ON THE COMPUTATION OF THE SAMPLING CHARACTERISTICS OF A GENERAL CLASS OF SEQUENTIAL DECISION PROBLEMS

By G. E. Albert

University of Tennessee

Summary. The connection between the theory of random walks and Wald's theory [10], [11] of sequential probability ratio tests of hypotheses has been remarked by several authors. In particular, Kemperman [5] has exploited that connection to obtain integral equations for the determination of the decision probabilities and the expected sample size of a Wald sequential test. It is the purpose of the present paper (1) to generalize Kemperman's integral equations to apply to a fairly extensive class of sequential multiple decision problems, and (2) to indicate methods of obtaining practical results from such integral equations.

Part I of the paper is purely theoretical. It presents the integral equations already mentioned and generalizes a method of obtaining upper and lower bounds for their solutions that seems to have been first published by Kemperman [5] and Snyder [7] simultaneously.

In Part II the possibilities for application of the general theory are illustrated by a discussion of Wald's sequential tests for simple alternatives on the parameter of a distribution, under the hypothesis that a sufficient statistic for that parameter exists. In particular, it is shown that the Kemperman-Snyder method for obtaining bounds for the solutions of the integral equations may be used to obtain substantial improvements over the bounds given by Wald for the operating characteristics of the test for simple alternatives on the mean of a normal distribution. Methods of numerical analysis are indicated that might be useful in a well-equipped computing laboratory for further improvement of the bounds.

It is clear from the results obtained here that the methods used, coupled with extensive numerical work, should yield definitive improvements over Wald's approximate methods for setting the decision boundaries and estimating the sample size moments for sequential tests. It is hoped that the decision rule adopted in Part I is sufficiently general that the theory will provide a useful tool in the design and study of multiple decision problems.

Part I. Theory

1. Random walks and decision problems. Let R be an abstract space of points x. For each fixed x in R let $P(A \mid x)$ denote a conditional probability measure defined on a Borel field $\mathfrak{F}$ of subsets A of R. It will be assumed that, for each set A of the field $\mathfrak{F}$, $P(A \mid x)$ is a Borel measurable function of x.

Let d_i, $i = 1, 2, \cdots, r$, denote a set of r distinct decisions, one of which is to be made about the probability function $P(A \mid x)$ as a result of a sequentially-performed experiment as described below. The symbol d_0 will denote the deci-

Received 1/6/53, revised 10/1/53.

sion to continue experimentation at any step of the sequence instead of making one of the terminal decisions. Wald [12] gives a more detailed description of this type of decision problem. Define r nonnegative measurable functions $\pi_i(x)$, $i = 1, 2, \cdots, r$, on R with the property $\sum_{i=1}^{r} \pi_i(x) \leqq 1$ for each x in R and let $\pi_0(x) \equiv 1 - \sum_{i=1}^{r} \pi_i(x)$.

The experiment takes the form of a random walk described as follows. Let x_0 be an arbitrary point in R. Make one and only one of the decisions d_i with respective probabilities $\pi_i(x_0)$, $i = 0, 1, 2, \cdots, r$. If the decision made is d_i, $i \geqq 1$, the experiment is terminated; if it is d_0, a point x_1 is drawn from R using the distribution specified by $P(A \mid x_0)$. Again make one and only one of the decisions d_i, $i = 0, 1, 2, \cdots, r$, with respective probabilities $\pi_i(x_1)$. If the decision made is d_i, $i \geqq 1$, the experiment is terminated; if it is d_0, a point x_2 is drawn from R using the distribution specified by $P(A \mid x_1)$ and so on until one of the desired decisions d_i, $i = 1, 2, \cdots, r$, is made at a point x_n, $n = 0, 1, 2, \cdots$. In order to guarantee that the duration n of the experiment be finite with probability one, the following assumption will be made.

Assumption 1. There is a constant ρ, $0 \leqq \rho < 1$, and an integer M such that for all x_0 in R and all $m \geqq M$ the inequality

$$(1) \quad \int_R \cdots \int_R \left[\prod_{i=1}^{m} \pi_0(x_i)\right] P(de_m \mid x_{m-1})\, P(de_{m-1} \mid x_{m-2}) \cdots P(de_1 \mid x_0) \leqq \rho$$

is satisfied.

The notation in (1) is similar in most respects to that used by Doob [1]. The integral is to be interpreted as an iterated Lebesgue-Stieltjes integral whenever such exists. Doob's paper gives further discussion. A subscript on the symbol de denotes the variable of integration.

Let $p_{ik}(x_0)$, $i = 1, 2, \cdots, r$, $k = 0, 1, 2, \cdots$, denote the probability that the terminal decision d_i is made at the stage $n = k$ if x_0 is the arbitrary starting point of the experiment. Evidently, for each $i = 1, 2, \cdots, r$,

$$(2) \quad \begin{cases} p_{i0}(x_0) = \pi_i(x_0), \\ p_{i\,k+1}(x_0) = \displaystyle\int_R \cdots \int_R \left[\prod_{j=0}^{k} \pi_0(x_j)\right] \pi_i(x_{k+1})\, P(de_{k+1} \mid x_k) \\ \qquad\qquad\qquad \cdot P(de_k \mid x_{k-1}) \cdots P(de_1 \mid x_0) \\ \qquad\quad = \pi_0(x_0) \displaystyle\int_R p_{ik}(y)\, P(de_y \mid x_0) \qquad k \geqq 0. \end{cases}$$

The probability that the experiment be terminated at the stage $n = k$, regardless of which one of the terminal decisions is made, is given by

$$(3) \quad p_k(x_0) \equiv \sum_{i=1}^{r} p_{ik}(x_0).$$

It is obvious that $p_0(x_0) = 1 - \pi_0(x_0)$ and that the rest of the functions (3) satisfy the same recurrence relations (2) relative to k as do the functions $p_{ik}(x_0)$.

That the duration n of the experiment is a random variable on the range $n = 0, 1, 2, \cdots$ will follow from

$$\lim_{k\to\infty} P(n > k) = \lim_{k\to\infty} \sum_{j=k+1}^{\infty} p_j(x_0) = 0. \tag{4}$$

To establish (4), note that by (2), for any $j \geqq 1$,

$$p_j(x_0) = \int_R \cdots \int_R \left[\prod_{s=0}^{j-1} \pi_0(x_s)\right] [1 - \pi_0(x_j)]\, P(de_j \mid x_{j-1}) \cdots P(de_1 \mid x_0).$$

Thus, for any integer k,

$$\sum_{j=0}^{k} p_j(x_0) = 1 - \int_R \cdots \int_R \left[\sum_{s=0}^{k} \pi_0(x_s)\right] P(de_k \mid x_{k-1}) \cdots P(de_1 \mid x_0). \tag{5}$$

Let N be any integer and suppose that $k \geqq NM$, where M is as specified in Assumption 1. Clearly the integral on the right in (5) is dominated by the quantity $\pi_0(x_0)\rho^N$ and the limit (4) follows.

