*The Annals of Mathematical Statistics*
1969, Vol. 40, No. 4, 1492–1495

# STRONG CONSISTENCY OF CERTAIN SEQUENTIAL ESTIMATORS[1]

By Robert H. Berk[2]

*Hebrew University*

**1. Introduction and summary.** Motivated by Loynes' (1969) treatment of (weak) consistency of sequential estimators, we establish here some allied results on strong consistency. The strengthened conclusion is achieved by imposing further restrictions, so that our results are not as broadly applicable as Loynes'. The reader is referred to Loynes' paper for additional motivation and discussion.

As in that paper, we are concerned with estimators that improve a given one by taking conditional expectations with respect to a sufficient statistic, in the case where the sample size may itself be a random variable. $x_1, x_2, \cdots$ will denote the data sequence, random variables defined on a measurable space $(\Omega, \mathcal{A})$. All probability measures considered on $(\Omega, \mathcal{A})$ will render the sequence iid. It is seen below (Theorem 3.4) that a given sequence $\{t_i\}$ of stopping times for the data sequence leads to a strongly consistent sequence of estimators if $\lim t_i = +\infty$ a.s., and the $t_i$ are $C$-ordered (Definition 2.3). This entails that the $t_i$ increase monotonically to $+\infty$ a.s., but requires additional structure as well. The specific considerations are in Section 3; Section 2 presents some general notions.

**2. Generalities.** Let $\{\mathcal{A}_n : 1 \leq n \leq \infty\}$ be an increasing sequence of subfields (of $\mathcal{A}$) with $\mathcal{A}_\infty = \bigvee_n \mathcal{A}_n$. Let $\{\mathcal{B}_n : 1 \leqq n \leqq \infty\}$ be another sequence of subfields and $\{\mathcal{C}_n : 1 \leqq n \leqq \infty\}$ be a decreasing sequence of subfields with $\bigcap_n \mathcal{C}_n = \mathcal{C}_\infty$. A random index is a positive (extended) integer-valued random variable on $(\Omega, \mathcal{A})$. These will be denoted by $M, N$, etc.

2.1 Definition. $\mathcal{B}_N = \{\bigcup_n B_n(N = n) : B_n \varepsilon \mathcal{B}_n\}$.

Remark. Note that $N$ is $\mathcal{B}_N$ measurable. If $\mathcal{B}_n$ is generated by a statistic $v_n$, $\mathcal{B}_N = \mathcal{B}(N, v_N)$, where $v_N = \sum v_n I_{(N=n)}$.

2.2 Proposition. *If* $f \varepsilon L_1(\Omega, \mathcal{A}, P)$,

$$E_P(f|\mathcal{B}_N) = \Sigma\{E_P(fI_{(N=n)}|\mathcal{B}_n)/E_P(I_{(N=n)}|\mathcal{B}_n)\}I_{(N=n)}[P].$$

Proof. Denoting the RHS by $f_N$, for $B_n \varepsilon \mathcal{B}_n$,

$$\begin{aligned}\int_{(N=n)B_n} f_N &= \int \{E_P(fI_{(N=n)}|\mathcal{B}_n)/E_P(I_{(N=n)}|\mathcal{B}_n)\}I_{(N=n)} \\ &= \int_{B_n}\{E(fI_{(n=n)}|\mathcal{B}_n)/P(N = n|\mathcal{B}_n)\}P(N = n|\mathcal{B}_n) \\ &= \int_{B_n} fI_{(N=n)} = \int_{(N=n)B_n} f. \ \square\end{aligned}$$

2.3 Definition. $M \leqq N$ are C-ordered if also, $\mathcal{C}_N \subset \mathcal{C}_M$.

2.4 Proposition. *If* $M \leqq N$, *the following are equivalent*:

(a) $N$ *is* $\mathcal{C}_M$ *measurable.*

(b) $M$ *and* $N$ *are C-ordered.*

Received 2 December 1968.

