A NOTE ON THE FUNDAMENTAL IDENTITY OF SEQUENTIAL ANALYSIS

By R. R. Bahadur

University of Chicago[1]

1. Introduction. This note points out that the fundamental identity of sequential analysis [1] can be regarded as a special case of a formula for the probability that the sampling terminates at some finite stage. This viewpoint, explored in Sections 2 and 3, provides proofs of the identity, and of its differentiability under the expectation sign, that seem more intuitive than the proofs in the literature ([1], [2], [3], [5], [6]).

The formula also has application to the well-known problem (cf., e.g., [7], [8]) of evaluating the probability of eventual termination of a random walk on the real line, in the case when there is one fixed barrier and a drift away from the barrier. Some upper and lower bounds on the probability in question are obtained in Section 4.

In concluding this introduction, the writer wishes to thank his colleague L. J. Savage for discussions and suggestions that have made a substantial contribution to this work.

2. A Formula for $P(n < \infty)$. Let x be a real valued random variable with distribution function F. It is assumed that the moment generating function

$$\phi(t) = \int_{-\infty}^{\infty} e^{tx}\, dF \tag{1}$$

exists for every real t in some neighbourhood of $t = 0$. Throughout this note, t is restricted to real values for which ϕ exists.

Let $x_{(\infty)} = (x_1, x_2, \cdots, \text{ad inf})$ denote a sequence of independent and identically distributed observations on x. Consider a fixed sequential sampling procedure, that is, a set of rules for observing the components $x_1, x_2, \cdots$, of $x_{(\infty)}$ one by one, such that at each stage the decision whether experimentation is to continue is a (possibly randomised) function of the observed values in hand at that stage. (Cf., e.g., [6], [9]). Let n denote the total number of components x_m observed in a given instance. It is assumed that the sampling procedure is closed under F, that is,

$$P(n < \infty \mid F) = 1. \tag{2}$$

The procedure is otherwise arbitrary.

Write $s = x_1 + \cdots + x_n$ and

$$\psi(t, n, s) = [\phi(t)]^{-n} e^{ts} \tag{3}$$

if $n < \infty$, and write $\psi = 1$ (say) if $n = \infty$.

Received February 7, 1957; revised July 12, 1957.

[1] Now at the Indian Statistical Institute, Calcutta.

THEOREM 1. *For every t,*

$$E(\psi(t, n, s)| F) = P(n < \infty \mid G_t), \tag{4}$$

where

$$dG_t = [\phi(t)]^{-1}e^{tx} dF. \tag{5}$$

For each t, G_t defined by (5) is clearly a probability distribution function, so that $\{G_t\}$ is an exponential family of alternative distributions of x, with F a member of this family. Such families of distributions have been studied in various statistical contexts, including that of sequential analysis (cf., e.g., [1], [10], [11]). It may be added, however, that the notion of alternative distributions is not essential to this paper, and the introduction here of the family $\{G_t\}$ could be regarded as a device in the study of the given sampling rule when F obtains. This device is, of course, a familiar one in probability theory (cf., e.g., [3], [12], [13], [14]).

To establish Theorem 1, for each $m = 1, 2, \cdots$ let $R^{(m)}$ denote the nonsequential sample space of exactly m observations, that is, of points

$$(x_1, \cdots, x_m) = x_{(m)}$$

say. For each m, let $\alpha_m(x_{(m)})$ be the conditional probability of the event $n = m$ given $x_{(\infty)}$. The sequence $\alpha_1, \alpha_2, \cdots$ of functions on $R^{(1)}, R^{(2)}, \cdots$ characterizes the given sampling procedure, and is, of course, independent of the distribution of x (cf., e.g., [9]). For each m, let $F^{(m)}$ denote the distribution function of $x_{(m)}$ when F obtains, that is, $F^{(m)}(x_1, \cdots, x_m) = \prod_{i=1}^{m} F(x_i)$.