2. Some integral equations. Let $q_k(x)$, $k = 0, 1, 2, \cdots$ be any sequence of functions defined and Borel-measurable on R which satisfy, on R, the conditions

$$\begin{cases} 0 \leqq q_0(x) \leqq K < \infty, \\ q_{k+1}(x) = \pi_0(x) \int_R q_k(y)\, P(de_y \mid x), & k \geqq 0. \end{cases} \tag{6}$$

The termination probabilities p_{ik} and p_k defined in the preceding section are admissible q_k.

THEOREM 1. *If the sequence $q_k(x)$, $k = 0, 1, 2, \cdots$ satisfies* (6) *and λ is any complex number for which*

$$\rho|\lambda|^M < 1, \tag{7}$$

where ρ and M are the constants specified in Assumption 1, *then the series*

$$\sum_{j=0}^{\infty} \lambda^j q_j(x) \tag{8}$$

converges uniformly and absolutely for x in R to the bounded solution $u(x; \lambda)$ of the integral equation

$$u(x; \lambda) = q_0(x) + \lambda\pi_0(x) \int_R u(y; \lambda)\, P(de_y \mid x). \tag{9}$$

PROOF. The convergence properties of the series (8) are easily established. Using (6) repeatedly, for any x_0 in R one has

$$q_{j+1}(x_0) = \int_R \cdots \int_R \left[\prod_{i=0}^{j} \pi_0(x_i)\right] q_0(x_{j+1})\, P(de_{j+1} \mid x_j) \cdots P(de_1 \mid x_0).$$

By Assumption 1, if $j \geqq NM$, one has

$$|\lambda^{j+1} q_{j+1}(x_0)| \leqq K\rho^N |\lambda|^{j+1}$$

so that the series $\sum_{j=m}^{\infty} | \lambda^{j+1} q_{j+1}(x_0) |$ is termwise dominated by the series of constants

$$K \sum_{N=1}^{\infty} \rho^N \sum_{j=NM}^{(N+1)M-1} | \lambda |^{j+1} = K \sum_{j=1}^{M} | \lambda |^j \sum_{N=1}^{\infty} (\rho | \lambda |^M)^N$$

which converges by (7). That the series (8) satisfies the integral equation (9) is an immediate consequence of the uniformity of convergence and the relation (6); for,

$$q_0(x_0) + \sum_{j=1}^{\infty} \lambda^j q_j(x_0) = q_0(x_0) + \lambda \pi_0(x_0) \int_R \sum_{j=1}^{\infty} \lambda^{j-1} q_{j-1}(y)\, P(de_y \mid x_0).$$

Uniqueness of the solution $u(x; \lambda)$ of (9) follows almost exactly as in the proof given by Wasow [14], pp. 201–202.

There are several important applications of Theorem 1. First, if the termination probabilities are chosen as the $q_k(x)$, the series (8) becomes the generating function of that sequence. Setting $\lambda = \exp(z)$, one obtains the moment generating function of the distribution of the duration n.

COROLLARY 1. *The moment generating function of the distribution of the duration n exists for all values of its argument z such that $\rho | \exp(z) |^M < 1$ and, when considered as a function of the starting point x_0 of the experiment, satisfies the integral equation* (9) *with $\lambda = \exp(z)$, $q_0(x) = 1 - \pi_0(x)$, and $x = x_0$.*

As a consequence of (4), the probability is unity that one of the decisions d_i, $i = 1, 2, \cdots, r$, will be made. If the arbitrary starting point of the experiment is x_0, the probability $P_i(x_0)$ of making the ith decision is given by $P_i(x_0) = \sum_{k=0}^{\infty} p_{ik}(x_0)$, $i = 1, 2, \cdots, r$. Since the functions (2) are admissible q_k in Theorem 1, and $\lambda = 1$ is allowed by (7), one has a second application of the theorem.

COROLLARY 2. *The probability $P_i(x_0)$ of making the decision d_i, $i = 1, 2, \cdots, r$, at some finite stage of the experiment satisfies the integral equation*

$$P_i(x_0) = \pi_i(x_0) + \pi_0(x_0) \int_R P_i(y)\, P(de_y \mid x_0). \tag{10}$$

It follows from Corollary 1 and the assumption $\rho < 1$ that the moment-generating function of the duration n exists for positive real values of its argument z. It follows that all of the moments of the distribution of the duration exist. Let $M_k(x_0)$ denote the kth moment. Formal differentiation of the moment generating function of n leads to the third result.

COROLLARY 3. *The kth moment $M_k(x_0)$ of the distribution of the duration n of an experiment that begins at x_0 satisfies the equation*

$$M_k(x_0) = \pi_0(x_0) \sum_{s=0}^{k-1} \binom{k}{s} \int_R M_s(y)\, P(de_y \mid x_0) + \pi_0(x_0) \int_R M_k(y)\, P(de_y \mid x_0).$$

In particular, the first two moments satisfy the equation (9) of Theorem 1 with $\lambda = 1$, $x = x_0$, and $q_0(x_0) = \pi_0(x_0)$ for the first moment, and with $q_0(x_0) = 2M_1(x_0) - \pi_0(x_0)$ for the second moment.

Wasow [14] established Theorem 1 and Corollary 3 in a special case using a similar proof. Kemperman [5] gave Corollaries 2 and 3 for cumulative sums on the real line. He assumed fixed decision boundaries for each step.

3. Approximate solutions. In many of the important applications of the integral equations of the last section, the equations are likely to be very difficult if not impossible to solve by methods that will give useful numerical results. It may be necessary to resort to approximate solutions. To this end, the following results are offered.

THEOREM 2. *Let $P_k(A \mid x)$, $k = 0, 1, 2, \cdots$ be a sequence of probability functions of the type defined in Section 1 which satisfy Assumption 1 with constants ρ and M that are independent of k. Let $q_0(x)$ be nonnegative and bounded on R, and suppose that λ is any complex constant satisfying (7). Denote by $u_k(x; \lambda)$, $k = 0, 1, 2, \cdots$ the solutions of the integral equations obtained by replacing P in (9) by P_k.*

If, for the sequence of functions

$$f_k(x) \equiv \lambda\pi_0(x) \left\{ \int_R u_0(y; \lambda)\, P_0(de_y \mid x) - \int_R u_0(y; \lambda)\, P_k(de_y \mid x) \right\}, \qquad k = 0, 1, 2, \cdots$$

it is true that $\lim_{k\to\infty} f_k(x) = 0$ is satisfied uniformly for x in R, then $\lim_{k\to\infty} u_k(x; \lambda) = u_0(x; \lambda)$ uniformly for x in R.

