A METHOD OF SEQUENTIAL ESTIMATION APPLICABLE TO THE HYPERGEOMETRIC, BINOMIAL, POISSON, AND EXPONENTIAL DISTRIBUTIONS

By William Knight

University of New Brunswick

1. Background and summary. Sequential estimation for the particular stopping rule, inverse or Haldane sampling, has been developed for the hypergeometric [4], [13], [15], binomial [10], [12], Poisson [17], and exponential [6], [16], distributions; more general stopping rules have been discussed for the binomial distribution by Girshick, Mosteller, and Savage [11] and for the exponential distribution by the author [14]. It is the purpose of this article to present a unified treatment of these distributions and to extend the known results for inverse sampling to a more general class of stopping rules called monotone. Estimates of parameters are obtained following [11], estimates of the variances of these estimates obtained, their distribution related to that of the fixed sample size stopping rule, and confidence methods suggested. Most of the methods described do not require complete knowledge of the stopping rule for their application.

2. Introduction. A system of points in a Cartesian plane called the *branch points* is given. These will be either the set of points with integral coordinates (discrete case) or the set of points with integral x coordinate (semi-continuous case). Branch points are denoted by lower case boldface letters and are added as vectors. The point, (0, 0), is denoted by $\mathbf{0}$, (1, 0) by $\mathbf{1}$, $(-1, 0)$ by $-\mathbf{1}$, etc. The coordinates of $\mathbf{a}$ are denoted $x(\mathbf{a})$ and $y(\mathbf{a})$. The notation, $\mathbf{a} \leqq \mathbf{b}$, means that both $x(\mathbf{a}) \leqq x(\mathbf{b})$ and $y(\mathbf{a}) \leqq y(\mathbf{b})$.

Imagine a path traced by a particle moving in the plane, motion at any instant being parallel and positively directed with respect to one of the axes. Direction of motion may change only at branch points, thus the path is a sequence of steps. A path is conveniently identified with the ordered set of its branch points. Moreover, a path is determined by its starting point and the numbers, $y_1, y_2, y_3, \cdots$, the y coordinates of the branch points from which a unit jump in the x direction occurs. These y's are taken to be a family of random variables, i.e. a stochastic process; in this sense the path also is random.

Define $P(\mathbf{a}, \mathbf{b})$ as the conditional probability that a path intersects $\mathbf{b}$ given that $\mathbf{a} \leqq \mathbf{b}$ is intersected, and $p(\mathbf{a}, \mathbf{b}, \mathbf{c})$ as the conditional probability that $\mathbf{b}$ is intersected, $\mathbf{a} \leqq \mathbf{b} \leqq \mathbf{c}$, given that $\mathbf{a}$ and $\mathbf{c}$ are intersected.

Three special cases are of interest.

The binomial case: A random experiment with two possible outcomes, call them positive and negative, is independently repeated every time the path reaches a branch point. The path moves one unit in the x direction if the out-

Received 9 July 1962; revised 18 January 1965.

come is positive, one unit in the y direction if negative. $P(\mathbf{a}, \mathbf{a} + \mathbf{1})$ is the same for all branch points. An estimate of the probability of a positive outcome, $P(\mathbf{a}, \mathbf{a} + \mathbf{1})$, is sought.

The hypergeometric case will be described by means of an example. Balls are drawn randomly from an urn without replacement. With each drawing the path moves one unit in the x direction if the ball be white, one unit in the y direction if black. Usually an estimate of the ratio of black to white balls is sought; sometimes an estimate of the number of balls in the urn is wanted.

The Poisson case is semi-continuous. The particle moves along lines parallel to the y axis with uniform velocity, jumping on occasion one unit in the x direction, the jumps being events in a Poisson process. Usually an estimate of the probability density of a jump in the x direction,

$$\lim_{h \to 0} P(\mathbf{a}, \mathbf{a} + \mathbf{d})/h \qquad \text{where} \quad \mathbf{d} = (1, h), \tag{2.1}$$

is sought; its reciprocal which is the expected y distance between jumps is also often of interest.

From these applications two assumptions are abstracted from which our estimation scheme will be built.