[1] Supported in part by NSF Postdoctoral fellowship 48097.

[2] On leave from The University of Michigan.

(c) $\forall k \leqq n, \exists C_{kn} \varepsilon \mathcal{C}_k$ *so that* $(N = n)(M = k) = C_{kn}(M = k)$.

PROOF. $(a) \Rightarrow (b)$. Choose $C \varepsilon \mathcal{C}_N$. For some $C_n \varepsilon \mathcal{C}_n$, $C(N = n) = C_n(N = n) = C_n(N = n)(M \leqq n) = [(N = n)] \cap [C_n(M \leqq n)] \varepsilon \mathcal{C}_M$ since each of the sets in brackets does. Thus $C \varepsilon \mathcal{C}_M$.

$(b) \Rightarrow (c)$. $\mathcal{C}_N \subset \mathcal{C}_M \Rightarrow N$ is $\mathcal{C}_M$ measurable. Thus for all $k$, $n$, $\exists C_{kn} \varepsilon \mathcal{C}_k$ so that $(N = n)(M = k) = C_{kn}(M = k)$.

$(c) \Rightarrow (a)$. For $k \leqq n$, $(N = n)(M = k) = C_{kn}(M = k) \varepsilon \mathcal{C}_M$. For $k > n$, $(N = n)(M = k) = \emptyset$. Hence $(N = n) \varepsilon \mathcal{C}_M \Rightarrow N$ is $\mathcal{C}_M$ measurable. □

2.5 COROLLARY. *If* $(N > n) \varepsilon \mathcal{C}_n$ *for all* $n$, *then for all* $M \leqq N$, $M$ *and* $N$ *are* $C$*-ordered.*

PROOF. For $k \leqq n$, $(N \leqq n)(M = k) \varepsilon \mathcal{C}_M$, while for $k >$ n, $(N \leqq n)(M = k) = \emptyset$. Thus $N$ is $\mathcal{C}_M$ measurable. □

REMARK. The condition in 2.5 is usually stated that $N$ is a reverse stopping time on the sequence $\{\mathcal{C}_n\}$. Another criterion for $C$-ordering is given in Section 3 (see Theorem 3.5).

2.6 DEFINITION. A collection $\{N_j : j \varepsilon J\}$ of random indices is $C$-ordered if every pair selected from it is $C$-ordered.

REMARK. Clearly, in order that $N_1 \leqq N_2 \leqq \cdots$ be $C$-ordered, it is necessary and sufficient that for $i = 1, 2, \cdots$, $N_i$ and $N_{i+1}$ be $C$-ordered.

2.7 THEOREM. *If* $N_1 \leqq N_2 \leqq \cdots$ *is* $C$*-ordered and* $N_\infty = \lim N_i$, *then* $\{N_i : 1 \leqq i \leqq \infty\}$ *is* $C$*-ordered and* $\mathcal{C}_{N_i} \downarrow \mathcal{C}_{N_\infty}$.

PROOF. By the $C$-ordering, it is clear that $\mathcal{C}_{N_i}$ decreases to $\mathcal{C}^*$, say. If $C \varepsilon \mathcal{C}_{N_\infty}$, for some $C_n \varepsilon \mathcal{C}_n$, $C(N_\infty = n) = C_n(N_\infty = n) = \lim_i C_n(N_i = n) \varepsilon \mathcal{C}^*$. Thus $\mathcal{C}_{N_\infty} \subset \mathcal{C}^*$, which in turn implies that $\{\mathcal{C}_i : 1 \leqq i \leqq \infty\}$ is $C$-ordered.