PROOF. The sequence of difference functions $w_k(x; \lambda) \equiv u_0(x; \lambda) - u_k(x; \lambda)$, $k = 1, 2, 3, \cdots$ satisfies the sequence of integral equations

$$w_k(x; \lambda) = f_k(x) + \lambda\pi_0(x) \int_R w_k(y; \lambda)\, P_k(de_y \mid x). \tag{11}$$

By iteration

$$w_k(x; \lambda) = f_k(x) + \pi_0(x) \sum_{N=1}^{\infty} \lambda^N \int_R \cdots \int_R \left[\prod_{s=1}^{N-1} \pi_0(x_s) \right] f_k(x_N) \cdot P_k(de_N \mid x_{N-1}) \cdots P_k(de_1 \mid x).$$

Let ϵ_k denote the least upper bound of $| f_k(x) |$ on R. Splitting the series at $N = M$ and treating the sum from $N = 1$ to $N = M$ by an obvious method and the remaining sum as in the convergence proof in Theorem 1, one obtains

$$| w_k(x; \lambda) | \leqq \epsilon_k \left\{ 1 + \sum_{N=1}^{M} | \lambda |^N + \sum_{N=1}^{\infty} \rho^N \sum_{j=NM}^{(N+1)M-1} | \lambda |^{j+1} \right\}.$$

The series in brackets is convergent and independent of x and k. Since $\lim_{k\to\infty} \epsilon_k = 0$, the theorem is proved.

The remaining results of this section extend and exploit a method published simultaneously by Kemperman [5] and Snyder [7]. For each point x in a space B let $G(A \mid x)$ be a measure function defined on a field $\mathfrak{F}'$ of Borel subsets of B, and suppose that for each A in $\mathfrak{F}'$, $G(A \mid x)$ is a measurable function of x. It will be assumed that for some constant ρ', $0 < \rho' < 1$, and integer M',

$$\text{(12)} \qquad \int_B \cdots \int_B G(de_1 \mid x_0)\, G(de_2 \mid x_1) \cdots G(de_{M'} \mid x_{M'-1}) \leqq \rho' \qquad \text{all } x_0 \text{ in } B.$$

Let $f(x)$ be measurable and $0 \leqq f(x) \leqq K' < \infty$ on B, and suppose that $w(x)$ is the solution of the integral equation

$$\text{(13)} \qquad w(x) = f(x) + \int_B w(y)\, G(de_y \mid x).$$

DEFINITION. A nonnegative function $h(x)$ will be called an upper (lower) function for $w(x)$ if its iterate

$$\text{(14)} \qquad h_1(x) \equiv f(x) + \int_B h(y)\, G(de_y \mid x)$$

is less than or equal (greater than or equal) to $h(x)$ for all x in B.

Let $\varphi_A(x)$ denote the characteristic function of the subset A of B; $\varphi_A(x)$ has the value one if x is in A and zero if x is not in A. For sets A of the Borel field $\mathfrak{F}'$, define the functions $G_1(A \mid x) \equiv G(A \mid x)$ and

$$G_s(A \mid x) \equiv \int_B \cdots \int_B \varphi_A(x_s)\, G(de_s \mid x_{s-1}) \cdots G(de_1 \mid x), \quad s = 2, 3, 4, \cdots;$$

also, define the functions $f_s(x) \equiv f(x) + \sum_{k=1}^{s-1} \int_B f(y)\, G_k(de_y \mid x)$,

$$s = 1, 2, 3, \cdots.$$

By iteration of (13), $w(x)$ also satisfies

$$\text{(15)} \qquad w(x) = f_s(x) + \int_B w(y)\, G_s(de_y \mid x), \qquad s = 2, 3, 4, \cdots.$$

Repeated iterations of a function $h(x)$ by the operator on the right in (14) define the functions

$$\text{(16)} \qquad h_s(x) \equiv f_s(x) + \int_B h(y)\, G_s(de_y \mid x), \qquad s = 1, 2, 3, \cdots,$$

of which h_1 is a special case.

THEOREM 3. *If $h(x)$ is an upper (lower) function for the solution $w(x)$ of* (13), *each of the functions $h_s(x)$, $s = 1, 2, 3, \cdots$, given by* (16) *is an upper (lower) function for $w(x)$.*

The proof is evident. The fundamental utility of upper and lower functions is expressed by

THEOREM 4. *An upper (lower) function for the solution $w(x)$ of* (13) *is an upper (lower) bound for $w(x)$ on B.*

PROOF. Consider an upper function for $w(x)$. (The proof is similar for the case of a lower function.) Let U be the least upper bound of $w(x) - h(x)$ on B and assume that $h(x)$ is not an upper function. Then $U > 0$ and for an arbitrary positive number ϵ there is a point x' in B for which $w(x') - h(x') > U - \epsilon$. By

Theorem 3, $h_s(x) \leqq h(x)$, $s \geqq 1$, so $w(x') - h_{M'}(x') > U - \epsilon$, where M' is the integer used in (12). By subtraction of (16) from (15) with $s = M'$,

$$\int_B [w(y) - h(y)] G_{M'}(de_y \mid x') > U - \epsilon.$$

By (12) and the definition of U, $U\rho' > U - \epsilon$; thus $U < \epsilon/(1 - \rho')$. Since ϵ was arbitrary, the assumption $U > 0$ has been contradicted.

The following results may be useful in connection with Theorem 3 for the determination of upper and lower functions.

THEOREM 5.

(i) *If* (12) *is satisfied for* $M' = 1$, *the least upper bound and greatest lower bound of the function* $f(x)/[1 - \int_B G(de_y \mid x)]$ *over* B *are, respectively, the smallest constant upper function and largest constant lower function for* (13).