ASSUMPTION I. The process, $y_1, y_2, y_3, \cdots$, is Markov. It follows that

$$p(\mathbf{a}, \mathbf{b}, \mathbf{c}) = P(\mathbf{a}, \mathbf{b})P(\mathbf{b}, \mathbf{c})/P(\mathbf{a}, \mathbf{c}). \tag{2.2}$$

ASSUMPTION II. The conditional distribution of paths connecting $\mathbf{a}$ and $\mathbf{b} \geqq \mathbf{a}$ (given $\mathbf{a}$ and $\mathbf{b}$ are intersected) is known where defined and independent of the parameter(s) to be estimated. In hypergeometric and binomial cases this follows from an *equipartition property*, all paths connecting two points are of equal probability. In the Poisson case, the conditional distribution of $y_1, y_2, \cdots, y_k$, where $k = x(\mathbf{b} - \mathbf{a})$, is that of an ordered sample of size k from the rectangular distribution on $[y(\mathbf{a}), y(\mathbf{b})]$. This seems generally known; a proof is in [7].

In particular, Assumption II implies that $p(\mathbf{a}, \mathbf{b}, \mathbf{c})$ is known.

3. Continuity problems. Some use of probability densities seems unavoidable. As the same results usually apply to probabilities and probability densities the same symbols and expressions usually refer to both, indeed both probabilities and densities sometimes appear in the same expression. The context is depended upon to make distinctions when necessary.

The density, $P(\mathbf{a}, \mathbf{a} + \mathbf{1})$, is defined by expression (2.1); $p(\mathbf{a}, \mathbf{a} + \mathbf{1}, \mathbf{b})$ can be defined either by a similar limit or by (2.2). The density, $P(\mathbf{a}, \mathbf{a} + \mathbf{2})$, can be defined as a limit of second order terms or simply as $P(\mathbf{a}, \mathbf{a} + \mathbf{1})P(\mathbf{a} + \mathbf{1}, \mathbf{a} + \mathbf{2})$. Such devices suffice for this discussion; although there is more unity in considering all such terms as probability densities with respect to a suitable measure. This is discussed in the following paragraph but may be omitted.

For this purpose the natural measure on paths starting at $\mathbf{a}$ is induced by a measure on the sequences, $\{y_i\}$. Take on y_i for each i the uniform measure on $[y(\mathbf{a}), \infty)$, discrete or continuous as appropriate. On the sequence, $\{y_i\}$, take the product measure restricted to the simplex, $y_1 \leqq y_2 \leqq y_3 \leqq \cdots$. Note that in discrete cases the measure of a set of paths is the number of paths in the set.

4. Stopping rules. In sequential estimation the path is started at some point, **0** unless otherwise stated, and stopped according to a stopping rule. This stopping rule can be defined by a region, **R**, observation ceasing at the first branch point not in **R** encountered which is called the *terminating point* and denoted **z**. To avoid trivial complications it is demanded that **0** be in **R** and that **R** be open.

A path or part thereof is called *proper* if all branch points save perhaps the last are in **R** and *very proper* if all branch points are in **R**. A branch point is called a *continuation point* if connected to the origin by a very proper path or if it has one or more negative coordinates, a *boundary point* if connected to the origin by a proper path and not a continuation point, and an *inaccessible point* otherwise. A point is *fully accessible* if a path connecting it to the origin is proper with probability one or it has one or more negative coordinates. A boundary is *closed* with respect to **a** if **a** is a continuation point and a path starting at **a** intersects the boundary with probability one.

5. Estimation. A model is completely specified by its branch points and transition probabilities, the former being known and the latter to be estimated. Usually estimates of the parameters, $F = P(\mathbf{0}, \mathbf{1})$ and $T = 1/P(-\mathbf{1}, \mathbf{0})$, are of chief interest. For example, in binomial models, all $P(\mathbf{a}, \mathbf{b})$ are determined by F; in Poisson models, F is the density of events and T the expected time between successive events. While F and T are simply related in the cases of Section 2, and estimates of one yield estimates of the other, unbiased estimates of one yield biased estimates of the other, so both need be considered.

It follows immediately from Assumption II that the terminating point is a sufficient statistic for all parameters.

Estimates will be obtained which are to some extent independent of the boundary used.

THEOREM 1. *The maximum likelihood estimate of any parameter is a function of the terminating point alone being the same for all boundaries.*

PROOF. The probability (density) of termination at a boundary point, **z**, is $P(\mathbf{0}, \mathbf{z})$ times the conditional probability (density) that a path starting at **0** is proper given that it intersects **z**. The first factor is independent of the boundary, the latter, by Assumption II, independent of the parameters.

