OPTIMUM DECISION PROCEDURES FOR A POISSON PROCESS PARAMETER[1]

By J. A. Lechner

Westinghouse Research Laboratories, Pittsburgh

0. Introduction and summary. This paper derives and exhibits the optimum Bayes solution to the following problem: Given a continuous-time Poisson process with unknown mean occurrence rate λ; to decide whether $\lambda > k$ or $\lambda < k$. The prior distribution is taken to be of Gamma type, with positive mean and finite variance. The cost of observation is taken proportional to the length of time the process is observed, and the cost of a *wrong* decision proportional to $|\lambda - k|$. The decision rule derived is optimum (in the sense of minimum expected cost) among all non-randomized sequential rules. Some of the results hold true, of course, for other cost functions and/or prior distributions. A method for treating the same problem with the inclusion of a constant setup cost is also given.

1. General remarks; relation to other work. The problem here is to decide between the two compound hypotheses, $H_1 : \lambda < k$ and $H_2 : \lambda > k$, where λ is the mean occurrence rate of a continuous-time Poisson process, and k is a given positive constant which can and will be adjusted to unity by changing the time scale. Two types of cost need to be balanced in attacking this problem, namely the cost of observing the process and the cost of making a wrong decision; in general, a decrease in the expected value of either can be achieved only with an increase in that of the other. The two types will be considered in turn.

The cost of observation is usually taken proportional to the observation time, and will be so taken in what follows (where by adjusting the monetary unit the constant of proportionality is made to be unity). However, the simple adjustment necessary when in addition there is a fixed setup cost is also presented, in Sec. 2.8. Some of the results, of course, hold true for more general cost functions.

The cost of a wrong decision will presumably be a function of $\lambda - 1$. We define $c_1(\lambda)$ as the cost of accepting H_1 when λ is the true value, and $c_2(\lambda)$ as the cost of accepting H_2 when λ is the true value. Now $c_1(\lambda) = 0$ for $\lambda < 1$, and $c_2(\lambda) = 0$ for $\lambda > 1$, but much latitude exists in choosing $c_1(\lambda)$ for $\lambda > 1$ and $c_2(\lambda)$ for $\lambda < 1$. The functions used throughout this paper are $c_1(\lambda) = \max[\phi(\lambda - 1), 0]$ and $c_2(\lambda) = \max[\phi(1 - \lambda), 0]$, where ϕ is a known constant greater than unity, but again some of the results herein hold true more generally.

The cost of observing the process for time t and then making a decision is

Received July 19, 1960; revised March 14, 1962.

[1] Most of this paper is based on a Ph.D. thesis submitted to Princeton University in 1959, which work was supported in part by the Office of Naval Research.

therefore $t + c_i(\lambda)$, where $i = 1$ or 2 according as the decision is to accept H_1 or H_2. For any proposed decision rule R, we can define a total expected cost $L(\lambda; R)$ by

$$L(\lambda; R) = E(t; \lambda) + P(\lambda)c_1(\lambda) + (1 - P(\lambda))c_2(\lambda), \tag{1.1}$$

where t is the duration of observation, $P(\lambda)$ is the probability that at the termination of observation H_1 is accepted, and $E(t; \lambda)$ is the expected value of t for given λ and the given decision rule. Figure 1 is an example of $L(\lambda; R)$ when $\phi = 1$, R is a fixed-sample-size experiment with $t = 1$, and H_1 is accepted if no event occurs.

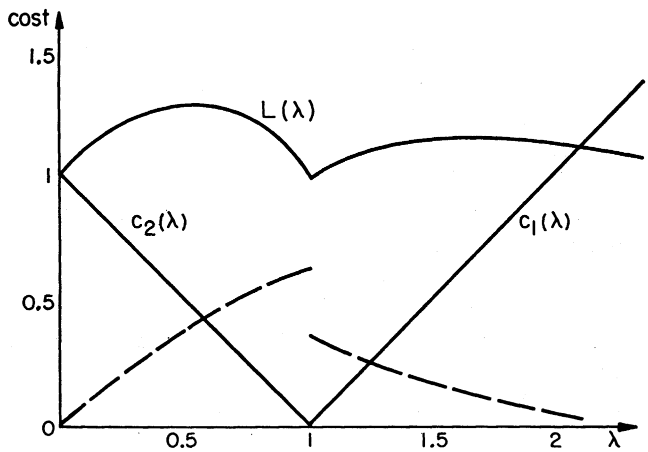

Fig. 1. $L(\lambda)$, the expected total cost for a fixed-sample-size procedure with $t = \phi = 1$ in which H_1 is accepted if no event occurs. Here $L(\lambda) = 1 + |\lambda - 1| \cdot \Pr$ (wrong decision). [The dashed curve gives the probability of a wrong decision.]