(ii) *Let* $G^*(A \mid x)$ *and* $f^*(x)$ *be functions of the types* $G(A \mid x)$ *and* $f(x)$. *If the solution* $w^*(x)$ *of the integral equation* $w^*(x) = f^*(x) + \int_B w^*(y)\, G^*(de_y \mid x)$ *is nonnegative, it is an upper function for* (13) *if* $f^*(x) \geqq f(x)$ *and* $G^*(A \mid x) \geqq G(A \mid x)$ *for all* x *in* B *and all sets* A *in* $\mathfrak{F}'$. *It is a lower function for* (13) *if these inequalities are reversed.*

(iii) *Let* $h(x)$ *be an upper function for* (13) *and* $h_1(x)$ *its iterate* (14) *and suppose that* $u(x)$ *is any measurable function satisfying* $[h_1(x)/h(x)] \leqq u(x) \leqq 1$. *Choosing* $f^*(x) = f(x)$ *and* $G^*(de_y \mid x) = u(y)\, G(de_y \mid x)$ *in (ii),* $w^*(x)$ *is a lower function for* (13) *and the average* $v(x) = \frac{1}{2}[h_1(x) + w^*(x)]$ *is an upper function for* (13) *having the property* $v(x) - w(x) \leqq \frac{1}{2}[h(x) - w(x)]$.

PROOF. The proofs of (i), (ii) and the first part of (iii) are trivial and are omitted. To see that $v(x)$ defined in (iii) is an upper function for (13), it must be shown that its iterate $v_1(x)$ by the operator in (14) is dominated by $v(x)$ over B. This will be so if

$$\int_B [h_1(y) + w^*(y)]\, G(de_y \mid x) \leqq \int_B [h(y) + w^*(y)u(y)]\, G(de_y \mid x).$$

Now $w^*(y) \leqq h(y)$ so

$$\int_B w^*(y)[1 - u(y)]\, G(de_y \mid x) \leqq \int_B h(y)[1 - u(y)]\, G(de_y \mid x)$$

$$\leqq \int_B h(y)[1 - h_1(y)/h(y)]\, G(de_y \mid x).$$

The desired result follows at once from this inequality.

In cases where the integral equation (13) takes the simple form

$$w(x) = f(x) + \int_a^b w(y)\, K(x, y)\, dy \tag{17}$$

with a continuous kernel $K(x, y)$, another device for improving upper and lower functions is available. Let $R(x, y)$ be the kernel that is reciprocal to $K(x, y)$. That is, for each x in (a, b), $R(x, y)$ satisfies the equation

$$R(x, y) = K(x, y) + \int_a^b K(x, z)\, R(z, y)\, dz.$$

Let $H(x)$ denote the difference $h(x) - h_1(x)$ between a function $h(x)$ and its iterate $h_1(x)$ by the operator on the right in (17).

Theorem 6. *If $h(x)$ is an upper [lower] function for the solution $w(x)$ of* (17) *and if, for each x in (a, b), the function $J(x, y)$ is a lower [upper] function for $R(x, y)$, then the function $H_1(x)$ defined by*

$$H_1(x) \equiv h(x) - H(x) - \int_a^b H(y)\, J(x, y)\, dy$$

is an upper [lower] function for $w(x)$ and $H_1(x) \leqq h(x)$ [$H_1(x) \geqq h(x)$].

Proof. The proof will be indicated for the first case. The iterate $H_2(x)$ of $H_1(x)$ by the operator on the right in (17) is easily shown to be given by

$$H_2(x) = H_1(x) + \int_a^b H(y) \left\{ J(x, y) - R(x, y) + \int_a^b K(x, z)[R(z, y) - J(z, y)]\, dz \right\} dy.$$

The function $J(x, y)$ is a lower function for $R(x, y)$ so that the bracketed expression in the integrand above is nonpositive. Thus $H_2(x) \leqq H_1(x)$.

The reader will recognize a similarity between the conditions imposed upon the measure function $G(A \mid x)$ and those imposed earlier upon the probability measure $P(A \mid x)$. The definitions and results on upper and lower functions may be rephrased in such a way as to apply directly to the integral equation (9) and its special cases, if the condition (7) is replaced by the requirement that λ be a positive real number for which $\rho\lambda^{M+1} < 1$. This may be useful in some problems. The applications of upper and lower functions to be made in the remainder of this paper are of a slightly different character and the discussion has been phrased in terms that are suitable for those applications.

Part II. Some Illustrative Applications

4. A sequential probability ratio test. Wald's sequential tests of hypotheses [10], [11] are based in part upon his theory of cumulative sums of independent, identically distributed random variables [9]. As Kemperman [5] has shown, an alternative treatment is available in terms of the integral equations given in Part I above. The remainder of this paper will be given to a study of the integral equations for the risk probabilities and the expected sample size of the sequential probability ratio test for simple alternatives on the parameter of a distribution. It will be assumed that a sufficient statistic for the parameter exists.

Let $g(u;\, \theta)$ be a probability density function on the real line of the form

$g(u; \theta) = \exp \{p(\theta)k(u) + r(u) + q(\theta)\}$ in which $k(u)$ and $p(\theta)$ are monotone increasing functions of the variable u and the parameter θ respectively. The sequential probability ratio test for the hypothesis $\theta = \theta_1$ against the alternative $\theta = \theta_2$ will be considered at some length. (Girshick [3] treated this problem by another method.)

Write $\xi = p(\theta)$ and $v = k(u)$. By the monotoneity assumptions on these functions, they have single-valued inverses $\theta = P(\xi)$ and $u = K(v)$. Define $Q(\xi) = q[P(\xi)]$ and $R(v) = r[K(v)]$. The probability density function for the variable v is

$$f(v; \xi) = K'(v) \exp \{\xi v + R(v) + Q(\xi)\}. \tag{18}$$

The test may be stated in terms of $\xi_1 = p(\theta_1)$ versus $\xi_2 = p(\theta_2)$, $\xi_2 > \xi_1$.

The logarithm of the probability ratio is

$$z = \log [g(u; \theta_2)/g(u; \theta_1)] = 2(v + \tau)\Delta$$

$$2\Delta = \xi_2 - \xi_1 \text{ and } \tau = [Q(\xi_2) - Q(\xi_1)]/2\Delta.$$

The cumulative distribution of z is

$$F(z; \xi) = \int_{-\infty}^{z/2\Delta} f(t - \tau; \xi)\, dt.$$

It will be assumed that a positive constant δ exists such that at least one of the inequalities $F(-\delta; \xi) > 0$ or $1 - F(\delta; \xi) > 0$ is satisfied.

To apply the general theory of sections 2 and 3 above to the Wald test of ξ_1 versus ξ_2, proceed as follows. There are two terminal decisions, $d_1 : \xi = \xi_1$ and $d_2 : \xi = \xi_2$. Choose two positive constants a and b and let the decision probability functions $\pi_i(x)$, $i = 1, 2$, be the characteristic functions of the sets $x \leqq -b$ and $x \geqq a$ respectively. The complementary probability $\pi_0(x)$ is the characteristic function of the interval $-b < x < a$.