Unbiased estimates of certain parameters, obtained by Girshick, Mosteller, and Savage [11] for the binomial model, will now be extended to the general case.

Define $G(\mathbf{a}, \mathbf{b})$ as the probability (density) that a path starting at **a** intersect **b** at or before its terminating point. Define $g(\mathbf{a}, \mathbf{b}, \mathbf{c})$ as the conditional probability (density) that a path starting at **a** intersect **b** at or before intersecting **c** given that it intersects **c** at or before its terminating point. It is remarked that

$$(5.1) \qquad g(\mathbf{a}, \mathbf{b}, \mathbf{c}) = G(\mathbf{a}, \mathbf{b})G(\mathbf{b}, \mathbf{c})/G(\mathbf{a}, \mathbf{c}) \quad \text{if} \quad \mathbf{a} \leqq \mathbf{b} \leqq \mathbf{c}$$
$$= 0 \qquad \text{otherwise,}$$

and that by Assumption II g is independent of the parameters of the process. Define $\mathbf{m} = \mathbf{m}(\mathbf{a})$ as $(\text{Min}\,[0, x(\mathbf{a})], \text{Min}\,[0, y(\mathbf{a})])$.

For any statistic, $s = s(\mathbf{z})$, the relation,

$$E\{s(\mathbf{z}) | \text{Event}\} \text{ Prob } \{\text{Event}\} = E\{s(\mathbf{z}) \text{ Prob } \{\text{Event} \mid \mathbf{z}\}\}, \tag{5.2}$$

can be used to obtain

$$\begin{aligned} E\{s \cdot g(\mathbf{m}, \mathbf{a}, \mathbf{z}) | \mathbf{0}\} G(\mathbf{m}, \mathbf{0}) &= E\{s \cdot g(\mathbf{m}, \mathbf{a}, \mathbf{z}) \cdot g(\mathbf{m}, \mathbf{0}, \mathbf{z})\} \\ &= E\{s \cdot g(\mathbf{m}, \mathbf{0}, \mathbf{z}) | \mathbf{a}\} G(\mathbf{m}, \mathbf{a}) \end{aligned} \tag{5.3}$$

where expectation is over paths starting at $\mathbf{m}$, $\mathbf{z}$ is the terminating point of the path, and the notation, $\mid \mathbf{b}$, denotes conditional expectation given that the path ntersects $\mathbf{b}$ at or before its terminating point. Note that if $\mathbf{z}$ is an arbitrary boundary point, a path intersects $\mathbf{z}$ at or before its terminating point if and only if $\mathbf{z}$ is the terminating point. Of course if the boundary is not closed with respect to $\mathbf{m}$ the expectations of (5.3) are undefined.

If s is taken as $1/g(\mathbf{m}, \mathbf{0}, \mathbf{z})$, Equation (5.3) can be rewritten as

$$G(\mathbf{m}, \mathbf{a})/G(\mathbf{m}, \mathbf{0}) = E\{[g(\mathbf{m}, \mathbf{a}, \mathbf{z})/g(\mathbf{m}, \mathbf{0}, \mathbf{z})] \mid \mathbf{0}\}. \tag{5.4}$$

As the conditional distribution of the terminating point given that the path intersects $\mathbf{0}$ is the same as the distribution of the terminating point of a path starting at $\mathbf{0}$, this yields an unbiased estimate of any parameter of the form of the left side of (5.4) provided the expectation exists. A special case is

THEOREM 2. *Assume the boundary closed with respect to* $\mathbf{0}$.

(i) *If* $\mathbf{a} \geqq \mathbf{0}$ *is fully accessible, an unbiased estimate of* $P(\mathbf{0}, \mathbf{a})$ *is* $g(\mathbf{0}, \mathbf{a}, \mathbf{z})$.

(ii) *If* $\mathbf{a} \leqq \mathbf{0}$ *and all points on the* y *axis are not inaccessible, an unbiased estimate of* $1/P(\mathbf{a}, \mathbf{0})$ *where defined is* $1/g(\mathbf{a}, \mathbf{0}, \mathbf{z})$.