The general problem is to pick a decision procedure which minimizes $L(\lambda; R)$. But λ is unknown, and no procedure exists which minimizes $L(\lambda; R)$ for *all* λ. At this point, decision theory logically splits into minimax-type and Bayes procedures. Wald [11] developed the minimax theory, which chooses from a given class of rules one for which $\max_\lambda L(\lambda; R)$ is a minimum. (There may be more than one rule satisfying this criterion.) We pursue the alternative, which looks for a rule minimizing the so-called "Bayes risk", $L_R(p)$, i.e., the expected value of $L(\lambda; R)$ with respect to some assumed "prior" distribution $f_p(\lambda)$,

$$L_R(p) = \int_0^\infty L(\lambda; R) f_p(\lambda)\, d\lambda. \tag{1.2}$$

This procedure was used very early by Wald (see [11] and [12]), merely as a tool for finding minimax solutions. It was mentioned again by Yates as an aside in a paper [13] concerned with an *estimation* problem, for which $L(\lambda; R)$ is actually independent of λ and therefore *can* be minimized directly. Grundy, Healy, and Rees [5] followed this suggestion, and produced a method of determining the optimum *fixed-sample-size* experiment, when the problem is to determine the sign of the mean of a Normal distribution with known variance, the cost of experimentation is proportional to the number of observations taken on this Normal, a wrong decision costs $k|\mu|$, and the prior distribution is Normal. From their results, which are really results concerning the Wiener process, I have shown that the optimum sequential boundary for the Wiener process is unbounded, using essentially the lemma of this paper. This has also been shown by Chernoff ([2] and [3]), who among others developed the differential equation and boundary conditions which determine the optimum boundary. The unlimited extent of the boundary causes considerable computational difficulty, however, which seems not to have been surmounted as yet. For *bounded* loss functions and discrete observations, it was shown by Sobel [10] that the optimum decision rule must be finite.

Much work has been done on the Poisson process. Representative items are the formulae for OC and AST, for Wald-type tests, in [6], and the tables in [7]; the minimax fixed-sample-size and the minimax Wald-type tests in [1], for Normal as well as Poisson. The minimax tests for the Wiener process with observation cost bt and error costs $c|\mu - \mu_0|^r$, $0 < r \leqq 2$, are derived in [4]. Another unique approach is that of Schwarz [9]. Assuming an indifference zone where neither decision is penalized, and letting the cost c of an observation go to zero, he proves that the optimal (Bayes) region approaches a limiting shape if both the t- and n- scales are reduced by the factor $\log (c^{-1})$.

The work for the Poisson process analogous to that of Grundy, Healy, and Rees [5] for the Wiener process has not been carried through. The expected loss as a function of observation time t is a *scalloped* curve, which has in general many local minima. (If the test is terminated as soon as one particular decision is assured, the loss is decreased of course; but then in addition to the scallops, the function has a positive jump at each integer value of t.)

2. The optimum sequential decision rule. In this section, the optimum rule is characterized and a method of calculating it is presented.

2.1. *The prior distribution and the sample paths.* We assume the prior distribution to be of the Gamma type, i.e., it has density

$$f(\lambda) = e^{-\lambda t}\lambda^{y-1}t^{y}/\Gamma(y), \qquad t > 0, y \geqq 1, \tag{2.1}$$

