$$\sum_{j=1}^{k} F_n(X_{j-1,k}) I_{A_j} \leqq F_n(x) \leqq \sum_{j=1}^{k} F_n(X_{jk} - 0) I_{A_j} \tag{4}$$

Inequality (6) should be replaced by

$$\begin{aligned} F(x \mid \mathfrak{Z}) - F_n(x) &\leqq \sum_{j=1}^{k} (F(X_{jk} - 0 \mid \mathfrak{Z}) - F_n(X_{j-1,k})) I_{A_j} \\ &= \sum_{j=1}^{k} (F(X_{jk} - 0 \mid \mathfrak{Z}) - F(X_{j-1,k} \mid \mathfrak{Z})) I_{A_j} \\ &\quad + \sum_{j=1}^{k} (F(X_{j-1,k} \mid \mathfrak{Z}) - F_n(X_{j-1,k})) I_{A_j} \\ &\leqq \max_{1 \leqq j \leqq k} \mid F_n(X_{jk}) - F(X_{jk} \mid \mathfrak{Z}) \mid + 1/k. \end{aligned} \tag{6}$$

Inequality (7) should be replaced by

$$F(x \mid \mathfrak{Z}) - F_n(x) \geqq -\max_{1 \leqq j \leqq k} \mid F_n(X_{jk} - 0) - F(X_{jk} - 0) \mid \mathfrak{Z}) \mid - 1/k. \tag{7}$$

Inequality (8) should be replaced by

$$\begin{aligned} \mid F_n(x) - F(x \mid \mathfrak{Z}) \mid \leqq 1/k + \max_{1 \leqq j \leqq k} \{ \mid F_n(X_{jk} - 0) \\ - F(X_{jk} - 0 \mid \mathfrak{Z}) \mid, \mid F_n(X_{jk}) - F(X_{jk} \mid \mathfrak{Z}) \mid \}. \end{aligned} \tag{8}$$

Immediately after inequality (8) the following sentence should be added: In a way similar to the proof on the bottom of page 829 one may easily verify that $P[F_n(X_{jk} - 0) \underset{n}{\rightarrow} F(X_{jk} - 0 \mid \mathfrak{Z})] = 1$.

CORRECTION TO "ON THE THEORY OF BAN ESTIMATES"[1]

By Robert A. Wijsman

University of Illinois

I am greatly indebted to Dr. Lucien LeCam for calling to my attention an error in the proof of Theorem 1 of the paper cited in the title (*Ann. Math. Stat.* Vol. 30 (1959), pp. 185–191). The transition from (12) to (13) is in general not justified. Worse, the theorem itself is false in general, as can be shown with a counter example. In order to remedy the situation, the assumptions have to be strengthened. This can be done either on the distributions of the Z_n, or on the estimator $\hat{\theta}$. As an example of the first, if the Z_n have densities which (when normalized) converge a.e. to the limiting normal density, then the transition

[1] Work supported by the National Science Foundation.

from (12) to (13) is valid [9] and with that the proof of Theorem 1 is correct. However, this seems too strong an assumption to be of much practical value, since so many examples deal with discrete random variables. Turning now to assumptions on $\hat{\theta}$, we could require (4) to be true for all sequences Z_n for which (1) holds. Taking then in particular $Z_n : N(\zeta(\theta), \Sigma(\theta)/n)$, the previous case (convergence of densities) applies, and the conclusion of Theorem 1 follows. A more attractive, even though slightly stronger, assumption on $\hat{\theta}$ is to require it to be differentiable in every point of U. This insures, of course, continuity in every point of U but not continuity in a neighborhood of U, leave alone differentiability in a neighborhood of U which would be the requirement for a regular (1) estimate. We are thus led to the following modification of Definition 2 and regular (2):

DEFINITION 3. $\hat{\theta}$ will be called regular (3) if (i) $\hat{\theta}(\zeta(\theta)) \equiv \theta$ identically in θ[2]; (ii) $\hat{\theta}$ is differentiable in every point $\zeta(\theta)$ of U.

Let the matrix derivative of $\hat{\theta}$ in the point $\zeta(\theta)$ be denoted by $A(\theta)$. Theorem 1 now follows immediately by differentiation of (i) of Definition 3 (which is the same as equation (2)). A few remarks about $A(\theta)$ are in order. In the first place, the existence of this derivative in every point of U implies (4) for every sequence Z_n satisfying (1). Secondly, it is not necessary to require A to be continuous in θ. However, if $\hat{\theta}$ is constructed according to Theorem 2, then $A = (BV)^{-1}B$ (see eq. (6)) so that A is continuous due to the continuity assumptions on B and V. Under all circumstances, the A corresponding to any BAN estimate is continuous since it is given by $A = (V'\Sigma^{-1}V)^{-1}V'\Sigma^{-1}$.