PROOF. The (i) denominators (ii) numerators of (5.4) become identically one. The conditions assure closure with respect to $\mathbf{m}(\mathbf{a})$, and that (i) $G(\mathbf{0}, \mathbf{a}) = P(\mathbf{0}, \mathbf{a})$, (ii) $G(\mathbf{a}, \mathbf{0}) = P(\mathbf{a}, \mathbf{0})$.

In particular, an unbiased estimate of F is $f = g(\mathbf{0}, \mathbf{1}, \mathbf{z})$, and an unbiased estimate of T is $t = 1/g(-\mathbf{1}, \mathbf{0}, \mathbf{z})$.

In some models, in particular those of Section 2, unbiased estimates of the variances of f and t are obtained from

COROLLARY 2A. *If the conditions of Theorem* 2 *hold for* (i) $\mathbf{a} = \mathbf{2}$ (ii) $\mathbf{a} = -\mathbf{2}$, *and if there exist known constants* A *and* B *such that*

$$\text{(i) } P(\mathbf{0}, \mathbf{1}) = A + BP(\mathbf{1}, \mathbf{2}) \qquad \text{(ii) } 1/P(-\mathbf{1}, \mathbf{0}) = A + B/P(-\mathbf{2}, -\mathbf{1}) \tag{5.5}$$

then an unbiased estimate of the variance of (i) f (ii) t *is*

$$\text{(i) } f^2 - Af - Bg(\mathbf{0}, \mathbf{2}, \mathbf{z}) \qquad \text{(ii) } t^2 - At - B/g(-\mathbf{2}, \mathbf{0}, \mathbf{z}). \tag{5.6}$$

Proof of (i) is given, that of (ii) being similar.

$$\begin{aligned} V(f) &= E(f^2) - [P(\mathbf{0}, \mathbf{1})]^2 \\ &= E(f^2) - P(\mathbf{0}, \mathbf{1})[A + BP(\mathbf{1}, \mathbf{2})] \\ &= E(f^2) - AP(\mathbf{0}, \mathbf{1}) - BP(\mathbf{0}, \mathbf{2}) \\ &= E(f^2 - Af - Bg(\mathbf{0}, \mathbf{2}, \mathbf{z})). \end{aligned} \tag{5.7}$$

6. Monotone boundaries. A stopping rule or the associated boundary is called *monotone* if all continuation points are fully accessible. There being no proper path from any boundary point, **b**, to a continuation point, **c**, the relation, $\mathbf{b} \leqq \mathbf{c}$, is impossible; thus the x coordinate of the boundary is nonincreasing in the y coordinate. Boundaries in this section are assumed monotone.

The boundary of the Wald sequential test is not monotone; however, monotonicity seems a reasonable property for the boundary of a sequential estimation plan as the sample size is independent of the order of observations except perhaps the last. If the equipartition property holds any order of observations is equally probable, and like treatment of all seems reasonable.

Boundary points may be divided into three *types* which it is convenient to name by the points, $\mathbf{e} = \mathbf{0}, \mathbf{1}, (0, 1)$. A boundary point is of type **0** if fully accessible; otherwise it is of type **1** or (0, 1) if (with probability one) it is approachable from the x or y direction respectively. For a boundary point, **b**, denote $\mathbf{b} - \mathbf{e}$ by $\mathbf{b}'$. Almost all paths terminating at **b** pass through $\mathbf{b}'$. An immediate consequence is

THEOREM 3. *If* **c** *is a point on a monotone boundary* $G(\mathbf{a}, \mathbf{c}) = P(\mathbf{a}, \mathbf{c}')$ *and* $g(\mathbf{a}, \mathbf{b}, \mathbf{c}) = p(\mathbf{a}, \mathbf{b}, \mathbf{c}')$.

This simplifies the estimates of Theorem 2 and its corollary, for while the function, g, depends on the boundary, the function, p, does not. Whereas the maximum likelihood estimates depend on the coordinates of the terminating point alone, these unbiased estimates depend only on its coordinates and type.

The distribution function of the terminating point likewise depends only on its coordinates and type. Define $\theta(\mathbf{a}) = \arctan [x(\mathbf{a})/y(\mathbf{a})]$.