Let v denote the total outcome of the sequential experiment, that is, $v = (x_1, \cdots, x_n)$ if $n < \infty$ and $v = x_{(\infty)}$ if $n = \infty$. We note that if h is a real valued function of v such that $E(| h | \mid F) < \infty$, and (2) holds, then

$$E(h \mid F) = \sum_{m=1}^{\infty} \int_{R^{(m)}} \alpha_m \cdot h_m \, dF^{(m)} \tag{6}$$

where $h_1, h_2, \cdots$ is the (essentially unique) sequence of functions on $R^{(1)}$, $R^{(2)}, \cdots$ such that $h_m = h$ when $n = m$, (cf., e.g., [9]). The right side of (6) is an absolutely convergent series; in fact, h is integrable if and only if

$$\sum_m \int \alpha_m \cdot | h_m | \, dF^{(m)} < \infty.$$

In accordance with the above notation, for each $m = 1, 2, \cdots$ let s_m denote the nonsequential random variable $x_1 + \cdots + x_m$. Then it is easy to establish (4) thus:

$$\begin{aligned} P(n < \infty \mid G_t) &= \sum_{m=1}^{\infty} P(n = m \mid G_t) \\ &= \sum_{m=1}^{\infty} \int_{R^{(m)}} \alpha_m \cdot dG_t^{(m)} \\ &= \sum_{m=1}^{\infty} \int_{R^{(m)}} \alpha_m \cdot \phi^{-m} \cdot e^{ts_m} \, dF^{(m)} \quad \text{by} \quad (5) \\ &= E(\psi \mid F) \quad \text{by} \quad (2), (3), \text{ and } (6). \end{aligned} \tag{7}$$

3. Wald's identity. It follows from Theorem 1 that Wald's identity, namely

$$E[\psi(t, n, s) | F] = 1, \tag{8}$$

holds for a given value of t if and only if

$$P(n < \infty \mid G_t) = 1 \tag{9}$$

for the same t value.

It follows easily from the preceding remark that a sufficient condition for the validity of (8) for all t is that there exist a finite k such that $P(n < k \mid F) = 1$. The same remark, together with the strong law of large numbers and the law of the iterated logarithm (cf. e.g., [15]), also yields the following sufficient condition for the same conclusion (assuming that F is a non-degenerate distribution): there exists an h, $-\infty < h < \infty$, and two sequences $\{a_m\}$ and $\{b_m\}$ such that

(i) $a_m = mh + o(\sqrt{m \log \log m})$, $b_m = mh + o(\sqrt{m \log \log m})$ as $m \to \infty$ and

(ii) for each $x_{(\infty)} = (x_1, x_2, \cdots)$, either $n < \infty$ or $a_m < s_m < b_m$ for all sufficiently large m. This condition seems weaker than other structural conditions of the same type in the literature for the validity of (8) for all t.

It may be noted here that the first paragraph of this section also suggests examples where Wald's identity fails to hold for all t. This was pointed out to the writer by L. J. Savage. The discussion in Section 4 concerns a general example of this sort.

Next, we shall describe an alternative sufficient condition for the validity of (9) (and thereby of (8)) in an assigned neighbourhood of $t = 0$, given (2) and (5). This condition does not depend on the detailed structure of the sampling rule; it is, essentially, that under F the joint moment generating function of n and s exist in a sufficiently large neighbourhood of the origin. As it happens, the condition also assures the validity of differentiation under the expectation sign in (8), that is, $D^k\psi(t, n, s)$ is integrable and

$$E(D^k\psi \mid F) = 0 \quad \text{for} \quad k = 1, 2, \cdots \tag{10}$$

and each t in the neighbourhood, where $D^k \equiv d^k / dt^k$.

Let I be an open interval including $t = 0$ such that $\phi(t)$ exists for each t in I.

THEOREM 2. *Suppose that corresponding to each t in I there exists a*

$$z > -\log_e \phi(t) \tag{11}$$

such that

$$E(e^{ts+zn} \mid F) < \infty. \tag{12}$$

Then (8), (9), *and* (10) *hold for t in I.*

The proof of Theorem 2 will be indicated later, but first some remarks by way of discussion of its hypothesis.

