ASYMPTOTIC PROPERTIES OF THE BLOCK UP-AND-DOWN METHOD IN BIO-ASSAY[1]

By R. K. Tsutakawa

University of Chicago

1. Introduction. The block up-and-down method consists of sequentially making $K \geqq 1$ quantal response observations per trial at levels belonging to some fixed countable collection of real numbers, such that the $(\nu + 1)$st level depends only on the level and frequency of responses on the νth trial. For situations in which the sequence of levels form a Markov chain with a positive class, the asymptotic properties of the frequency of trials and responses at each level are studied and used to derive the asymptotic distribution of the maximum likelihood estimator. The method is illustrated by the random walk design where the successive levels form a random walk.

The case where $K = 1$ and the $(\nu + 1)$st level is a unit amount above or below the νth depending on whether the νth trial shows nonresponse or response, respectively, is known as the up-and-down method and, together with modifications, has been studied in [3], [7] and [8] for estimating the mean of the normal distribution. Comparisons with other procedures, such as the Robbins-Monro process have been made in [5] and [13]. With numerical results for the normal and logistic distributions, these studies have emphasized designing efficient sequential sampling procedures. The purpose here will be to study large sample properties. The problem of seeking efficient procedures has been discussed in a separate paper [12] and will not be considered here.

2. Sample properties. Let L denote a countable collection $\{d_i\}$ of distinct real numbers indexed by a subset, I, of the integers, R the integers $\{0, 1, \cdots, K\}$, and D a function defined on the Cartesian product $I \times R$ taking values in I. In bio-assay L corresponds to dosage levels and we will refer to d_i as the ith level or level i. Now consider taking independent quantal response observations, K at a time, starting at level Y_1 in I and sequentially at $Y_2, \cdots, Y_n$, defined by

$$Y_{\nu+1} = D(Y_\nu, J_\nu), \qquad \nu = 1, \cdots, n - 1,$$

where J_ν is the number of responses on the νth trial and n a predetermined total number of trials. D will be called the design function since, together with L, it specifies the sampling procedure. This method gives rise to a sequence of levels and responses $\{Y_\nu, J_\nu; 1 \leqq \nu \leqq n\}$ and, if the sampled population has distribution function F, the probability of obtaining a particular sequence

Received 1 September 1966; revised 9 May 1967.

[1] Part of this paper is based on the author's doctoral dissertation. This research was supported by USPHS Grant No. GM 201 from the Division of General Medical Sciences of the National Institutes of Health, and in part by NONR 2121(23).

$\{y_\nu, j_\nu; 1 \leqq \nu \leqq n\}$ is

$$\prod_{\nu=1}^{n} \binom{K}{j_\nu} P_{y_\nu}^{j_\nu} Q_{y_\nu}^{K-j_\nu},$$

where $P_i = 1 - Q_i = F(d_i)$, if $0 \leqq j_\nu \leqq K$ and $y_{\nu+1} = D(y_\nu, j_\nu)$ and is zero otherwise. The two processes $\{Y_\nu, J_\nu\}$ and $\{Y_\nu\}$ are Markov chains with many related properties. (It may be noted that the results given in this section and the next can be generalized, with minor changes, to include randomized rules by treating $D(Y_\nu, J_\nu)$ as a random variable over I for each value of (Y_ν, J_ν).)

In a sequence of n trials, let n_i, $i \varepsilon I$, be the number of trials at the ith level and let r_i be the corresponding number of responses. If C_F is a positive class for $\{Y_\nu\}$ and $Y_1 \varepsilon C_F$, it follows from the strong law of large numbers (cf. [4], p. 87) that n_i/n and r_i/n, $i \varepsilon C_F$, converge with probability one, as n increases, to π_i and $K\pi_i P_i$, respectively. Thus if we define p_i as r_i/n_iK or 0 according as $n_i > 0$ or $n_i = 0$, p_i converges with probability one to P_i for all $i \varepsilon C_F$. The joint distribution of $\{p_i, n_i/n\}$ is given by the following:

THEOREM 1. *If the sequence $\{Y_\nu\}$ obtained by the block up-and-down method has a positive class C_F with stationary distribution $\{\pi_i, i \varepsilon C_F\}$ and the second moment of the recurrence time, $m_{ii}^{(2)}$, is finite for some $i \varepsilon C_F$, then for any finite set $C = \{i_1, \cdots, i_s\}$ in C_F the distribution of the vector*

$$\{n^{\frac{1}{2}}(p_i - P_i), n^{\frac{1}{2}}(f_i - \pi_i); i \varepsilon C\}$$

where $f_i = n_i/n$, converges as n increases to the multivariate normal distribution with mean zero and some covariance matrix $\mathfrak{T}$.