THEOREM 4.

$$\text{Prob}\,\{\theta(\mathbf{z}) \leqq \theta(\mathbf{b})\} = \Phi_x(x(\mathbf{b}'), y(\mathbf{b})), \tag{6.1}$$

$$\text{Prob}\,\{\theta(\mathbf{z}) \geqq \theta(\mathbf{b})\} = \Phi_y(x(\mathbf{b}), y(\mathbf{b}')), \tag{6.2}$$

takes the same for all monotone boundaries containing **b**, *and this defines* Φ_x *and* Φ_y.

Proof of (6.1) is given, (6.2) being the dual. On a monotone boundary $\theta(\mathbf{z}) \leqq \theta(\mathbf{b})$ if and only if the path reaches the closed horizontal line joining the y axis to **b** or $(x(\mathbf{b}'), y(\mathbf{b}))$, intersection of the two lines being equivalent with probability one. All paths to the latter line being proper, the probability of its intersection is independent of the boundary. My thanks to the referee for simplification of the above proof.

As a consequence of Theorem 4 the distribution of the termination point and associated likelihood functions and confidence regions are functions solely of the coordinates and type of the terminating point (and the parameters of the process) being independent of the boundary. This simplifies tabulation as a table appropriate to one class of boundaries, e.g. fixed sample size, is applicable to all.

7. Hypergeometric model. This and the following three sections are devoted to particular models. The reader interested in the semi-continuous Poisson model rather than the discrete binomial and hypergeometric models may skip to Section 10 without loss of continuity; conversely the reader interested in discrete models only may omit Section 10.

Hypergeometric models are parametrized by a point, $\mathbf{k} \geqq \mathbf{0}$, with integral coordinates. Points $\mathbf{b} \leqq \mathbf{k}$ with integral coordinates are branch points. Transition probabilities are $P(\mathbf{b}, \mathbf{b}+\mathbf{1}) = x(\mathbf{k}-\mathbf{b})/n(\mathbf{k}-\mathbf{b})$ where $n(\mathbf{a}) = x(\mathbf{a}) + y(\mathbf{a})$. By the equipartition property

$$(7.1) \qquad p(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \binom{n(\mathbf{b}-\mathbf{a})}{x(\mathbf{b}-\mathbf{a})}\binom{n(\mathbf{c}-\mathbf{b})}{x(\mathbf{c}-\mathbf{b})}\Big/\binom{n(\mathbf{c}-\mathbf{a})}{x(\mathbf{c}-\mathbf{a})}.$$

Expressions for this and other models are simplified by the notations, $x = x(\mathbf{z})$, $y = y(\mathbf{z})$, $n = n(\mathbf{z})$, $x' = x(\mathbf{z}')$, $y' = y(\mathbf{z}')$, $n' = n(\mathbf{z}')$, $X = x(\mathbf{k})$, $Y = y(\mathbf{k})$, $N = n(\mathbf{k})$.

Maximum likelihood estimates are the same as for the fixed sample size boundary. If (i) N (ii) X is known, the maximum likelihood estimate of (i) X (ii) N is the largest integer no greater than (i) xN/n (ii) nX/x [9].

The remainder of this section applies only to monotone boundaries. An unbiased estimate of $F = X/N$ is $f = p(\mathbf{0}, \mathbf{1}, \mathbf{z}') = x'/n'$. Using the relation

$$(7.2) \qquad P(\mathbf{0}, \mathbf{1}) = [1/N] + [(N-1)/N]P(\mathbf{1}, \mathbf{2}),$$

Corollary 2A yiělds, after some computation, an unbiased estimate of the variance of f, $f(1-f)(N-n')/(n'-1)N$. Similarly an unbiased estimate of $T = (N+1)/(X+1)$ is $t = (n'+1)/(x'+1)$, an unbiased estimate of the variance of which is $t(t-1)(X-x')/(x'+2)(X+1)$, this being obtained from Corollary 2A and

$$(7.3) \qquad P(-\mathbf{1}, \mathbf{0}) = [-1/(X+1)] + [(X+2)/(X+1)]P(-\mathbf{2}, -\mathbf{1}).$$

The one sided confidence interval bounding F below is that which would apply if a sample of fixed size n' had been taken and x successes obtained; the interval bounding F above is that which would apply if a sample of fixed size n' had been taken and x' successes obtained. Two sided intervals can be constructed from two one sided intervals.

8. Binomial model. The binomial model is similar to the hypergeometric and will merely be outlined. The function, $p(\mathbf{a}, \mathbf{b}, \mathbf{c})$, is the same as in the hypergeometric model. The maximum likelihood estimate of F is x/n.

