To do this we use the inequality established by McMillan [2],

$$\text{(i)} \qquad \int_{\{m \leqq g_k < m+1\}} g_k \leqq s(m+1)2^{-m}.$$

We confine our attention to the cylinder set $Z_i \subset \Omega$, $Z_i = \{\omega;\, x_0 = a_i\}$. On Z_i we have

$$g_k(\omega) = -\log_2 p(x_0 = a_i \mid x_{-1}, \cdots, x_{-k}).$$

Since $p(x_0 = a_i \mid x_{-1}, \cdots, x_{-k})$ is a martingale, it follows from the convexity of $-\log$ and inequality (i) that the sequence $\{g_k\}$ is a semi-martingale (see [3], p. 295). Therefore, g_k converges a.s. on Z_i and hence on Ω.

Furthermore, by a semi-martingale inequality, [3] p. 317, we have, on Z_i,

$$\int_{Z_i} \left(\sup_{0 \leqq k \leqq n} g_k \right) \leqq \frac{e}{e-1} + \frac{e}{e-1} \int_{Z_i} (g_n \log^+ g_n).$$

By using inequality (i) again, we bound the last term on the above right;

$$\int_{Z_i} (g_n \log^+ g_n) = \sum_{m=0}^{\infty} \int_{Z_i\{m \leqq g_n < m+1\}} (g_n \log^+ g_n) \leqq \sum_{m=0}^{\infty} s(m+1) \log (m+1) 2^{-m}.$$

Therefore $\int_{Z_i} (\sup_k g_k) < \infty$, by addition $E(\sup_k g_k) < \infty$, and the theorem is proved.

It is a pleasure to acknowledge our debt to Professor David Blackwell who suggested to us the problem treated herein.

REFERENCES

[1] C. E. SHANNON, "A mathematical theory of communication," *Bell System Technical Journal*, Vol. 27 (1948), pp. 379–423, pp. 623–656.

[2] B. MCMILLAN, "The basic theorems of information theory," *Ann. Math. Stat.*, Vol. 24 (1953), pp. 196–219.

3] J. L. DOOB, *Stochastic Processes*, John Wiley & Sons, Inc., New York, 1953.

A COUNTEREXAMPLE TO A THEOREM OF KOLMOGOROV[1,2]

BY LEO BREIMAN

University of California, Berkeley

1. Introduction. In 1928 Kolmogorov [1] presented the now well-known degenerate convergence theorem (weak law of large numbers) as follows (see, for

Received August 10, 1956; revised January 18, 1957.

[1] This paper was prepared with the support of the Office of Ordnance Research, U. S. Army under Contract DA-04-200-ORD-171.

[2] After this note was submitted the author was informed that C. Derman had constructed a similar counterexample. While the note was in proof, a similar counterexample appeared in a paper by Hartley Rogers, Jr., *Proc. Am. Math. Soc.*, Vol. 8 (1957), pp. 518–520.

example, Loève [2]): let $X_1, X_2, \cdots$ be independent random variables such that $EX_k = 0$, $k = 1, 2, \cdots$, and let

$$X_{nk} = \begin{cases} X_k & \text{if} \quad |X_k| < n, \\ 0 & \text{if} \quad |X_k| \geqq n. \end{cases}$$

Then

$$\frac{1}{n} \sum_{k=1}^{n} X_k \xrightarrow{P} 0$$

if and only if

(i) $\sum_{k=1}^{n} P(|X_k| \geqq n) \to 0$,

(ii) $\frac{1}{n} \sum_{k=1}^{n} EX_{nk} \to 0$,

(iii) $\frac{1}{n^2} \sum_{k=1}^{n} \sigma^2 X_{nk} \to 0$.

Presented without proof in the same paper was a sharpened version of the above theorem with condition (iii) replaced by

(iii′) $\frac{1}{n^2} \sum_{k=1}^{n} EX_{nk}^2 \to 0$.

A proof of this last theorem was given in Gnedenko and Kolmogorov's 1949 book and carried over into the English edition ([3], pp. 135–137). Unfortunately, the proof contains a slight gap and the sharpened theorem is not correct. Since it appears in several places in the literature, for example, in Loève ([2], p. 278), and follows from Theorems 3.2 (p. 124) and 3.3 (p. 125) of Doob's book [4] the following simple counterexample may be of interest to the reader:

We will show that conditions (i), (ii), (iii′) are not necessary by proceeding as follows: define the independent random variables $X_1, X_2, \cdots$ by

$$P(X_1 = 0) = 1,$$

$$\left.\begin{aligned} P(X_k &= (-1)^k k^{5/2}) = k^{-2}, \\ P(X_k &= (-1)^{k+1} k^{1/2}(1 - k^{-2})^{-1}) = 1 - k^{-2}, \end{aligned}\right\} \qquad k \geqq 2.$$

We verify immediately that $EX_k = 0$, $k = 1, 2, \cdots$. Then we demonstrate that conditions (i), (ii), (iii) above are satisfied. Finally we show that, contrary to the theorem,

$$\frac{1}{n^2} \sum_{k=1}^{n} EX_{nk}^2 \nrightarrow 0.$$

In the following proofs we take $n \geqq 4$.

Proof of (i). If $k^{5/2} < n$, then $X_{nk} = X_k$, and if $k \leqq n$, then $k^{1/2}(1 - k^{-2})^{-1} < n$. Hence

$$P(|X_k| \geqq n) = \begin{cases} 0 & \text{if} \quad 1 \leqq k^{5/2} < n, \\ k^{-2} & \text{if} \quad n \leqq k^{5/2} \quad \text{and} \quad k \leqq n, \end{cases}$$

and

$$\sum_{k=1}^{n} P(|X_k| \geqq n) = \sum_{k=[n^{2/5}]}^{n} \frac{1}{k^2} \to 0,$$

where $[\cdot]$ denotes next higher integer.