which will also be denoted by $f_{t,y}(\lambda)$. For convenience of language only, we take y to be an integer. Then the *posterior* distribution for λ, having observed y' events in time t', is simply $f_{t+t',y+y'}(\lambda)$. For this reason, we will consider the results of experimentation as represented by a plot of $y_0 + y(t)$ against $t_0 + t$ where $y(t)$ is the number of events that have occurred up to the time t and $f_{t_0,y_0}(\lambda)$ is the prior distribution for λ. We restrict ourselves to boundaries consisting of stopping points (t, y) with a definite decision attached to each such stopping point, since (t, y) together with (t_0, y_0) determine the likelihood function for λ. The sample path thus starts at the point $p = (t_0, y_0)$, for the prior distribution $f_p(\lambda) = f_{t_0, y_0}(\lambda)$, and continues to some point $b = (t(b), y(b)) \varepsilon B$, a boundary set, at which time a decision is made dependent only on $t(b)$ and $y(b)$. Now for any point q, we write $L(q, \lambda)$ or $L(t, y, \lambda)$ for the loss incurred by an immediate decision, being at q; i.e., $L(q, \lambda) = c_i(\lambda)$, where $i = 1$ or 2 according as the decision is to accept H_1 or H_2. Since the posterior distribution for λ, having reached $q = (t, y)$, is $f_q(\lambda)$, the *expected* loss incurred by an immediate decision is given by

$$L_S(q) = L_S(t, y) = \int_0^\infty L(q, \lambda) f_q(\lambda)\, d\lambda. \tag{2.2}$$

The function of B to be minimized (by choice of B) will be denoted by $L_B(p)$, the expected loss incurred by starting at p and using B:

$$\begin{aligned} L_B(p) &= \int_0^\infty E\,[L(b, \lambda) + t(b) - t(p) \mid \lambda]\, f_p(\lambda)\, d\lambda \\ &= \int_0^\infty L(\lambda; B) f_p(\lambda)\, d\lambda, \end{aligned} \tag{2.3}$$

where $E(u \mid \lambda)$ is the expectation of u, for fixed λ, taken over the boundary B under consideration.

Having stopped at the point (t, y), we have from (2.2)

$$L_S(t, y) = \int_0^\infty c_i(\lambda) f_{t,y}(\lambda)\, d\lambda,$$

where $i = 1$ or 2 according as (t, y) is an acceptance point for H_1 or H_2. For $i = 1$, $L_S(t, y) = \phi \int_1^\infty (\lambda - 1) f_{t,y}(\lambda)\, d\lambda$ and for

$$i = 2,\ L_S(t, y) = \phi \int_0^1 (1 - \lambda) f_{t,y}(\lambda)\, d\lambda.$$

The difference Δ of these two expressions is given by

$$\Delta = \phi \int_0^\infty (\lambda - 1)\, e^{-\lambda t}\, [\lambda^{y-1} t^y / \Gamma(y)]\, d\lambda = \phi\, [(y - t)/t].$$

We minimize $L_S(t, y)$ by accepting H_1 whenever $\Delta < 0$, i.e., wherever $y < t$, and accepting H_2 when $y > t$. For $y = t$, $\Delta = 0$, and the decision is immaterial; but if H_1 is accepted, then $L(t, y, \lambda)$ is continuous from the right in t. There is as yet no assurance that minimizing $L_S(t, y)$ will in fact minimize $L_B(p)$ for fixed B; this will however be proved later on. Therefore, we will not specify $L(t, y, \lambda)$ at this point, but merely assume that $L(t, y, \lambda)$ is continuous from the right in t for each y and λ.

There is an alternative representation of the sample paths, as points, which will be useful later on. Consider the sample path which starts at $(0, 0)$, and for which the ith event occurs at t_i, $i = 1, 2, \cdots$. We represent it by the point (in infinite-dimensional Cartesian space) with coordinates $(t_1, t_2, \cdots)$. This gives us a representation for every path, and every infinite-dimensional vector $(t_1, t_2, \cdots)$ represents a path *if* $0 \leqq t_1 \leqq t_2 \leqq \cdots$; i.e., the set of all paths is mapped onto the wedge $0 \leqq t_1 \leqq t_2 \leqq \cdots$. Every path which goes through the point (t, n) has $t_n \leqq t \leqq t_{n+1}$; and every vector $(t_1, t_2, \cdots)$ with $0 \leqq t_1 \leqq \cdots \leqq t_n \leqq t \leqq t_{n+1} \leqq \cdots$ represents a path which goes through (t, n). We will be considering correspondences between sets T^* in this space, which we call the star-space, and sets T in the (t, y)-space.