Choose $C \varepsilon \mathcal{C}^*$. $\forall i \, \exists C_{ni} \varepsilon \mathcal{C}_n$ so that $C(N_i = n) = C_{ni}(N_i = n)$. $C(N_i = n) \to C(N_\infty = n)$, thus $\lim_i I_{C_{ni}} = 1$ on $C(N_\infty = n)$. Let $C' = \Omega - C$. $C'(N_i = n) = C'_{ni}(N_i = n)$, so that $\lim I_{C'_{ni}} = 1$ on $C'(N_\infty = n)$. Or,

$$\begin{aligned} \lim I_{C_{ni}} &= 1 \quad \text{on} \quad C(N_\infty = n), \\ &= 0 \quad \text{on} \quad C'(N_\infty = n). \end{aligned}$$

Let $C_n = \limsup_i C_{ni} \varepsilon \mathcal{C}_n$. $C(N_\infty = n) = C_n(N_\infty = n) \epsilon \mathcal{C}_{N_\infty}$. Thus $C \varepsilon \mathcal{C}_{N_\infty}$ and hence $\mathcal{C}^* = \mathcal{C}_{N_\infty}$. □

**3. Strong consistency.** Let $z_n$ be the order statistic generated by

$$(x_1, \cdots, x_n) : z_n \equiv \{x_1, \cdots, x_n\}.$$

Let $\mathcal{P}$ be a family of distributions on $(\Omega, \mathcal{A})$ and for every $n$, let $v_n$ be $\mathcal{B}(x_1, \cdots, x_n)$ measurable and sufficient for $(x_1, \cdots, x_n)$. We suppose $v_n$ satisfies.

CONDITION A. For all $n$, (i) $v_n$ is $\mathcal{B}(z_n)$ measurable and (ii) $v_{n+1}$ is $\mathcal{B}(v_n, x_{n+1})$ measurable.

REMARK. A(i) requires $v_n$ to be symmetric in $x_1, \cdots, x_n$, a condition usually met in applications. A(ii) implies that $\{v_n\}$ is a transitive sequence; see Bahadur (1954).

Let $\mathcal{A}_n = \mathcal{B}(x_1, \cdots, x_n)$, $\mathcal{B}_n = \mathcal{B}(v_n)$ and $\mathcal{C}_n = \mathcal{B}(v_n, x_{n+1}, x_{n+2}, \cdots)$. Under condition $A$, $\mathcal{C}_n \downarrow \mathcal{C}_\infty$, say, and by the Hewitt-Savage $0 - 1$ law, $\mathcal{C}_\infty$ is a.s. trivial $[\mathcal{P}]$.

3.1 LEMMA. *If* $T \epsilon L_1(\Omega, \mathcal{A}_n, \mathcal{P})$, $E(T \mid \mathcal{B}_n) = E(T \mid \mathcal{C}_n)$ $[\mathcal{P}]$.

PROOF. Follows from independence of $\mathcal{A}_n$ and $x_{n+1}, x_{n+2}, \cdots$. □

3.2 DEFINITION. A random index $t$ is a stopping time on $\{\mathcal{C}_n\}$ if $(t = n) \epsilon \mathcal{A}_n$, $\forall_n$. All stopping times will be on $\{\mathcal{A}_n\}$.

REMARK. If $t$ is a stopping time, $\mathcal{A}_t = \{A \epsilon \mathcal{A} : A(t = n) \epsilon \mathcal{A}_n\}$.

3.3 COROLLARY. *If $t$ is a stopping time, for all* $T \epsilon L_1(\Omega, \mathcal{A}_t, \mathcal{P})$, $E_P(T \mid \mathcal{B}_t) = E_P(T \mid \mathcal{C}_t)$ *and neither expression depends on* $P \varepsilon \mathcal{P}$.