Proof of (ii).

$$EX_{nk} = \begin{cases} 0 & \text{if } 1 \leqq k^{5/2} < n \\ (-1)^{k+1}k^{1/2} & \text{if } n \leqq k^{5/2} \text{ and } k \leqq n. \end{cases}$$

Hence

$$\frac{1}{n}\sum_{k=1}^{n} EX_{nk} = \frac{1}{n}\sum_{k=[n^{2/5}]}^{n} (-1)^{k+1}k^{1/2}.$$

We use the inequality, valid for $s > 1$,

$$\sqrt{s} - \sqrt{s-1} < \frac{1}{2\sqrt{s-1}}$$

to get

$$\left|\frac{1}{n}\sum_{k=1}^{n} EX_{nk}\right| \leqq \left(\frac{1}{2n}\sum_{k=1}^{n-1}\frac{1}{\sqrt{k}} + \frac{1}{n}\sqrt{[n^{2/5}]}\right) \to 0.$$

Proof of (iii).

$$\sigma^2 X_{nk} = \begin{cases} k^3 + k(1-k^{-2})^{-1} & \text{if } 2 \leqq k^{5/2} < n, \\ k(1-k^{-2})^{-1} - k & \text{if } n \leqq k^{5/2} \text{ and } k \leqq n. \end{cases}$$

For $k \geqq 2$,

$$k^3 + k(1-k^{-2})^{-1} \leqq k^3 + \tfrac{4}{3}k \leqq 2k^3,$$
$$k(1-k^{-2})^{-1} - k = k^{-1}(1-k^{-2})^{-1} \leqq \tfrac{4}{3}k^{-1}.$$

Hence

$$\frac{1}{n^2}\sum_{k=2}^{n}\sigma^2 X_{nk} \leqq \frac{1}{n^2}\sum_{k=2}^{[n^{2/5}]} 2k^3 + \frac{1}{n^2}\sum_{k=[n^{2/5}]}^{n}\frac{4}{3k} \leqq \frac{1}{2n^2}([n^{2/5}])^4 + \frac{4}{3n^2}\log(n) \to 0.$$

Finally, we show that $1/n^2\left(\sum_{k=1}^{n} EX_{nk}^2\right) \nrightarrow 0$. We have

$$\frac{1}{n^2}\sum_{k=1}^{n} EX_{nk}^2 \geqq \frac{1}{n^2}\sum_{k=[n^{2/5}]}^{n} EX_{nk}^2 = \frac{1}{n^2}\sum_{k=[n^{2/5}]}^{n} k(1-k^{-2})^{-1} \geqq \frac{1}{n^2}\sum_{k=[n^{2/5}]}^{n} k \to \tfrac{1}{2},$$

which completes the counterexample.

It is a pleasure to be able to acknowledge our debt to M. Loève, who brought the question to our attention and suggested further inquiry. We are also indebted to R. K. N. Patell whose letter to M. Loève was the cause of the re-examination of this theorem.

REFERENCES

[1] A. Kolmogorov, "Ueber die Summen durch den Zufall bestimmter unabhangigen Grossen," *Math. Ann.*, Vol. 99 (1928), pp. 309–319.

[2] M. Loève, *Probability Theory*, Van Nostrand, New York, 1955.

[3] B. Gnedenko and A. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables*, Addison-Wesley, Cambridge, 1954.

[4] J. L. Doob, *Stochastic Processes*, John Wiley & Sons, New York, 1953.

ON THE COMBINING OF INTERBLOCK AND INTRABLOCK ESTIMATES

By D. A. S. Fraser

University of Toronto

In a recent paper Sprott [1] has considered methods for combining interblock and intrablock estimates of variety contrasts for incomplete block designs. The intrablock estimates are derived from treatment contrasts obtained within blocks. The interblock estimates presuppose that the block effects are random, independent, and identically distributed, and they are derived from contrasts among the block averages. Under normality the intrablock estimates are independent of the interblock estimates.

Sprott compares two methods for producing combined estimates. The first method, introduced by Yates [2], is the familiar method of combining by weighting with the reciprocal of the variances, and is known to produce minimum variance when two real estimates of the same quantity are combined linearly. The second method, discussed by Rao [3] and Cochran and Cox [4], is to apply the method of maximum likelihood to the joint density function, and the resulting estimate is linear in terms of the interblock and intrablock estimates. Sprott shows that, in general, the two methods are not equivalent. The second method is direct and has considerable theoretical weight behind its use. We are left then with the implication that one of the methods is incorrect for obtaining good estimates. In a sense this is not the case. Rather, one of the methods may be *inappropriately* applied. Weighting with reciprocal variances is *appropriate* to combining real estimates but if applied to vector estimates it ignores any covariances and may not be optimum.

Suppose $x = (x_1, \cdots, x_r)$ and $y = (y_1, \cdots, y_r)$ are independent estimates of the parameter $\eta = (\eta_1, \cdots, \eta_r)$ and have nonsingular covariance matrices V and W respectively. Also suppose, for the moment, that x and y are normal. Then, the joint density function is a constant times

$$\exp\left[-\tfrac{1}{2}(x - \eta)V^{-1}(x - \eta)' - \tfrac{1}{2}(y - \eta)W^{-1}(y - \eta)'\right],$$

Received October 29, 1956; revised January 10, 1957.