REMARK 1. Let C denote the set of all points (t, z) in the plane such that

(12) holds. Then (a) C is a convex set, (b) $(t, z) \varepsilon C$ implies $(t, z_0) \varepsilon C$ for all $z_0 \leqq z$, since n is non-negative, and (c) each point on the graph of the function $z = -\log \phi(t)$ is in C, by (3) and (4). It follows, in particular, that (d) the hypothesis of Theorem 2 is that the graph of $-\log \phi(t)$ lie in the interior of C, at least when t is restricted to I. It should be noted also (e) that $z = -\log \phi$ is a concave function of t, possessing derivatives of all orders, with $z(0) = 0$, and $z'(0) = -E(x \mid F)$, by (1). The facts (a), (b), (c), (d), and (e) are useful in the proof of Theorem 2, and also in the deduction of special sufficient conditions for the validity of (9) and (10). (Cf. remarks 3 and 4 below).

Remark 2. In the statement of Theorem 2, as also in remark 1(d) above, the hypothesis is stated in terms of the given distribution F. The hypothesis can also be stated in terms of the associated distributions G_t, as follows: for each t in I, the *conditional* moment generating function of n given $n < \infty$ exists in some neighbourhood of zero when G_t obtains, that is, $E(e^{\delta n} \mid G_t, n < \infty) = \sum_m e^{\delta m} P(n = m \mid G_t, n < \infty) < \infty$ for some $\delta > 0$. This alternative formulation follows readily from (5), (6), (11) and (12).

Remark 3. Suppose that ϕ exists for all t. If

$$E(e^{zn} \mid F) < \infty \tag{13}$$

for some $z > \sup \{-\log \phi(t): -\infty < t < \infty\}$, and if

$$E(e^{ts} \mid F) < \infty \tag{14}$$

for all t, then (8), (9), and (10) hold for all t. A stronger sufficient condition for the same conclusion is that (13) hold for all z.

Remark 4. If (13) holds for some $z > 0$, then n and s possess moments of all orders, and there exists a neighbourhood of $t = 0$ in which (8), (9) and (10) hold. A stronger sufficient condition for the same conclusion is that $E(x \mid F) \neq 0$ and (14) holds for some t of the same sign as $E(x \mid F)$. These conditions are of interest since the validity of (10) in a neighbourhood of $t = 0$ is sufficient for most applications (cf. [2]) of differentiation.

Remark 5. A theorem of Albert [3], [4] states that if ϕ exists for all t, if

$$P(x > 0 \mid F) > 0$$

and $P(x < 0 \mid F) > 0$, and if the sampling procedure is a random walk based on the cumulative sums s_m (with fixed barriers a and b, $a < 0 < b$), then (8) and (10) hold for all t. It can be shown (cf. Lemma 2 of [3] and Remark 3 above) that in this case the hypothesis of Theorem 2 is satisfied by $I = (-\infty, \infty)$, so that Albert's theorem is a special case of Theorem 2.

Remark 6. In his applications of martingale theory to sequential analysis, Doob [6] derives from Theorem 2.2 of [6] a sufficient condition for the validity of (8) for a given t. It can be shown that if this condition holds for each t in I then the hypothesis of Theorem 2 is satisfied. A fuller discussion of the relation between Doob's Theorem 2.2 in its application to the present case and Theorems 1 and 2 of this note would be worthwhile, but cannot be undertaken here.

REMARK 7. It is easy to see that $P(n < k \mid F) = 1$ for some finite k implies the hypothesis of Theorem 2. L. J. Savage has constructed examples showing that the other condition stated in the second paragraph of this section does not imply the hypothesis of Theorem 2; in fact, (10) fails for $k = 2$ in these examples.