The random walk begins at an arbitrarily chosen real number z_0 (Wald's specialization to $z_0 = 0$ will be made later). The successive points x_N, $N = 0, 1, 2, \cdots$ of the walk are the cumulative sums $x_N = \sum_{i=0}^{N} z_i$ of a sequence $\{z_i\}$ of independent values of the probability ratio z. By Corollaries 2 and 3, if ξ is the true value of the parameter, the probability $P_i(x_0 ; \xi)$ of making the decision d_i, $i = 1, 2$, and the expected duration $M_1(x_0 ; \xi)$ for a test that begins at $z_0 = x_0$ satisfy the integral equations:

$$2P_1(x_0; \xi)\Delta = \begin{cases} 1 & x_0 \leqq -b, \\ \displaystyle\int_{-\infty}^{-b} f\left(\frac{y - x_0}{2\Delta} - \tau; \xi\right) dy \\ \displaystyle\qquad + \int_{-b}^{a} P_1(y; \xi) f\left(\frac{y - x_0}{2\Delta} - \tau; \xi\right) dy & -b < x_0 < a, \\ 0 & x_0 \geqq a; \end{cases} \tag{19}$$

$$(20)\quad 2P_2(x_0;\xi)\Delta = \begin{cases} 0 & x_0 \leqq -b, \\ \int_a^\infty f\left(\frac{y-x_0}{2\Delta} - \tau;\xi\right) dy & \\ \qquad + \int_{-b}^{a} P_2(y;\xi) f\left(\frac{y-x_0}{2\Delta} - \tau;\xi\right) dy & -b < x_0 < a, \\ 1 & x_0 \geqq a; \end{cases}$$

$$(21)\quad 2M_1(x_0;\xi)\Delta = \begin{cases} 0 & x_0 \leqq -b \text{ or } x_0 \geqq a, \\ 2\Delta + \int_{-b}^{a} M_1(y;\xi) f\left(\frac{y-x_0}{2\Delta} - \tau;\xi\right) dy & -b < x_0 < a. \end{cases}$$

The risk probabilities for Wald's test of θ_1 against θ_2 are as follows. The probability of rejecting θ_2 when it is correct is the value $P_1(0;\xi_2)$ of the solution of (19). The probability of rejecting θ_1 when it is correct is the value $P_2(0;\xi_1)$ of the solution of (20). Some consideration will be given later to the approximate evaluation of these risks and the expected sample size $M_1(0;\xi)$.

The dispersion of the sample size could be studied from its variance. By Corollary 3 the second moment of the distribution of the sample size satisfies the equation obtained by replacing the leading term 2Δ on the right in (21) by

$$2\Delta[2M_1(x_0\,;\xi) - 1].$$

It might be better to study the dispersion of n directly from its distribution. If x_0 and ξ have the same meanings as above and if k is an arbitrary integer, $P(n \geqq k; x_0, \xi) = \sum_{N=k}^{\infty} p_N(x_0\,;\xi)$, where the terms on the right are (3) with their dependence upon ξ put into evidence. Applying the recurrence relations (2),

$$(22)\quad P(n \geqq k; x_0, \xi) = p_k(x_0;\xi) + (2\Delta)^{-1} \int_{-b}^{a} P(n \geqq k; y, \xi) f\left(\frac{y-x_0}{2\Delta} - \tau;\xi\right) dy \quad -b < x_0 < a.$$

The initial term on the right in (22) is given by the iterated integral

$$p_k(x_0;\xi) = \int_{-b}^{a} dx_1 \cdots \int_{-b}^{a} dx_{k-1} \left(\int_{-\infty}^{-b} + \int_a^\infty\right) \left[\prod_{i=1}^{k} \frac{1}{2\Delta} f\left(\frac{x_i - x_{i-1}}{2\Delta} - \tau;\xi\right)\right] dx_k.$$

Since the normal distribution has the form (18) with the mean as the parameter, the results to be obtained are applicable to that case. The binomial and Poisson probability functions also have the form (18). Thus, with proper care in interpretation of integrals as sums, the results may also be used for tests of hypotheses on the proportion p of the binomial distribution or the mean of the Poisson distribution. In the latter connection see Herbach [4].

5. Monotone character of the risk probabilities. It is important to note conditions under which, for each fixed x_0 in the interval $(-b, a)$, the probabilities

$P_1(x_0\,;\,\xi)$ and $P_2(x_0\,;\,\xi)$ are respectively monotone decreasing and monotone increasing functions of the parameter ξ. Consider (19). Substituting $t = y/2\Delta$ and $t_0 = x_0/2\Delta$ and using the identity

$$f(t - t_0 - \tau;\, \xi) = f(t - t_0 - \tau;\, \xi_2) \exp\{(\xi - \xi_2)(t - t_0 - \tau) + Q(\xi) - Q(\xi_2)\}$$

to display dependence upon ξ, iteration of (19) gives the series solution

$$\begin{aligned} P_1(2t_0\Delta;\, \xi) &= \int_{-\infty}^{-b/2\Delta} f(t - t_0 - \tau; \xi_2) \exp\{(\xi - \xi_2)(t - t_0 - \tau) + Q(\xi) - Q(\xi_2)\}\, dt \\ &+ \sum_{N=1}^{\infty} \int_{-b/2\Delta}^{a/2\Delta} dt_1 \cdots \int_{-b/2\Delta}^{a/2\Delta} dt_N \int_{-\infty}^{-b/2\Delta} \left[\prod_{i=1}^{N+1} f(t_i - t_{i-1} - \tau; \xi_2)\right] \\ &\exp\{(\xi - \xi_2)[t_{N+1} - t_0 - (N+1)\tau] \\ &\qquad + (N+1)[Q(\xi) - Q(\xi_2)]\}\, dt_{N+1}. \end{aligned} \tag{23}$$

Differentiating (23) with respect to ξ termwise under the integral signs has the effect of inserting the quantities $\{t_{N+1} - t_0 + (N+1)\,[Q'(\xi) - \tau]\}$ as factors in the respective integrands of the terms of (23). Since the ranges of t_0 and t_{N+1} make the difference $t_{N+1} - t_0$ negative, it is clear that the condition

$$Q''(\xi) \leqq 0 \qquad \text{for all } \xi \tag{24}$$

is sufficient to make $P_1(x_0\,;\,\xi)$ monotone decreasing in ξ in the range $\xi \geqq \xi_2$. Operating similarly with (20), the condition (24) is also sufficient for $P_2(x_0\,;\,\xi)$ to be monotone increasing in ξ in the range $\xi \leqq \xi_1$. It is assumed that the formal manipulations can be justified.

These monotone properties of the exact risk functions are important in the extension of the validity of the test of the simple alternatives ξ_1 and ξ_2 to composite alternatives $\xi \leqq \xi_1$ and $\xi \geqq \xi_2$.

