insures the existence of an unbiased sufficient estimate. Since, from (2), the maximum likelihood estimate of θ is a function of z, a natural method for finding this unique estimate when it exists is to first find the maximum likelihood estimate and then, if necessary, determine what function of it will be unbiased. Formula (8) with u chosen as z could also be used to find this unique estimate.

REFERENCES

[1] B. O. KOOPMAN, "On distributions admitting a sufficient statistic," *Trans. Am. Math. Soc.*, Vol. 39 (1936), p. 399.
[2] E. J. G. PITMAN, "Sufficient statistics and intrinsic accuracy," *Proc. Cambridge Philos. Soc.*, Vol. 32 (1936), p. 567.
[3] E. L. LEHMANN AND H. SCHEFFÉ, "On the problem of similar regions," *Proc. Nat. Acad. Sci.*, Vol. 33 (1947), p. 383.
[4] D. V. WIDDER, *The Laplace Transform*, Princeton University Press, 1946.
[5] D. BLACKWELL, "Conditional expectation and unbiased sequential estimation," *Annals of Math. Stat.*, Vol. 18 (1947), p. 105.