For monotone boundaries an unbiased estimate of F is $f = x'/n'$, and an unbiased estimate of the variance of f is $f(1-f)/(n'-1)$. The same statements apply to confidence intervals as in the hypergeometric case.

9. Special boundaries. Some monotone stopping rules for hypergeometric and binomial models are discussed in this section. Formulae are given for hypergeometric models, binomial forms being limiting cases.

The fixed sample size stopping rule, stop when n reaches some predetermined value, n^*, and the inverse sampling rule, stop when x reaches some predetermined value, x^*, are well known, references for the latter being [9], [10], [12], (binomial) and [4], [13], [15], (hypergeometric). Using a fixed sample size bounds the variance of f; inverse sampling bounds the coefficient of variation of t, with

$$V(t)/E(t)^2 \leqq [(x^*)^{-1} - (X+1)^{-1}][(X+1)/(X+2)]. \tag{9.1}$$

Of use might be truncated inverse sampling: Stop when x reaches a predetermined value, x^*, or n reaches a predetermined value, n^*, whichever happens first. (In other words, use inverse sampling until the money is exhausted!)

The rule, stop when

$$n \geqq n^*, \qquad x \geqq x^* + 1, \qquad \text{or} \quad y \geqq y^* + 1 \tag{9.2}$$

whichever happens first, where

$$\begin{aligned} n^* &= (\tfrac{1}{4}\sigma^{-2} + 1)/(1 + \tfrac{1}{4}\sigma^{-2}N^{-1}), \\ x^* = y^* &= [(4/27)\sigma^{-2} + 1]/[1 + (4/27)\sigma^{-2}N^{-1}], \end{aligned} \tag{9.3}$$

assures that $V(f) \leqq \sigma^2$ where σ^2 is set in advance. Comparing this with the rule of fixed sample size for the same bound on $V(f)$, stop when

$$n \geqq \tfrac{1}{4}\sigma^{-2}/[1 + \tfrac{1}{4}\sigma^{-2}N^{-1} - N^{-1}], \tag{9.4}$$

it can be seen that the sample size obtained using Rule (9.4) can be at most one greater than that obtained using (9.2) and may be smaller by as much as a factor of 16/27.

It will be shown that the estimated variance of f,

$$s^2 = f(1-f)(N-n')/(n'-1)N = x'y'(N-n')/(n')^2(n'-1)N, \tag{9.5}$$

is bounded by σ^2 at all points on the boundary defined by (9.2). The estimate being unbiased, the same bound holds for $V(f)$. If $n' = n^*$, $s^2 = f(1-f)4\sigma^2 \leqq \sigma^2$. If $x' = x^*$, the inequality, $x'(n'-1)/n'(x'-1) \geqq 1$ if $x' \leqq n'$, can be used to obtain

$$s^2 \leqq (x')^2y'(N-x')/(n')^3(x'-1)N = f^2(1-f)(27/4)\sigma^2 \leqq \sigma^2. \tag{9.6}$$

The boundary points determined by the stopping rule, stop when $x \geqq x^* + 1$, are of type **1**. Thus $x' = x - 1 \geqq x^*$.

10. Poisson model. Poisson models differ from discrete models in that there are no boundary points of type (0, 1), for if approach to a point from the x direction is cut off, the probability of the paths so made improper is zero, and the point is still fully accessible. It is convenient to denote $x = x(\mathbf{z})$, $x' = x(\mathbf{z}')$, $y = y(\mathbf{z}) = y(\mathbf{z}')$.

Under the well known stopping rule, stop when $x = x^*$, where x^* is a predetermined value, y has a gamma distribution, the sum of x^* exponentially distributed random variables with mean, T. The maximum likelihood estimate of T is y/x for this boundary and must be the same for all others. Under this boundary the maximum likelihood estimate is unbiased. Written as a function of $\mathbf{z}'$ it is $t = y/(x' + 1)$. By the completeness of the gamma distribution, t is the only unbiased estimate, must be the one determined by Theorem 2, and thus is valid for all monotone boundaries with no points on the y axis. For the fixed x boundary an unbiased estimate of the variance of t is $t^2/(x' + 2)$. This again is the only unbiased estimate, must be that determined by Corollary 2A, and thus is valid for all monotone boundaries with no points on the y axis. (The relation, $1/P(-\mathbf{1}, \mathbf{0}) = 1/P(-\mathbf{2}, -\mathbf{1})$, satisfies (5.5).)