6. Bounds for the solutions of (19), (20) and (21). Convergent series solutions such as (23) are available for all of the integral equations associated with the sequential test described in Section 4. They appear to be useless in general for computational purposes. The possibilities for applying the upper and lower function concepts discussed in Section 3 will be illustrated here by a particularly simple analysis that yields bounds for the solutions of (19), (20) and (21). Specializations are given in the next section to the test of a simple hypothesis and alternative on the mean of a normal distribution with known variance. A comparison with Wald's bounds [10] and [11] on the risk probabilities and expected sample size of the test is given there. Kemperman [5] obtained much weaker results of the same type under weaker hypotheses on $F(z;\,\xi)$. His methods were similar to those to be used here. The potentialities of the theory for further work should be evident. To obtain bounds for the risk probabilities and expected sample size that are substantially better than those given below is clearly possible, but involves extensive numerical work.

Consider the problem of finding upper and lower functions for the solution of the integral equation (19) when $\xi^* = \xi_2$. Iteration of a constant by the operator on the right in (19) is simple. Also, since

$$e^{-y}f\left(\frac{y-x}{2\Delta} - \tau; \xi_2\right) \equiv e^{-x}f\left(\frac{y-x}{2\Delta} - \tau; \xi_1\right), \tag{25}$$

iteration of the function e^{-x} is simple. Upper and lower functions for $P_1(x; \xi_2)$ will be obtained from linear combinations of the pair $(1, e^{-x})$. It is convenient to use the form

$$h(x; \delta, \gamma) \equiv (e^{-x} - \gamma)/(\delta - \gamma), \qquad \delta - \gamma > 0, \tag{26}$$

where δ and γ are constants. Write $G(z; \xi) = 1 - F(z; \xi)$ where $F(z; \xi)$ is as defined in Section 4. For (26) to be an upper function for $P_1(x; \xi_2)$, the inequality

$$\delta F(-b - x; \xi_2) + \gamma G(a - x; \xi_2) \leqq e^{-x}F(-b - x; \xi_1) + e^{-x}G(a - x; \xi_1) \tag{27}$$

is sufficient. Reversal of this inequality will make (26) a lower function for $P_1(x; \xi_2)$. Evidently (27) will be satisfied by the pair of constants $\delta = \delta_1$ and $\gamma = \gamma_1$ defined by

$$\delta_1 = \min_{(-b,a)} \frac{e^{-x}F(-b - x; \xi_1)}{F(-b - x; \xi_2)}, \quad \gamma_1 = \min_{(-b,a)} \frac{e^{-x}G(a - x; \xi_1)}{G(a - x; \xi_2)}. \tag{28}$$

The reverse inequality to (27) will be satisfied by the pair $\delta = \delta_2$ and $\gamma = \gamma_2$ defined by

$$\delta_2 = \max_{(-b,a)} \frac{e^{-x}F(-b - x; \xi_1)}{F(-b - x; \xi_2)}, \quad \gamma_2 = \max_{(-b,a)} \frac{e^{-x}G(a - x; \xi_1)}{G(a - x; \xi_2)}. \tag{29}$$

Thus, the functions $h(x; \delta_i, \gamma_i)$, $i = 1, 2$, obtained by using (28) and (29) in (26) will be respectively upper and lower bounds for the solution $P_1(x; \xi_2)$ of (19) when $\xi = \xi_2$.

A similar argument using the integral equation (20) with $\xi = \xi_1$ and the form $g(x; \delta', \gamma') \equiv (e^{x} - \gamma')/(\delta' - \gamma')$ leads to upper and lower functions for $P_2(x; \xi_1)$. For an upper function use

$$\delta_1' = \min_{(-b,a)} \frac{e^{x}G(a - x; \xi_2)}{G(a - x; \xi_1)}, \quad \gamma_1' = \min_{(-b,a)} \frac{e^{x}F(-b - x; \xi_2)}{F(-b - x; \xi_1)}. \tag{30}$$

For a lower function use the pair δ_2' and γ_2' obtained by using maxima in (30) instead of minima.

Consider next the integral equation (21) for the first moment $M_1(x; \xi)$ of the duration of the sequential experiment. Let $\mu = \int_{-\infty}^{\infty} (v + \tau)f(v; \xi)\, dv$ and $\nu =$

$\int_{-\infty}^{\infty} (v + \tau)^2 f(v; \xi)\, dv$ where $f(v; \xi)$ is given by (18). For any choice of a constant k one has

$$
(31)\quad \begin{cases} \dfrac{x + k}{2\Delta} + \mu \equiv (2\Delta)^{-1} \displaystyle\int_{-\infty}^{\infty} \dfrac{y + k}{2\Delta} f\left(\dfrac{y - x}{2\Delta} - \tau; \xi\right) dy, \\ \left(\dfrac{x + k}{2\Delta}\right)^2 + \dfrac{\mu(x + k)}{\Delta} + \nu \equiv (2\Delta)^{-1} \displaystyle\int_{-\infty}^{\infty} \left(\dfrac{y + k}{2\Delta}\right)^2 f\left(\dfrac{y - x}{2\Delta} - \tau; \xi\right) dy. \end{cases}
$$

Let λ_1, λ_2 and k be any constants and define the function

$$
(32)\quad R(x; \xi, \lambda_1, \lambda_2, k) \equiv (\lambda_1 \mu + \lambda_2 \nu)\, M_1(x; \xi) + \lambda_1 \frac{x + k}{2\Delta} + \lambda_2 \left(\frac{x + k}{2\Delta}\right)^2.
$$

Using the integral equation (21) and the identities (31), it is easily shown that the function (32) satisfies the integral equation

$$
(33)\quad \begin{aligned} R(x; \xi, \lambda_1, \lambda_2, k) &= g(x; \xi, \lambda_1, \lambda_2, k) \\ &\quad + (2\Delta)^{-1} \int_{-b}^{a} R(y; \xi, \lambda_1, \lambda_2, k)\, f\left(\frac{y - x}{2\Delta} - \tau; \xi\right) dy. \end{aligned}
$$

The leading term is defined by

$$
(34)\quad \begin{aligned} &g(x; \xi, \lambda_1, \lambda_2, k) \\ &\equiv -\Delta^{-1}\lambda_2 \mu(x + k) + (2\Delta)^{-1} \left(\int_{-\infty}^{-b} + \int_{a}^{\infty}\right) f\left(\frac{y - x}{2\Delta} - \tau; \xi\right) \\ &\qquad \cdot \left[\lambda_1 \frac{y + k}{2\Delta} + \lambda_2 \left(\frac{y + k}{2\Delta}\right)^2\right] dy. \end{aligned}
$$