We turn now to the proof of Theorem 2. The first step is to show that (9) holds in a sufficiently small neighbourhood of zero, that is,

$$\sum_{m=1}^{\infty} p_m(t) = 1 \tag{15}$$

for all t in the neighbourhood, where $p_m(t)$ is an abbreviation of $P(n = m \mid G_t)$, so that

$$p_m(t) = \int_{R^{(m)}} \alpha_m \cdot \{\phi(t)\}^{-m} \cdot e^{ts_m}\, dF^{(m)} \tag{16}$$

by (5). Write $\beta_m = 1 - \sum_{i=1}^{m} \alpha_i$, and $\rho(t) = \phi(2t)/\phi^2(t)$. Then, for any t in I and any m,

$$\begin{aligned} P(n > m \mid G_t) &= \int_{R^{(m)}} \beta_m\, dG_t^{(m)} \quad \text{since} \quad \beta_m = P(n > m \mid x_{(m)}) \\ &= \phi^{-m} \int_{R^{(m)}} \beta_m e^{ts_m}\, dF^{(m)} \quad \text{by} \quad (5) \\ &\leqq \phi^{-m} \left\{ \int_{R^{(m)}} \beta_m^2\, dF^{(m)} \right\}^{1/2} \cdot \left\{ \int_{R^{(m)}} e^{2ts_m}\, dF^{(m)} \right\}^{\frac{1}{2}} \\ &= \rho^{m/2} \left\{ \int_{R^{(m)}} \beta_m^2\, dF^{(m)} \right\}^{1/2} \\ &\leqq \rho^{m/2} \left\{ \int_{R^{(m)}} \beta_m\, dF^{(m)} \right\}^{1/2} \quad \text{since} \quad 0 \leqq \beta_m \leqq 1 \\ &= \sqrt{\rho(t)^m \cdot P(n > m \mid F)}. \end{aligned} \tag{17}$$

It follows from (11) and (12) with $t = 0$ that, for some $\lambda > 1$,

$$\lambda^m P(n > m \mid F) \to 0$$

as $m \to \infty$. Hence, by (17), $P(n > m \mid G_t) \to 0$ as $m \to \infty$ for each t such that $\rho(t) \leqq \lambda$. Thus (15) holds whenever $\rho(t) \leqq \lambda$. This establishes the desired conclusion, since $\lambda > 1$, ρ is continuous, and $\rho(0) = 1$.

The next step is to extend the validity of (15) to all t in I by analytic continuation, as follows. Let w denote the complex variable $t + iu$, and let $\phi(w)$ be defined by (1) in the strip $\{w : t \varepsilon I\}$. For each m let $p_m(w)$ be defined by (16) whenever $\phi(w)$ is defined and $\neq 0$. It follows from the differentiability of moment generating functions that the functions p_m are differentiable everywhere in their domain of definition. A straightforward argument based on the continuity of ϕ, Remark 1(d) above, formula (6), and the convexity of the exponential function shows that corresponding to each t in I there exists a complex neighbourhood of t, $N(t)$ say, and a convergent series of positive terms, $\sum_m c_m(t)$ say, such that $|p_m(w)| \leqq c_m(t)$ for all $w \varepsilon N(t)$ and each $m = 1, 2, \cdots$. The details of this argument are omitted. It follows hence that $\sum_m p_m(w)$ is a uniformly convergent series of analytic functions, so that $\sum_m p_m(w)$ is well defined and analytic in $N(t)$. (Cf. e.g., [16]). Since this holds for each t, it follows from the preceding paragraph that $\sum_m p_m(w) = 1$ for $w \varepsilon N(t)$ and $t \varepsilon I$; in particular, (15) holds for all $t \varepsilon I$.

It follows from uniform convergence (cf. [16]) and the conclusion of the preceding paragraph that for each t in I and every $k = 1, 2, \cdots$,

$$\sum_m (d^k/dw^k) p_m(w)$$

is well defined and $= 0$ for $w \varepsilon N(t)$; in particular, $\sum_m D^k p_m(t) = 0$. Since, as is readily seen, D^k commutes with the integral sign in (16), we have

$$\sum_m D^k p_m(t) = \sum_m \int_{R^{(m)}} \alpha_m \cdot D^k\{\phi^{-m} e^{ts_m}\}\, dF^{(m)} = 0 \text{ for } k = 1, 2, \cdots \tag{18}$$

and each t in I. Assuming for the moment that each of the functions $D^k\psi(t, n, s)$ is integrable when F obtains, it follows by inspection from (3), (6), and (18) that (10) holds for each t in I.