PROOF. $E_P(T \mid \mathcal{C}_t) = \Sigma\{E_P(T1_{(t=n)} \mid \mathcal{C}_n / P(t = n \mid \mathcal{C}_n)\} I_{(N=n)}[P]$. Since $TI_{(t=n)}$ and $I_{(t=n)}$ are $\mathcal{A}_n$ measurable, the conclusion follows from Lemma 3.1 and the sufficiency of $\mathcal{B}_n$. □

3.4 THEOREM. *Let* $s \leqq t_1 \leqq t_2 \leqq \cdots$ *be C-ordered stopping times so that* $\lim_i t_i = +\infty$ $[\mathcal{P}]$. *Then for all* $T \epsilon L_1(\Omega, \mathcal{A}_s, \mathcal{P})$ *and* $P \epsilon \mathcal{P}$, $E(T \mid \mathcal{B}_{t_i}) \rightarrow E_P T[P]$.

PROOF. We remark first that if $s \leqq t$ are stopping times, then $\mathcal{A}_s \subset \mathcal{A}_t$. Thus for all $i$, $T \epsilon L_1(\Omega, \mathcal{A}_{t_i}, \mathcal{P})$, so that by Corollary 3.3, $E(T \mid \mathcal{B}_{t_i}) = E(T \mid \mathcal{C}_{t_i})$ $[\mathcal{P}]$. By Theorem 2.7, $\mathcal{C}_{t_i} \downarrow \mathcal{C}_\infty \equiv (\emptyset, \Omega)$ $[\mathcal{P}]$; hence $E_P(T \mid \mathcal{C}_{t_i}) \rightarrow E_P T[P]$. □

Thus if the $t_i$ increase to $+\infty$ a.s. and are C-ordered, projecting $T$ onto the sufficient $\sigma$-fields $\mathcal{B}_{t_i} = \mathcal{B}(t_i, v_{t_i})$ produces a sequence of strongly consistent estimators of $E_p T$. If $s$ and the $t_i$ are not random, one obtains the strong consistency of the usual Blackwell-Rao estimators. If, in addition, $v_n = z_n$, $E(T \mid \mathcal{B}_{t_i})$ is just the $U$-statistic formed from $t_i$ observations, based on the kernel $T = T(x_1, \cdots, x_s)$. Strong consistency in this last case is already known: see Berk (1966).

For the more specialized structure of this section, we present a criterion that assures $C$-ordering and which is more useful than Corollary 2.5.

3.5 THEOREM. *Let* $\{V_{ni}, n = 1, 2, \cdots, i = 1, 2, \cdots\}$ *be a collection of measurable sets,* $V_{ni}$ *being a subset of range* $v_n$ *for all* $i$ *and so that* $V_{ni} \supset V_{n(i+1)}$ *for all* $n$ *and* $i$. *Let* $t_i$ *be the first* $n \geqq 1$ *so that* $v_n \varepsilon V_{ni}$ *or be* $+\infty$ *if no such* $n$ *occurs. Then* $t_1 \leqq t_2 \leqq \cdots$ *are C-ordered.*

PROOF. Since $V_{ni}$ is decreasing in $i$, it is clear that $t_i$ is increasing in $i$. Moreover, for $k \leqq n$, since $V_{ji} \supset V_{j(i+1)}$,

$$\begin{aligned}(t_i = k)(t_{i+1} = n) &= (v_1 \varepsilon V'_{1i}, \cdots, v_{k-1} \varepsilon V'_{(k-1)i}, v_k \varepsilon V_{ki}) \\ &\quad \cap (v_k \varepsilon V'_{k(i+1)}, \cdots, v_{n-1} \varepsilon V'_{(n-1)(i+1)}, v_n \varepsilon V_{n(i+1)}) \\ &= (t_i = k) C_{kn}, \qquad C_{kn} \varepsilon \mathcal{C}_{kn}.\end{aligned}$$