Let A_1 and A_2 be further constants and define the functions

$$
(35)\quad S_i(x; \xi, \lambda_1, \lambda_2, k) \equiv (-1)^{i-1}[R(x; \xi, \lambda_1, \lambda_2, k) - A_i P_i(x; \xi)], \quad i = 1, 2.
$$

Using (19), (20) and (33), one finds that the functions (35) satisfy the integral equations

$$
(36)\quad \begin{aligned} S_i(x; \xi, \lambda_1, \lambda_2, k) &= f_i(x; \xi, \lambda_1, \lambda_2, k) \\ &\quad + \frac{1}{2\Delta} \int_{-b}^{a} S_i(y; \xi, \lambda_1, \lambda_2, k) f\left(\frac{y - x}{2\Delta} - \tau; \xi\right) dy \qquad i = 1, 2, \end{aligned}
$$

with leading terms defined in terms of (34) and the constants A_i by

$$
(37)\quad f_i(x; \xi, \lambda_1, \lambda_2, k) = \begin{cases} g(x; \xi, \lambda_1, \lambda_2, k) - A_1 F(-b - x; \xi) & i = 1, \\ A_2 G(a - x; \xi) - g(x; \xi, \lambda_1, \lambda_2, k) & i = 2. \end{cases}
$$

If, for some specific choice of λ_1, λ_2 and k, the constants A_1 and A_2 are chosen so that the functions (37) are nonnegative on the interval $-b \leqq x \leqq a$, the solutions (35) of the integral equations (36) will be nonnegative on that interval.

It follows that, for such choices of the A_i with distinct values k_i, $i = 1, 2$, of the constant k, one has bounds for the quantity $(\lambda_1\mu + \lambda_2\nu) M_1(x; \xi)$ given by

$$(38)\quad \begin{aligned} A_1P_1(x;\xi) - \lambda_1 \frac{x+k_1}{2\Delta} - \lambda_2\left(\frac{x+k_1}{2\Delta}\right)^2 &\leqq (\lambda_1\mu + \lambda_2\nu) M_1(x;\xi) \\ &\leqq A_2P_2(x;\xi) - \lambda_1\frac{x+k_2}{2\Delta} - \lambda_2\left(\frac{x+k_2}{2\Delta}\right)^2. \end{aligned}$$

If μ is not close to zero, the choices $\lambda_1 = 1$ and $\lambda_2 = 0$ are useful and the bounds (38) are of the type given by Wald in [10]. If μ is close to zero, a better choice is $\lambda_1 = 0$ and $\lambda_2 = 1$; in this case the bounds (38) are of the type given by Wald [13], using $x = 0$.

7. Examples. Suppose that the distribution (18) is,

$$f(v; \xi) = \varphi(v - \xi) = (2\pi)^{-\frac{1}{2}} \exp(\xi v - \tfrac{1}{2}v^2 - \tfrac{1}{2}\xi^2).$$

The sequential test under consideration provides a test on one value ξ_1 of the mean against a larger value ξ_2. It is clear that all of the hypotheses that have been placed upon (18) are satisfied.

In this case the functions to be minimized or maximized in (28), (29) and (30) are monotone. For example, consider, in (28) and (29), the functions

$$\chi_1(x) = e^{-x}F(-b - x; \xi_1)/F(-b - x; \xi_2)$$

and

$$\chi_2(x) = e^{-x}G(a - x; \xi_1)/G(a - x; \xi_2).$$

One finds that $\tau + \xi_1 = -\Delta$ and $\tau + \xi_2 = \Delta$. On setting $t = (b + x)/2\Delta$ in $\chi_1(x)$ and $t' = (a - x)/2\Delta$ in $\chi_2(x)$, these functions take the forms

$$\chi_1(x) = e^{b-2t\Delta}G(t - \Delta)/G(t + \Delta)$$

and

$$\{\chi_2(x)\}^{-1} = e^{a-2t'\Delta}G(t' - \Delta)/G(t' + \Delta).$$

Wald ([10], pp. 140–141,) has shown that the function of t or of t' involved here is a monotone decreasing function of its argument. It follows that $\chi_1(x)$ and $\chi_2(x)$ are monotone decreasing functions of x.

Using $G_1(u) = (2\pi)^{-\frac{1}{2}}\int_u^\infty \exp(-\tfrac{1}{2}x^2)\,dx$ and $A = (a + b)/2\Delta$, the definitions (28) and (29) give

$$(39)\quad \begin{cases} \delta_1 = e^{-a}G_1(A - \Delta)/G_1(A + \Delta), & \gamma_1 = e^{-a}G_1(\Delta)/G_1(-\Delta), \\ \delta_2 = e^{b}G_1(-\Delta)/G_1(\Delta), & \gamma_2 = e^{b}G_1(A + \Delta)/G_1(A - \Delta). \end{cases}$$

Using (39) in (26) and evaluating at $x = 0$, one has the bounds

$$(40)\quad (1 - \gamma_2)/(\delta_2 - \gamma_2) \leqq P_1(0; \xi_2) \leqq (1 - \gamma_1)/(\delta_1 - \gamma_1)$$

for the probability of accepting ξ_1 when ξ_2 is correct in Wald's test. Replacement of δ_1 by e^b and γ_2 by e^{-a} in (40) gives the bounds obtained by Wald. It will be shown that $\delta_1 > e^b$ and $\gamma_2 < e^{-a}$, from which it will follow that (40) gives better bounds than those quoted. Actually, the improvement is only very slight. Further improvements will be indicated in the next section.

From the asymptotic series $G_1(u) = \varphi(u)[u^{-1} - u^{-3} + 3u^{-5} - \cdots]$, valid for large positive u, one obtains

$$\delta_1 = e^b[(A - \Delta)^{-1} - (A - \Delta)^{-3} + \cdots]/[(A + \Delta)^{-1} - (A + \Delta)^{-3} + \cdots],$$

$$\gamma_2 = e^{-a}[(A + \Delta)^{-1} - (A + \Delta)^{-3} + \cdots]/[(A - \Delta)^{-1} - (A - \Delta)^{-3} + \cdots],$$

where A has the same meaning as in (39). Obviously $\delta_1 > e^b$ and $\gamma_2 < e^{-a}$. The bounds for $P_2(x; \xi_1)$ may be treated in a similar way.