The asymptotic distribution of the proportions of responses can be derived without requiring the finiteness of the second moment of the recurrence time and is given by the following result:

THEOREM 2. *If $\{Y_\nu, J_\nu; \nu \geqq 1\}$ is a Markov chain obtained by a block up-and-down method and the marginal process $\{Y_\nu, \nu \geqq 1\}$ has a positive class C_F with stationary distribution $\{\pi_i, i \varepsilon C_F\}$, then for any finite set $C = \{i_1, \cdots, i_s\}$ in C_F the distribution of the vector $\{(nK)^{\frac{1}{2}}(p_i - P_i), i \varepsilon C\}$ converges as n increases to the multivariate normal distribution with zero means, zero covariances, and variances P_iQ_i/π_i, $i \varepsilon C$.*

PROOF. For each i in C_F and $\nu \geqq 1$, let

$$u_{i\nu} = K^{\frac{1}{2}}(J_\nu/K - P_i)/\pi_i \quad \text{if} \quad Y_\nu = i$$
$$= 0 \quad \text{otherwise.}$$

Without loss of generality assume that $(i_1, \cdots, i_s) = (1, \cdots, s)$ and for any s real numbers $a_1, \cdots, a_s$ consider the linear combination $u_\nu = \sum_{i=1}^{s} a_i u_{i\nu}$, $\nu = 1, 2, \cdots$. Then, for each $\nu \geqq 1$, u_ν is a function of (Y_ν, J_ν) such that the conditional expectation

$$E\{u_\nu \mid Y_m, J_m; 1 \leqq m < \nu\} = 0$$

with probability one. If (Y_1, J_1) has the stationary distribution,

$$E\{u_{i\nu}u_{j\nu}\} = P_iQ_i/\pi_i \qquad \text{if} \quad i = j$$
$$= 0 \qquad \text{otherwise}$$

and

$$E\{u_\nu^2\} = \sum_{i=1}^s a_i^2 P_iQ_i/\pi_i . \tag{1}$$

It follows from a result by Billingsley [2] that the distribution of

$$n^{-\frac{1}{2}} \sum_{\nu=1}^n u_\nu = (nK)^{\frac{1}{2}} \sum_{i=1}^s a_i f_i(p_i - P_i)/\pi_i$$

converges as n increases to the normal distribution with mean zero and variance (1). Since f_i converges in probability to π_i, $i \varepsilon C_F$, the distribution of

$$(nK)^{\frac{1}{2}} \sum_{i=1}^s a_i(p_i - P_i)$$

converges to the same normal distribution. The proof is completed by applying the Cramér-Wold [6] theorem.

3. Maximum likelihood estimation. Suppose that quantal response data are obtained from a population with a distribution function, F_θ, which depends on an unknown real or vector valued parameter θ. For nonsequential designs, the asymptotically optimal estimators of θ are often discussed (cf., e.g., [9]) under the theory of regular best asymptotically normal (RBAN) estimates (cf. [11]). In RBAN estimation we consider a fixed sample design, taking $N_1, \cdots, N_s$ observations at levels $d_1, \cdots, d_s$, respectively, and analyze the asymptotic properties of the estimators as the total sample size $N = \sum_{i=1}^s N_i$ increases while N_i/N converges to some positive constant c_i, $i = 1, \cdots, s$. Under certain regularity conditions, we will show for the block up-and-down method that, if there is a finite positive class C_F with stationary distribution $\{\pi_i\}$ for $\{Y_\nu\}$, then the asymptotic distribution of the maximum likelihood estimator is, essentially, identical to that of an RBAN estimator using $c_i = \pi_i$, $i = 1, \cdots, s$.

CONDITIONS. (i) The set X of x's for which $0 < F_\theta(x) < 1$ is independent of $\theta = (\theta_1, \cdots, \theta_m) \varepsilon \Omega$, where Ω is an open set in m-dimensional Euclidean space.

(ii) For every $x \varepsilon X$, $F_\theta(x)$ has continuous partial derivatives of order 2 with respect to $\theta_1, \cdots, \theta_m$.

(iii) The process $\{Y_\nu, \nu \geqq 1\}$ has a finite positive class $C_F = \{i_1, i_2, \cdots, i_s\}$ such that $d_i \varepsilon X$ if $i \varepsilon C_F$.