The next and final step in the proof is therefore to verify that each

$$D^k\psi(t, n, s)$$

is integrable. Since $D^k\psi$ is of the form $\psi \cdot \eta$ where η is a polynomial in n and s, it suffices to show that, for each t, $\psi \cdot |s|^i \cdot n^j$ is integrable for $i, j = 0, 1, 2, \cdots$. This may be established by showing that corresponding to each t in I there exist positive numbers ϵ and δ such that $\psi(t, n, s) \cdot \exp(\epsilon|s| + \delta n)$ is integrable. Since $\exp(\epsilon|s|) < \exp(\epsilon s) + \exp(-\epsilon s)$, it is easily seen from (3) and Remark 1(d) that this last condition is satisfied.

In concluding this section, we remark that Theorems 1 and 2 can be generalized, by straightforward extensions of the arguments used here, to the case when $x_1, x_2, \cdots$ is a sequence of independent but not necessarily identically distributed random variables, and each x_m takes values in k-dimensional Cartesian space, $1 \leqq k < \infty$. Another straightforward generalization that may be worth mentioning here is to the case where the sampling rule is defined for a sequence $w_1, w_2, \cdots$ of independent abstract random variables, and for each $i = 1, 2, \cdots x_i$ is a real (or vector) function of w_i.

4. An application. Let c be a positive constant, and let the sampling rule be defined thus: for any sequence $x_{(\infty)} = (x_1, x_2, \cdots)$, n is the smallest integer

m such that $s_m = x_1 + \cdots + x_m > c$, and $n = \infty$ if no such integer exists. Let G be a given distribution function such that $P(x > 0 \mid G) > 0$ and

$$P(x < 0 \mid G) > 0,$$

and such that $E(x \mid G)$ exists and is negative. It is shown in this section that if G admits a moment generating function, Theorem 1 can be used to obtain upper and lower bounds for $P(n < \infty \mid G)$. In certain special cases this method yields the exact value of $P(n < \infty \mid G)$.

The probability in question can be interpreted as the probability of ultimate ruin in playing an advantageous gamble long enough (with x_m the amount lost in the mth play, each x_m distributed according to G, and c the initial fortune of the player), and has been studied in connection with insurance theory (cf. [8], [12]). In [8] Dubourdieu has given derivations and original references to an upper bound, due to de Finetti, for this probability. The upper bounds obtained here are improvements of de Finetti's.

It is assumed henceforth that

$$\eta(h) = \int_{-\infty}^{\infty} e^{hx}\, dG \tag{19}$$

exists in a neighbourhood of $h = 0$. It then follows from the preceding hypotheses concerning G, by well-known properties of moment generating functions, that there exist uniquely determined points a and b (say), in the interior of the interval in which η exists, such that $0 < a < b$, $\eta'(a) = 0$, and $\eta(b) = 1$. We note that

$$\eta'(h) \geqq 0 \qquad \text{for } h \geqq a \tag{20}$$

and that

$$\eta(h) \begin{cases} <1 & (a \leqq h < b) \\ =1 & (h = b) \\ >1 & (h > b). \end{cases} \tag{21}$$

In (20), (21), and in what follows, h is understood to be restricted to the interior of the interval in which η exists.

Now choose and fix an $h \geqq a$ and define

$$dF_h = [\eta(h)]^{-1} e^{hx}\, dG. \tag{22}$$

Then the moment generating function of F_h, say ϕ, is given by

$$\phi(t) = \eta(t + h)/\eta(h).$$

Since $E(x \mid F_h) = \phi'(0) = \eta'(h)/\eta(h)$, it follows from (20) and the choice of h that $E(x \mid F_h) \geqq 0$. Consequently, by well-known properties of cumulative sums, $P(n < \infty \mid F_h) = 1$. Since $dG = \eta(h)e^{-hx}\, dF_h$ by (22), and

$$\eta(h) = [\phi(-h)]^{-1},$$

Theorem 1 yields the identity

$$P(n < \infty \mid G) = E(e^{-hs}[\eta(h)]^n \mid F_h), \tag{23}$$

valid for all $h \geqq a$.