Wald [10] gave an explicit calculation of bounds for the expected duration of the experiment for the case in which μ is not close to zero. His results are derivable from (38). To see this, choose $\lambda_1 = 1$, $\lambda_2 = 0$, $k_1 = -a$ and $k_2 = b$. For the functions (37) to be nonnegative on $-b \leqq x \leqq a$, it is sufficient that

$$A_1 \leqq g(x; \xi, 1, 0, -a)/F(-b - x; \xi),$$

$$A_2 \geqq g(x; \xi, 1, 0, b)/G(a - x; \xi)$$

be satisfied on that range. These inequalities are satisfied by

$$(41)\qquad \begin{aligned} A_1 &= \min_{(-b,a)} \left[-A + \chi_3\left(\frac{b + x}{2\Delta} + \xi + \tau\right)\right], \\ A_2 &= \max_{(-b,a)} \left[A - \chi_3\left(\frac{a - x}{2\Delta} - \xi - \tau\right)\right], \end{aligned}$$

where A has the same meaning as in (39) and

$$\chi_3(t) \equiv t - \int_t^\infty x\varphi(x)\, dx/G_1(t).$$

The function $\varphi(x)$ is the standardized normal probability density. By Wald [10, p. 144], $\chi_3(t)$ is a monotone increasing function of t. It follows from (41) that

$$A_1 = -A + \xi + \tau - (\varphi(\xi + \tau))/(G_1(\xi + \tau))$$

and

$$A_2 = A + \xi + \tau + (\varphi(\xi + \tau))/(F_1(\xi + \tau)).$$

Evaluation of (38) at $x = 0$ using the various constants chosen above gives results that are in agreement with (4.13), (4.14), (4.20) and (4.24) of Wald [10].

As remarked in Section 4, the binomial and Poisson probability functions have the form (18). With proper care in the interpretation of integrals as sums, the results of Section 6 may be adapted for tests of simple hypotheses and alternatives on the parameters of these distributions. Again the results are com-

TABLE 1

Bounds on $P_1(0, \xi_2)$ for $a = b = 2$ and $\Delta = 0.25$

	Lower	Upper
By Wald	0.07941	0.12459
By (39), (40)	0.08008	0.11717
Optimal	0.08089	0.10997

parable with those obtained by Wald [10] and Herbach [4]. Since the normal example given above furnishes an adequate comparison of this type, explicit calculations will be omitted here.

8. Possible further improvements and applications. The upper and lower bounds derived in Section 6 for the risk probabilities and expected sample size tend toward the exact solutions of their respective integral equations as Δ approaches zero. For small values of Δ these bounds may be sufficiently close to the exact solutions to be regarded as solutions by the designer of practical experiments. For values of Δ greater than .01, say, more accurate solutions are needed.

For specific problems such as the test considered in detail in Section 7, it should be possible to choose the constants involved in the form (26) in such a manner as to minimize the upper function and maximize the lower function. Explicitly, find the largest values of δ and γ such that $\delta - \gamma > 0$ and (27) is satisfied and the smallest values for which the reverse inequality to (27) is satisfied. The procedure is numerical and involves trial and error. In the normal example of Section 7, it is found that the values of γ_1 and δ_2 given in (28) and (29) are the best possible values, but that δ_1 and γ_2 may be improved. In the special case of that example defined by $a = b = 2$ and $\Delta = 0.25$, the equations (39) give $\delta_1 = 7.8514$, $\gamma_1 = 0.09071$, $\delta_2 = 11.024$ and $\gamma_2 = 0.12736$. The optimal values are $\gamma_2 = 0.11785$ and $\delta_1 = 8.3593$. For this case the bounds on $P_1(0, \xi_2)$ given by Wald [10], by (39) and (40), and by the above optimal choices of constants are shown in Table 1. This example shows that further improved bounds are needed.

It is interesting to note that the approximate value $\overline{P_1(0, \xi_2)} = (1 - e^{-a})/(e^b - e^{-a})$ obtained from equation (3.35) of Wald [10] gives the value 0.11920 for the numerical example tabulated above. Since this value lies outside the range allowed by the optimal bounds, a need is indicated for a better approximate formula for $P_1(0, \xi_2)$.

Various devices are known for the numerical approximation of solutions of integral equations. One or more such devices might be used in combination with the upper and lower function concepts with considerable success in a well equipped computing laboratory. In particular, careful attention should be given to the possible utility of Theorem 6.

The versatility of the theory presented in Part I should make it a useful tool in the definition and study of (1) more general termination rules than that used in Part II, (2) multidimensional problems, and (3) decision problems that involve more than two terminal decisions. The writer hopes to investigate such applications at a later date and hopes that this paper will serve to interest others in joining the search.

REFERENCES

[1] J. L. Doob, "Stochastic processes with an integral valued parameter," *Trans. Amer Math. Soc.*, Vol. 44 (1938), pp. 87–103.

[2] M. A. Girshick, "Contributions to the theory of sequential analysis, I," *Ann. Math. Stat.*, Vol. 17 (1946), pp. 123–143.

[3] M. A. Girshick, "Contributions to the theory of sequential analysis, II, III," *Ann. Math. Stat.*, Vol. 17 (1946), pp. 282–298.

[4] L. H. Herbach, "Bounds for some functions used in sequentially testing the mean of a Poisson distribution," *Ann. Math. Stat.*, Vol. 19 (1948), pp. 400–405.

[5] F. Kemperman, *The General One-Dimensional Random Walk with Absorbing Barriers with Applications to Sequential Analysis*, University of Amsterdam thesis, Excelsiors Foto-offset, 's-Gravenhage, 1950.

[6] Jack Silber, "Multiple sampling for variables," *Ann. Math. Stat.*, Vol. 19 (1948), pp. 246–256.

[7] W. S. Snyder, "Calculations for maximum permissible exposure to thermal neutrons," *Nucleonics*, Vol. 6 (1950), pp. 46–50.

[8] Charles Stein, "A note on cumulative sums," *Ann. Math. Stat.*, Vol. 17 (1946), pp. 498–499.

[9] A. Wald, "On cumulative sums of random variables," *Ann. Math. Stat.*, Vol. 15 (1944), pp. 283–296.

[10] A. Wald, "Sequential tests of statistical hypotheses," *Ann. Math. Stat.*, Vol. 16 (1945), pp. 117–186.

[11] A. Wald, *Sequential Analysis*, John Wiley and Sons, 1947.

[12] A. Wald, *Statistical Decision Functions*, John Wiley and Sons, 1950.

[13] A. Wald, "Some improvements in setting limits for the expected number of observations required by a sequential probability ratio test," *Ann. Math. Stat.*, Vol. 17 (1946), pp. 466–474.

[14] W. Wasow, "On the duration of random walks," *Ann. Math. Stat.*, Vol. 22 (1951), pp. 199–216.