(iv) The $m \times s$ matrix $\|\partial F_\theta(d_i)/\partial\theta_k\|$, where $i \varepsilon C_F$, of partial derivatives has rank m.

Note that if (iii) is replaced by the condition that $N_i/N \to c_i$, $c_i > 0$, $i = 1, \cdots, s$, then the conditions for RBAN estimation are essentially satisfied. Note also that (iv) requires that $s \geqq m$.

THEOREM 3. *Suppose that the distribution function F_θ, $\theta \varepsilon \Omega$, and the Markov chain $\{Y_\nu, J_\nu; \nu \geqq 1\}$ associated with the block up-and-down method satisfy* *conditions* (i) *through* (iv) *and that* $Y_1 \varepsilon C_F$. *Then, with probability one, if* n *is sufficiently large there exists a solution vector* $(\hat{\theta}_1, \cdots, \hat{\theta}_m)$ *to the maximum likelihood equations and the distribution of the vector* $\{(nK)^{\frac{1}{2}}(\hat{\theta}_k - \theta_k), 1 \leqq k \leqq m\}$ *converges, as* $n \to \infty$, *to the multivariate normal distribution with mean zero and covariance matrix the inverse of* $\|G_{uv}\|$ *defined by* (2) *below.*

Proof. If $\{Y_\nu, J_\nu; \nu = 1, \cdots, n\}$ is the outcome of the first n trials and $Y_1 \varepsilon C_F$, then under conditions (i) and (iii) the log-likelihood function is given by

$$L = nK \sum_{i \varepsilon C_F} f_i(p_i \log P_i + q_i \log Q_i)$$

where $P_i = 1 - Q_i = F_\theta(d_i)$ and $q_i = 1 - p_i$, $i \varepsilon C_F$. Without loss of generality, assume that $C_F = \{1, \cdots, s\}$. Under conditions (i), (ii), and (iii) the partial derivatives, L_k, of L with respect to θ_k, $k = 1, \cdots, m$, exist. Moreover, under the same conditions $L_k = L_k(p, f, \theta)$ possesses continuous partial derivatives with respect to the arguments $p = (p_1, \cdots, p_s)$, $f = (f_1, \cdots, f_s)$ and $\theta = (\theta_1, \cdots, \theta_m)$. For a fixed point $\theta^0 = (\theta_1^0, \cdots, \theta_m^0)$ in Ω, let P^0 be the value of $P = (P_1, \cdots, P_s)$ and π^0 the value of the stationary distribution $\{\pi_1^0, \cdots, \pi_s^0\}$ for C_F when $\theta = \theta^0$. Then, $L_k(P^0, \pi^0, \theta^0) = 0$, $k = 1, \cdots, m$. Define

$$G_{uv} = \sum_{i=1}^{s} \pi_i^0 P_{iu}^0 P_{iv}^0 / P_i^0 Q_i^0, \qquad u, v = 1, \cdots, m, \tag{2}$$

where P_{iu}^0 is the partial derivative of P_i with respect to θ_u evaluated at $\theta = \theta^0$, $i = 1, \cdots, s$; $u = 1, \cdots, m$. These are the second partial derivatives of L with respect to $\theta_1, \cdots, \theta_m$ evaluated at (P^0, π^0, θ^0) and divided by $-nK$. By (iv) the matrix $\|G_{uv}\|$ is nonsingular. Thus, by the implicit function theorem there exists m unique functions $\hat{\theta}_k(p, f)$, $k = 1, \cdots, m$, which are continuously differentiable and satisfy the maximum likelihood equations,

$$L_k(p, f, \theta) = 0, \qquad k = 1, \cdots, m, \tag{3}$$

for (p, f) in a neighborhood of the point (P^0, π^0) such that $\theta_k^0 = \hat{\theta}_k(P^0, \pi^0)$, $k = 1, \cdots, m$. Furthermore, the partial derivatives a_{ki} and b_{ki} of $\hat{\theta}_k$ with respect to p_i and f_i, respectively, evaluated at (P^0, π^0) are given by the equations

$$\sum_{u=1}^{m} G_{ku} a_{ui} = (P_{ik}^i \pi_i^0)/(P_i^0 Q_i^0), \tag{4}$$

$$\sum_{u=1}^{m} G_{ku} b_{ui} = 0, \tag{5}$$

$k = 1, \cdots, m$; $i = 1, \cdots, s$.