It is somewhat remarkable that Theorem 2 remains true if, in the conclusion, regular (2) is replaced by the stronger regular (3). The surprise is that $\hat{\theta}$ turns out to be differentiable in each point of U, even though no differentiability assumptions are made on B. Therefore, a proof of Theorem 2, with regular (2) replaced by regular (3), seems to be in order. Before doing this, it may be of interest to point out that Ferguson's estimates [5] are also differentiable in each point of U since they are generated by (5) with $B(z, \theta)$ satisfying even stronger assumptions than in Theorem 2. Comparing now the various kinds of regular estimates, we have that regular (1) estimates are continuously differentiable in a neighborhood of U, Ferguson's estimates are continuous in a neighborhood of U and differentiable in every point of U, while regular (3) estimates are differentiable in every point of U.

PROOF OF THEOREM 2, with regular (2) replaced by regular (3). It suffices to show that in each point of U there is a neighborhood possessing the properties ascribed to the neighborhood N in the conclusion of Theorem 2. Then N can be taken as the union of the individual neighborhoods. Consider any point of U. We may take this as the origin of the coordinate system in Z. For the purpose of the proof we may make the same transformations as in Section 4 (observe

[2] The assumption (i) of Definition 3 is the same as equation (2). Instead, we could have made the same assumption as in Definition 1 (i). The two assumptions are equivalent since $\hat{\theta}$ is supposed to be continuous in each point of U.

that ζ^{-1} is differentiable due to Assumption 2 (iii) and (iv)). We may suppose then that U is a linear subspace of Z, spanned by the first m coordinate axes, and that ζ is the identity function from U to U. Thus we have identified U with the parameter space Ω. A point u of U has its last $k - m$ components equal to 0; the m-vector formed by its first m components will be written θ. Let I_{km} be a $k \times m$ matrix whose elements are 1 on the "main diagonal" and 0 otherwise. We can write then $u = I_{km}\theta$. The transformations which we have employed replace in (5) $\zeta(\theta)$ by $I_{km}\theta$, and $B(z, \theta)$ by some other matrix, which, however, we shall again denote by $B(z, \theta)$. The matrix $V(0)$ is replaced by I_{km}. Put $B(z, \theta)\, I_{km} = C(z, \theta)$, then by assumption $C(0, 0)$ is non-singular. Furthermore, C is continuous in (z, θ) at $(0, 0)$. Put $C^{-1}B = D$, then $D(z, \theta)$ exists in a neighborhood of $(0, 0)$, is continuous in (z, θ) at $(0, 0)$ and is continuous in θ for each fixed z. Let $S_1 \times S_2$ be such a neighborhood, where S_1 is a solid k-sphere about $z = 0$ and S_2 a solid m-sphere about $\theta = 0$. In addition, we may choose the radii r_1 and r_2 of S_1 and S_2 so that for $(z, \theta) \varepsilon S_1 \times S_2$ we have $\| D(z, \theta) \| \leqq r_2/r_1$. We now write (5) as

$$\theta = D(z, \theta)z. \tag{25}$$

For each $z \varepsilon S_1$, the right hand side of (25) is a continuous transformation of S_2 into itself. According to the Brouwer fixed point theorem [7] there is a fixed point of the transformation, therefore a solution $\hat{\theta}(z)$ to (25). Write

$$\hat{\theta}(z) = D(z, \hat{\theta}(z))z. \tag{26}$$

For $z \varepsilon S_1$, $\| D(z, \hat{\theta}(z)) \|$ is bounded, so $\hat{\theta}(z) \to 0$ as $z \to 0$. Hence $\hat{\theta}$ is continuous at 0. From this we have $D(z, \hat{\theta}(z)) \to D(0, 0)$ as $z \to 0$, and from (26) it follows then that $\hat{\theta}$ is differentiable at $z = 0$, with matrix derivative $D(0, 0)$. This proves that on S_1 $\hat{\theta}$ is regular (3). In the original coordinate system the matrix $D(0, 0)$ takes the form $(BV)^{-1}B$, evaluated at some point $(\zeta(\theta), \theta)$. This leads immediately to (6). The last assertion in the conclusion of Theorem 2 is proved in [3].

REFERENCE

[9] Henry Scheffé, "A useful convergence theorem for probability distributions," *Ann. Math. Stat.*, Vol. 18 (1947), pp. 434–438.