Letting $h = b$ in (23), we have

$$P(n < \infty \mid G) = E(e^{-bs} \mid F_b), \tag{24}$$

by (21). Since $n < \infty$ implies $s \geqq c$, since $P(n < \infty \mid F_b) = 1$, and since $b > 0$, (24) yields

$$P(n < \infty \mid G) \leqq e^{-bc}, \tag{25}$$

which is de Finetti's inequality.

It is clear from the preceding derivation of (25) that the equality sign holds in (25) if and only if $P(s = c \mid F_b) = 1$. This condition can be shown to be equivalent to the condition that x be a discrete random variable taking only one positive value, say d, and that the negative values of x be integral multiples of d[2]. The condition is satisfied, in particular, if x takes only two values.

We turn now to the case when s can exceed c with positive probability. In this case, the effect of the 'excess over the boundary' can be estimated by means of an argument due to Wald [1][3]. It is possible and advantageous to apply the argument to (23) rather than (24), as follows.

Suppose that F_h obtains. Write $y = c$ if $n = 1$, and

$$y = c - (x_1 + \cdots + x_{n-1})$$

if $1 < n < \infty$. Then y is well defined and $0 < y < \infty$ with probability 1. Let ξ denote the conditional expectation of e^{-hx_n} given n and y. It is not difficult to see that ξ depends only on y and h; in fact

$$\xi = E(e^{-hx} \mid x \geqq y, F_h). \tag{26}$$

We observe next that $s = c - y + x_n$ with probability one. Consequently, the right side of (23) can be written as $e^{-hc} \cdot E(e^{hy} \cdot \xi(y) \cdot \eta^n \mid F_h)$. It follows hence, by regarding y as a real variable confined to positive values, and setting

$$f(h) = \inf_y \{e^{hy} \cdot \xi\}, \qquad g(h) = \sup_y \{e^{hy} \cdot \xi\}, \tag{27}$$

that

$$e^{-hc} \cdot f(h) \cdot E(\eta^n \mid F_h) \leqq P(n < \infty \mid G) \leqq e^{-hc} \cdot g(h) \cdot E(\eta^n \mid F_h). \tag{28}$$

---

[2] In this case a slight extension of the methods of this paper can be used to obtain the probability distribution of n, with ∞ a possible value of n.

[3] Wald used the argument, in the context of a random walk with two absorbing barriers, to find the maximum possible effect of the excess over a barrier on the probability of absorption in that barrier. However, the argument also yields the minimum possible effect, in Wald's context as well as in the present one.

Next, an easy calculation using (22), (26), and (27) shows that

$$f(h) = \inf_y \{e^{hy}/E(e^{hx} \mid x \geqq y, G)\},$$
$$g(h) = \sup_y \{e^{hy}/E(e^{hx} \mid x \geqq y, G)\}. \tag{29}$$

Finally, since $n \geqq 1$, we see from (21) and (28) that

$$P(n < \infty \mid G) \leqq e^{-hc} \cdot g(h) \cdot \eta(h) \tag{30}$$

for $a \leqq h \leqq b$, and

$$P(n < \infty \mid G) \geqq e^{-hc} \cdot f(h) \cdot \eta(h) \tag{31}$$

for $h \geqq b$.

The infimum of the right side of (30) with h restricted to $[a, b]$ gives, of course, the best upper bound obtainable by this method, while the supremum of the right side of (31) with h restricted to $[b, \infty]$ gives the best lower bound. In particular, taking $h = b$ in (30) and (31), we have

$$e^{-bc} \cdot f(b) \leqq P(n < \infty \mid G) \leqq e^{-bc} \cdot g(b). \tag{32}$$

Another special bound is

$$P(n < \infty \mid G) \leqq \inf_{a \leqq h \leqq b} \{e^{-hc} \cdot \eta(h)\}; \tag{33}$$

this follows from (30) since $0 < g \leqq 1$.