By Proposition 2.4, the $t_i$ are $C$-ordered. □

REMARK. We note that the above proof works equally well if

$$\mathcal{C}_n = \mathcal{B}(v_n, v_{n+1}, \cdots).$$

**4. Examples.** We illustrate the idea of $C$-ordered stopping times. We suppose the $x_i$ are real-valued and, for convenience, we suppress $\mathcal{P}$. (i) Let $t_i$ be the first $n \geqq 1$ so that $\max\{x_1, \cdots, x_n\} \geqq a_i$, or be $+\infty$ if no such $n$ occurs. We take $a_i$ increasing and $\mathcal{C}_n = \mathcal{B}(z_n, x_{n+1}, \cdots)$. Since $a_i$ increases, $t_i$ increases. Moreover, $(t_i > n) = (x_i < a_i, i = 1, \cdots, n) \varepsilon \mathcal{C}_n$, so that by Corollary 2.5, the $t_i$ are $C$-ordered.

(ii) Let $t_i$ be the first $n > i$ so that $x_n > \max\{x_1, \cdots, x_i\}$ or be $+\infty$ if no such $n$ occurs. Take $\mathcal{C}_n = \mathcal{B}(z_n, x_{n+1}, \cdots)$. Again, $t_i$ increases. If $t_i > i + 1$, $t_{i+1} = t_i$. Thus for $n > i + 1$, consider $(t_i = i + 1, t_{i+1} = n)$. $(t_{i+1} = n) \epsilon \mathcal{C}_{i+1}$, thus $(t_i = i + 1, t_{i+1} = n) \epsilon \mathcal{C}_{t_i}$, which implies that $t_{i+1}$ is $\mathcal{C}_{t_i}$ measurable. By Proposition 2.4, the $t_i$ are C-ordered. The first two examples are taken from Loynes (1969). The next is motivated by the stopping rule for the Wald SPRT; the last example, by a sequential procedure proposed by Chow and Robbins (1965).

(iii) Let $v_n = x_1 + \cdots + x_n$ and $t_i$ be the first $n \geqq 1$ so that $|v_n| \geqq a_i$, or be $+\infty$ if no such $n$ occurs. We take $a_i$ increasing and $\mathcal{C}_n = \mathcal{B}(v_n, x_{n+1}, \cdots)$. On taking $V_{ni} = (-a_i, a_i)$, it is seen that Theorem 3.5 applies.

(iv) Let $v_n = (\bar{x}_n, s_n)$, where $\bar{x}_n = \sum_1^n x_i/n$ and $s_n = \sum_1^n (x_i - \bar{x}_n)^2$. We consider the continuum of stopping times: $t_d$ is the first $n \geqq 1$ so that $s_n \leqq b_n\, a(d)$, or is $+\infty$ if no such $n$ occurs. We take $a(\cdot)$ increasing, $\{b_n\}$ to be an arbitrary sequence of positive constants and $\mathcal{C}_n = \mathcal{B}(v_n, x_{n+1}, \cdots)$. Clearly $c \leqq d \Rightarrow t_c \geqq t_d$ and, letting $V_{nd} = [0, b_n a(d)]$, Theorem 3.5 shows that $\{t_d : d \geqq 0\}$ is C-ordered. Let $t_{c+} = \lim_{d \downarrow c} t_d$ and $t_{c-} = \lim_{d \uparrow c} t_d$. Since these monotone limits are sequential, Theorem 2.7 shows that the augmented system

$$\{t_d, t_{d+}, t_{c-} : c > 0, d \geqq 0\}$$

is $C$-ordered.

## REFERENCES

[1] BAHADUR, R. R. (1954). Sufficiency and statistical decision functions. *Ann. Math. Statist.* **25** 423–462.

[2] BERK, R. H. (1966). Limiting behaviour of posterior distributions when the model is incorrect. *Ann. Math. Statist.* **37** 51–58.

[3] CHOW, Y. S. and ROBBINS, H. (1965). On the asymptotic theory of fixed-width sequential confidence intervals for the mean. *Ann. Math. Statist.* **36** 457–462.

[4] LOYNES, R. M. (1969). The consistency of certain sequential estimators. *Ann. Math. Statist.* To appear.