Since the matrix $\|G_{uv}\|$ is nonsingular, (5) implies $b_{ki} = 0$, $k = 1, \cdots, m$; $i = 1, \cdots, s$. As indicated in the last section (p, f) converges, as $n \to \infty$, to (P^0, π^0) with probability one. Thus, with probability one, if n is sufficiently large there exists a unique solution $(\hat{\theta}_1, \cdots, \hat{\theta}_m)$ to the maximum likelihood equations (3). Moreover, using Theorems 1 and 2 together with (4) and (5), it follows that the distribution of the vector

$$\{(nK)^{\frac{1}{2}}(\hat{\theta}_k - \theta_k^0), 1 \leqq k \leqq m\}$$

and the vector

$$\{(nK)^{\frac{1}{2}} \sum_{i=1}^{s} a_{ki}(p_i - P_i^0),\ 1 \leqq k \leqq m\},$$

converge to the same distribution, which is the multivariate normal distribution with mean zero and covariance matrix the inverse of $\|G_{uv}\|$, completing the proof of the theorem.

An extension of the above result to cases where C_F is countably infinite can be made by applying a result given by Billingsley [1]. The extension requires a few additional conditions involving the third order partial derivatives of F_θ with respect to $\theta_1, \cdots, \theta_m$. It may be noted that the asymptotic covariance matrix of the maximum likelihood estimator, for both the finite and infinite cases, depends on the Markov property of the design only through the stationary distribution.

4. Random walk design. A class of block up-and-down methods which can be analyzed in some detail is the random walk designs. For this design L is a sequence, $\{\cdots, d_{-1}, d_0, d_1, \cdots\}$, of increasing real numbers having no accumulation points. The design function is now given by $D(i, r) = i + 1, i$, or, $i - 1$ according as $0 \leqq r \leqq k_1$, $k_1 < r < k_2$, or $k_2 \leqq r \leqq K$, where k_1 and k_2, $0 \leqq k_1 < k_2 \leqq K$, are fixed constants; thus $\{Y_\nu\}$ forms a random walk. We shall now consider the finiteness of various moments which appear in the asymptotic study of this design.

Let α_i and β_i denote the transition probabilities of moving up or down, respectively, from $i \varepsilon I$. (These are just the tail probabilities of the binomial distribution with parameter P_i.) (Random walks with $\alpha_i + \beta_i = 1$, $i \varepsilon I$, are discussed in [4] and [10].) Since F is a distribution function and L has no accumulation points, $\{\alpha_i, \beta_i; i \varepsilon I\}$ satisfies the following properties:

$$\alpha_i + \beta_i > 0, \tag{6}$$

$$\alpha_{i+1} \leqq \alpha_i \quad \text{and} \quad \beta_i \leqq \beta_{i+1}, \tag{7}$$

$$\beta_{i_1} < \alpha_{i_1} \quad \text{and} \quad \alpha_{i_2} < \beta_{i_2}, \tag{8}$$

for some integers i_1 and i_2, $i_1 < i_2$. Now let

$$C_F = \{i, F(d_{i-1}) < 1 \text{ and } F(d_{i+1}) > 0\}.$$

Without loss of generality assume $0 \varepsilon C_F$ and for each integer i let

$$\begin{aligned} L_i &= \textstyle\prod_{k=1}^{i} \alpha_{k-1}/\beta_k && \text{if } i > 0 \\ &= 1 && \text{if } i = 0 \\ &= \textstyle\prod_{k=i}^{-1} \beta_{k+1}/\alpha_k && \text{if } i < 0. \end{aligned} \tag{9}$$

From (6), (7), and (8) it follows that

$$\textstyle\sum_{i \varepsilon C_F} |i|^r L_i < \infty \tag{10}$$

for each positive integer r. It is easily verified that C_F forms a positive class with stationary distribution $\pi_i = L_i/\sum L_j$, $i \varepsilon C_F$.

Given $i, j \varepsilon I$, $i \neq j$, let ρ_{ij} be the probability that the random walk reaches j before it returns to i, given that it starts at i. Define ρ_{ii} to be 1 or 0 depending on whether i belongs to C_F or not. We will now show that for any fixed $j_0 \varepsilon C_F$,

$$\inf_{i\varepsilon C_F} \rho_{ij_0} > 0. \tag{11}$$

A difference equation technique (cf. [4], p. 66) can be used to show that, for $j < i$, with $i, j \varepsilon C_F$,

$$\rho_{ij} = \beta_i\{1 + \textstyle\sum_{k=j+1}^{i-1} \prod_{u=k}^{i-1} \alpha_u/\beta_u\}^{-1}, \tag{12}$$

where we define the empty sum $\sum_i^{i-1} = 0$. (Cases where both i and j do not belong to C_F may be treated in the obvious manner. By symmetry, the expression for the case $i < j$ is similar.) The inequality (11) follows upon applying (6), (7), and (8) to (12).

