CORRECTION NOTES

CORRECTION TO

"THE INDIVIDUAL ERGODIC THEOREM OF INFORMATION THEORY"

By Leo Breiman

University of California at Los Angeles

Mr. James Abbott has pointed out that the argument on Page 811 of the above-cited work, *Ann. Math. Stat.*, Vol. 28, No. 3 (1957), pp. 809–811, is incorrect. The results of the paper are valid, however, and Page 811 may be replaced by the following discussion.

Note that

$$E\left(\frac{p(x_{-k}, \cdots, x_{-1})}{p(x_{-k}, \cdots, x_0)} \,\middle|\, x_0, \cdots, x_{-k+1}\right) \leqq \frac{p(x_{-k+1}, \cdots, x_{-1})}{p(x_{-k+1}, \cdots, x_0)}$$

with probability one. By the concavity of log, it follows that the g_k sequence,

$$g_k = -\log_2 \left(\frac{p(x_{-k}, \cdots, x_0)}{p(x_{-k}, \cdots, x_{-1})}\right)$$

satisfies

$$E(g_k \mid x_0, \cdots, x_{-k+1}) \leqq g_{k-1}.$$

Since $g_k \geqq 0$, and $Eg_0 < \infty$, the g_k sequence forms a non-negative lower semi-martingale and hence converges a.s. Actually, the convergence of the g_k sequence has been previously established by McMillan in [2].

Now consider $P(\sup_{k \leqq n} g_k > \lambda)$, and define the disjoint sets

$$E_j = \{g_j > \lambda, \ \sup_{k<j} g_k \leqq \lambda\},$$

whence $P(\sup_{k \leqq n} g_k > \lambda) = \sum_{j=1}^{n} P(E_j)$. Let Z_i be the cylinder sets $\{x_0 = a_i\}$ and $f_k^{(i)}$ the functions $-\log_2 P(x_0 = a_i \mid x_{-1}, \cdots, x_{-k})$. If $\sum_A f(x_0, x_{-1}, \cdots)$ indicates the sum of $f(x_0, x_{-1}, \cdots)$ over all sequences $(x_0, x_{-1}, \cdots) \,\varepsilon\, A$, then

$$P(E_j) = \sum_{E_j} p(x_{-j}, \cdots, x_0) = \sum_i \sum_{E_j \cap Z_i} \frac{p(x_{-j}, \cdots, x_0)}{p(x_{-j}, \cdots, x_{-1})} p(x_{-j}, \cdots, x_{-1}).$$

But on E_j we have the inequality

$$\frac{p(x_{-j}, \cdots, x_0)}{p(x_{-j}, \cdots, x_{-1})} = 2^{-g_j} \leqq 2^{-\lambda},$$

leading to

$$P(E_j) \leqq 2^{-\lambda} \sum_i \sum_{E_j \cap Z_i} p(x_{-j}, \cdots, x_{-1}) = 2^{-\lambda} \sum_i P(f_j^{(i)} > \lambda, \sup_{k<j} f_k^{(i)} \leqq \lambda).$$

Finally, then,

$$P\,(\sup_{k \leqq n} g_k > \lambda) \leqq 2^{-\lambda} \sum_i P\,(\sup_{k \leqq n} f_k^{(i)} > \lambda) \leqq s \cdot 2^{-\lambda},$$

where s is the number of values that the process ranges over. This last inequality gives $P(\sup_k g_k > \lambda) \leqq s \cdot 2^{-\lambda}$, which quickly leads to $E(\sup_k g_k) < \infty$.

CORRECTION TO

"BOUNDS ON NORMAL APPROXIMATIONS TO STUDENT'S AND THE CHI-SQUARE DISTRIBUTIONS"

BY DAVID L. WALLACE

University of Chicago

The following correction should be made on p. 1127 of the above-titled article (*Ann. Math. Stat.*, Vol. 30 (1959), pp. 1121–1130): In the conclusion of Corollary 2 to Theorem 4.2, the exponent of n should be $-\frac{1}{2}$ and not $\frac{1}{2}$.