Under the stopping rule, stop when $y = y^*$, x has a Poisson distribution with expectation, y^*F. An unbiased estimate of F is $f = x'/y$, and an unbiased estimate of the variance of f is f/y. By the completeness of the Poisson distribution these estimates are unique, must be the ones determined by Theorem 2 and Corollary 2A, and thus are valid for all monotone boundaries for which $\mathbf{2}$ is a continuation point.

One sided confidence intervals bounding F below are calculated from the point, $\mathbf{z}$, those bounding F above from $\mathbf{z}'$. Tables for either the gamma or Poisson distribution can be used. Two sided intervals are obtained from two one sided intervals.

The remainder of this section is devoted to illustrations of applications of the Poisson model.

In making bacterial counts, the number of colonies on a plate of standard area is counted, this count being Poisson distributed. Sandelius [17] suggests a modification whereby part of the plate is viewed, the area of the viewed part being increased until a predetermined number of colonies are contained therein; if one plate is insufficient the process is continued with additional plates. The two methods correspond to stopping at a predetermined y and x respectively in our formulation. The second procedure could be truncated at one, or some other number of plates.

Life testing applications have received much attention, e.g. [6], [8]. A population of articles with exponentially distributed failure time is given; the object is the estimation of the mean life, T. Articles are placed on test, being removed from test by failure. Additional articles may be added according to some predetermined plan which may include replacement upon failure. Call the number of articles on test at time τ, $r(\tau)$. The path used for sequential estimation is that described by an imaginary particle moving in the y direction with velocity, $r(\tau)$, which jumps one unit in the x direction with each failure. The jumps will then be events in a Poisson process with transition probability density $F = 1/T$. It is well known [1] that the process is independent of the replacement policy, i.e. $r(\tau)$, thus estimates will be identical for any replacement policy. The x coordinate is commonly called the *number of failures*, the y coordinate the *observed life*. The duration of a test is sometimes more conveniently taken as elapsed observed life than as elapsed time.

Two particular stopping rules which have received considerable attention in the literature are stopping at a predetermined observed life and stopping at a predetermined number of failures. Truncation of either rule yields the rule, stop at x^* failures or at an observed life of y^*, whichever happens first, x^* and y^* being predetermined values. The truncated rule is monotone. Both this and the following rule may be useful where resources are limited.

If it costs D_1 dollars to operate testing apparatus during one unit of observed life, and if each article which fails must be replaced at a cost of D_2 dollars, D dollars being available to conduct the test, the stopping rule, stop when $(x + 1) \cdot D_2 + yD_1 \geqq D$, assures a test costing between D and $D - D_2$ dollars.

11. Other models. Models other than the binomial, hypergeometric, and Poisson exist, though the author does not know of applications. Paths can be generated by urn schemes such as Pólya's ([9], pp. 82–83) in which each ball is replaced after drawing together with another of the same color. Another urn model consists of a pilot urn and m other urns. The pilot urn contains balls numbered from one to m, not necessarily in equal proportions; the other urns contain black and white balls in varying proportions. A ball is drawn without replacement from the pilot urn and another ball drawn without replacement from the urn thus numbered, the path then moving one unit in the direction indicated by the second ball. Other models can be based on some of the distributions described in [13].

12. Multivariate models. Although only two dimensional models have been described, similar models of higher dimension arise, as in sampling without replacement from an urn containing balls of more than two colors, in sampling from a multinomial distribution, and in treating a vector valued Poisson process. A digression in this direction seems appropriate.