For each $i, j \varepsilon C_F$, let m_{ij} and $m_{ij}^{(2)}$ be the first and second moments of the first passage time from i to j. (Here, i and j need not be distinct.) We will now show that for each $i \varepsilon C_F$, $m_{ii}^{(2)} < \infty$. It will then follow that $m_{ij}^{(2)} < \infty$, for every $i, j \varepsilon C_F$.

Suppose $Y_1 = i$ for some $i \varepsilon C_F$. Let t be the first recurrence time of i. Let h be a real valued function on I and let $U = \sum_{\nu=1}^{t-1} h(Y_\nu)$. The expression for the second moment of U given in [4], p. 83, can be simplified for random walks by noting that

$$m_{ik} = m_{ij} + m_{jk} \tag{13}$$

if $i < j < k$ or $i > j > k$ and using the result, shown by Harris [10], that for distinct j and k in C_F

$$\rho_{jk}(m_{jk} + m_{kj}) = m_{jj}. \tag{14}$$

Substitution of (13) and (14) into the expression in [4] leads to

$$E\{U^2\} = m_{ii}\{4\textstyle\sum\sum_{(j,k)\varepsilon C_i} h(j)h(k)\pi_k/\rho_{ji} + \sum_{j\varepsilon C_F} (2 - \rho_{ji})h^2(j)\pi_j/\rho_{ji} + 2h(i)\sum_{j\neq i, j\varepsilon C_F} h(j)\pi_j\}, \tag{15}$$

provided the series converge absolutely, where $C_i = \{(j, k) \varepsilon C_F \times C_F; i < j < k$ or $i > j > k\}$. The finiteness of $m_{ii}^{(2)}$ follows upon setting $h \equiv 1$ in (15) and applying (10) and (11).

Acknowledgment. I appreciate the helpful comments from the referee and Professors P. Billingsley, W. H. Kruskal, and D. L. Wallace. I am grateful for the direction I received from Professor P. Meier.

REFERENCES

[1] BILLINGSLEY, P. (1961). *Statistical Inference for Markov Processes.* Institute of Mathematical Statistics, Univ. of Chicago Statistical Research Monographs, Univ. of Chicago Press.

[2] BILLINGSLEY, P. (1961). The Lindeberg-Lévy theorem for martingales. *Proc. Amer. Math. Soc.* **12** 788–792.

[3] BROWNLEE, K. A., HODGES, J. L., JR., and ROSENBLATT, M. (1953). The up-and-down method with small samples. *J. Amer. Statist. Assoc.* **48** 262–277.

[4] CHUNG, K. L. (1960). *Markov Chains with Stationary Transition Probabilities.* Springer, Berlin.

[5] COCHRAN, W. G. and DAVIS, M. (1964). Stochastic approximation to the median effective dose in bioassay, *Stochastic Models in Medicine and Biology*, edited by John Gurland. Univ. of Wisconsin Press, Madison. 281–300.

[6] CRAMÉR, H and WOLD, H. (1936). Some theorems on distribution functions. *J. London Math. Soc.* **11** 290–294.

[7] DIXON, W. J. (1965). The up-and-down method for small samples. *J. Amer. Statist. Assoc.* **60** 967–978.

[8] DIXON, W. J. and MOOD, A. M. (1948). A method for obtaining and analyzing sensitivity data. *J. Amer. Statist. Assoc.* **43** 109–126.

[9] FINNEY, D. J. (1964). *Statistical Method in Biological Assay* (2nd Ed.). Hafner, New York.

[10] HARRIS, T. E. (1952). First passage and recurrence distributions. *Trans. Amer. Math. Soc.* **73** 471–486.

[11] NEYMAN, J. (1949). Contribution to the theory of the χ^2 test. *Proc. Berkeley Symp. Math. Statist. Prob.* 239–273. Univ. of California Press.

[12] TSUTAKAWA, R. K. (1967). Random walk design in bio-assay. *J. Amer. Statist. Assoc.* **62** 842–856.

[13] WETHERILL, G. B. (1963). Sequential estimation of quantal response curves (with discussion). *J. Roy. Statist. Soc. Ser. B* **25** 1–48.