It is easy to see from the preceding argument that in case x is bounded from above, η can be replaced by η^k in (30), (31) and (33), where k is the least positive integer such that $P(n = k \mid G) > 0$.

In concluding this section let us consider an example. In this example,

$$G = pH + (1 - p)K \tag{34}$$

where $0 < p < 1$, H is some distribution function (possibly degenerate) confined to $(-\infty, 0]$, and

$$dK(x) = \begin{cases} \lambda e^{-\lambda x}\, dx & \text{for} \quad x > 0 \\ 0 \text{ otherwise,} \end{cases} \tag{35}$$

where λ is a positive constant. It is assumed that H (and therefore G) admits a moment generating function in a neighbourhood of the origin, and that

$$E(x \mid G) \equiv pE(x \mid H) + (1 - p)/\lambda < 0.$$

It then follows that the equation

$$\eta(b) \equiv pE(e^{bx} \mid H) + (1 - p)(\lambda/(\lambda - b)) = 1 \tag{36}$$

has a unique non-zero solution b, with $0 < b < \lambda$.

A simple calculation, which is omitted, shows that in the present case we have $f(h) = g(h) = (\lambda - h)/\lambda$ for all h, where f and g are defined by (29). Consequently, it follows from (32) that

$$P(n < \infty \mid G) = e^{-bc} \cdot (\lambda - b)/\lambda. \tag{37}$$

REFERENCES

[1] A. Wald, *Sequential Analysis*, John Wiley and Sons, Inc., New York (1947), pp. 157–164.

[2] A. Wald, "Differentiation under the expectation sign in the fundamental identity of sequential analysis", *Ann. Math. Stat.* Vol. 17 (1946), pp. 493–497.

[3] G. E. Albert, "A note on the fundamental identity of sequential analysis", *Ann. Math. Stat.*, Vol. 18 (1947), pp. 593–596.

[4] G. E. Albert, "Correction to 'A note on the fundamental identity of sequential analysis'," *Ann. Math. Stat.* Vol. 19 (1948), p. 426.

[5] D. Blackwell and M. A. Girshick, "On functions of sequences of independent chance vectors with applications to the problem of the "random walk" in k dimensions", *Ann. Math. Stat.*, vol. 17 (1946), pp. 310–317.

[6] J. L. Doob, *Stochastic processes*, John Wiley and Sons, Inc., New York (1953), pp. 350–352.

[7] W. Feller, *An introduction to probability theory and its applications*, John Wiley and Sons., Inc., New York (1950), Chap. 14.

[8] J. Dubourdieu, *Theorie mathêmatique du risque dans les assurances de repartition*, Gauthier-Villars, Paris (1952), pp. 163–174.

[9] R. R. Bahadur, "Sufficiency and statistical decision functions," *Ann. Math. Stat.*, Vol. 25 (1954), pp. 423–462, Section 8.

[10] M. A. Girshick and L. J. Savage, "Bayes and minimax estimates for quadratic loss functions", *Proc. Second Berkeley Symp. on Math. Stat. and Probability*, University of California Press (1951), p. 53.

[11] B. O. Koopman, "On distributions admitting a sufficient statistic," *Trans. Amer. Math. Soc.*, Vol. 39 (1936), p. 399.

[12] A. Wald, "Some generalizations of the theory of cumulative sums of random variables," *Ann. Math. Stat.*, Vol. 16 (1945), pp. 287–293.

[13] M. A. Girshick, "Contributions to sequential analysis. II, III," *Ann. Math. Stat.*, Vol. 17 (1946), pp. 291–298.

[14] H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," *Ann. Math. Stat.*, Vol. 23 (1952), pp. 493–507.

[15] P. Hartman and A. Wintner, "On the law of the iterated logarithm", *Amer. J. Math.*, Vol. 63 (1941), pp. 169–176.

[16] E. C. Titchmarsh, *Theory of functions*, Oxford Univ. Press., New York, (Second Edition, 1939).