Theorems 1 and 2 generalize directly to higher dimensions. It is also simple to alter Corollary 2A to obtain unbiased estimates of the covariances of the estimated transition probabilities (or densities) from $\mathbf{0}$ to the adjacent points. For example, in the discrete case, calling $\mathbf{1} = (1, 0, 0, 0, \cdots, 0)$ and $\mathbf{a} = (0, 1, 0, 0, \cdots, 0)$, an unbiased estimate of $P(\mathbf{0}, \mathbf{1})$ is $g(\mathbf{0}, \mathbf{1}, \mathbf{z})$, etc. If $P(\mathbf{0}, \mathbf{a}) = A + BP(\mathbf{1}, \mathbf{1} + \mathbf{a})$, and unbiased estimate of the covariance of $g(\mathbf{0}, \mathbf{1}, \mathbf{z})$ and $g(\mathbf{0}, \mathbf{a}, \mathbf{z})$ is

$$g(\mathbf{0}, \mathbf{1}, \mathbf{z})g(\mathbf{0}, \mathbf{a}, \mathbf{z}) - Ag(\mathbf{0}, \mathbf{a}, \mathbf{z}) - Bg(\mathbf{0}, \mathbf{1} + \mathbf{a}, \mathbf{z}). \tag{12.1}$$

Extension of results based on monotone boundaries is less straightforward. It is still true that the unbiased estimates of Theorem 2 are functions of the coordinates and type of the terminating point only, but as there are, in a k dimensional model, $2^k - 2$ ways in which access to a point can be cut off partially, $2^k - 1$ types exist. Moreover, it is no longer true that every boundary point is either fully accessible or has a unique fully accessible predecessor; that is, z' cannot be defined.

13. A remark on inference. The coordinates of the terminating point are a sufficient statistic for a particular boundary, and no more than these coordinates and the shape of the boundary need be reported. If the boundary be monotone, the coordinates and type of the terminating point determine unbiased estimates, the likelihood function hence maximum likelihood estimates, and confidence curves [2], [3] hence confidence intervals of all sizes. As this information is enough to specify the usual estimates, inferences, etc., the author submits that when the stopping rule is monotone the least that should and the most that need be reported are the fact of monotonicity and the coordinates and type of the terminating point.

REFERENCES

[1] BIRNBAUM, ALLAN (1954). Statistical methods for Poisson processes and exponential populations. *J. Amer. Statist. Assoc.* **49** 254–266.

[2] BIRNBAUM, ALLAN (1961). A unified theory of estimation, I. *Ann. Math. Statist.* **32** 112–135.

[3] BIRNBAUM, ALLAN (1961). Confidence curves: an omnibus technique in testing statistical hypotheses. *J. Amer. Statist. Assoc.* **56** 246–249.

[4] CHAPMAN, DOUGLAS G. (1952). Inverse, multiple and sequential sample censuses. *Biometrics* **8** 286–306.

[5] DEGROOT, MORRIS H. (1959). Unbiased sequential estimation for binomial populations. *Ann. Math. Statist.* **30** 80–101.

[6] EPSTEIN, BENJAMIN (1959). Statistical techniques in life testing. Duplicated report, Wayne State University, Michigan.

[7] EPSTEIN, BENJAMIN (1960). Tests of the validity of the assumption that the underlying distribution is exponential, part I. *Technometrics* **2** 83–101.

[8] EPSTEIN, BENJAMIN and SOBEL, MILTON (1954). Truncated life tests in the exponential case. *Ann. Math. Statist.* **25** 555–564.

[9] FELLER, WILLIAM (1950). *An Introduction to Probability Theory and its Applications*, **1**. Wiley, New York.

[10] FINNEY, D. J. (1949). On a method of estimating frequencies. *Biometrika* **36** 223–234.

[11] GIRSHICK, M. A., MOSTELLER, FREDERIC and SAVAGE, L. J. (1946). Unbiased estimates for certain sampling plans with applications. *Ann. Math. Statist.* **17** 13–23.

[12] HALDANE, J. B. S. (1945). On a method of estimating frequencies. *Biometrika* **33** 222–225.

[13] KEMP, C. D. and KEMP, A. W. (1956). Generalized hypergeometric distributions. *J. Roy. Statist. Soc. Ser. B* **18** 202–211.

[14] KNIGHT, WILLIAM R. (1959). Exponential and subexponential distributions in statistical life testing. PhD thesis, University of Toronto, Toronto.

[15] MATUSZEWSKI, T. I. (1962). Some properties of Pascal distribution for finite population. *J. Amer. Statist. Assoc.* **57** 172–174.

[16] NADLER, JACK (1960). Inverse binomial sampling plans when an exponential distribution is sampled with censoring. *Ann. Math. Statist.* **31** 1201–1204.

[17] SANDELIUS, MARTIN (1950). An inverse sampling procedure for bacterial plate counts. *Biometrics* **6** 291–292.