The convenience of this representation lies in the existence of a probability density on any finite-dimensional subspace $\{(t_1, t_2, \cdots, t_k)\}$ of the star-space. Thus there is defined for any Borel set in this k-dimensional space a probability, which gives the probability of our sample path fulfilling certain conditions (which can be expressed using only the part of the sample path below the line $y = y_0 + k$).

We henceforth restrict our attention to that part of the star-space given as $\{(t_{y_0+1}, t_{y_0+2}, \cdots): t_0 \leqq t_{y_0+1} \leqq \cdots\}$; we will denote it by S^*.

2.2. *The boundary, and its interior.* (This section is independent of the particular cost functions used, given our assumption that $L(q, \lambda)$ is continuous from the right in $t(q)$, which is necessary to permit consideration only of closed sets B.)

We consider a particular starting point p, and any particular set B in the (t, y)-plane, where for every point $b \varepsilon B$, $y(b)$ is an integer $\geqq y(p)$, and $t(b) \geqq t(p)$. Let $\bar{B}$ denote the closure of B. We will show that $P(\lambda)$, and $E(t(b) - t(p))$ if it exists, are the same for $\bar{B}$ as for B. The points of $\bar{B} - B$ can be separated into three classes: those points (t, y) for which each interval from $(t - \epsilon, y)$ to (t, y) contains points of B, but there exists an interval from (t, y) to $(t + \epsilon, y)$ which does not contain any points of B, which we call left-accumulation points of B; the correspondingly defined right-accumulation points of B; and the two-sided accumulation points, i.e., those points such that both intervals of length ϵ having (t, y) as an end point contain points of B, for any $\epsilon > 0$. We consider these in turn.

For any left-accumulation point, there is an open interval to its right that does not contain points of $\bar{B}$. Since there can be at most a countable number of non-overlapping open intervals on the real line, the left-accumulation points on the line $y = n$ are at most countable; there is at most a countable number of lines $y = 1, y = 2, \cdots$, having left-accumulation points of B; therefore, the set of left-accumulation points is countable. Similarly, the set of right-accumulation points is countable.

Consider a particular left-accumulation point (t, n). Any path which reaches the line $y = n$ at $t_1 < t$ will not reach (t, n) since there are points of B in the open interval from (t_1, n) to (t, n). Therefore, the only paths which are affected are those which reach the line $y = n$ at t; but the probability of a jump occurring at the particular point $(t, n - 1)$ is zero.

Consider a right-accumulation point (t, n). Any path which reaches (t, n) and does *not* have a jump there will stop before going a distance ϵ, for any $\epsilon > 0$; so $E(t)$ is not altered by the addition of the point. Also, having assumed $L(q, \lambda)$ continuous from the right in $t(q)$, it follows that $L_B(p)$ is not altered. Thus we say that the path is not significantly altered by making (t, n) a point of B. Also, the probability of having a jump precisely at (t, n) is zero, so the probability of a significant change in $L_B(p)$ resulting from the addition to B of a right-accumulation point in zero. Since the one-sided accumulation points are countable, we can add them all to B without changing $L_B(p)$.

Consider the two-sided accumulation points of B. The only paths significantly changed by the addition of these two-sided accumulation points to B are those which have a double jump at one of these points, i.e., a jump from $y = n - 1$ to $y = n + 1$ at some value t, where (t, n) is a two-sided accumulation point of B. But the probability of a double jump anywhere is zero. Therefore the addition to B of the two-sided accumulation points of B does not change $L_B(p)$.

This shows that for any boundary set B for which $L_B(p)$ is defined, we can substitute $\bar{B}$, the closure of B; and $L_{\bar{B}}(p) = L_B(p)$. *Henceforth we consider only closed sets B.* Furthermore, we can add to B the points $(t(p), y(p) + k)$ for $k = 1, 2, \cdots$, since the probability of reaching any of these is zero. For convenience later, we assume these points are in B.

We will now establish the existence of a probability distribution on B for the stopping point. For this purpose, it is sufficient to show that for any *closed* one-dimensional subset B_1 of $B \cap \{y = n\}$, the set of paths reaching B_1 is measurable. Let B_1^* represent the set of paths which would reach B_1 if $B \cap \{y < b\}$ were empty. For $n = y(p)$, B_1^* is the set of paths having $t_{n+1} \geqq \inf \{t: (t, n) \in B_1\}$, and thus is Borel. Now suppose $n > y(p)$. Let $\tilde{B} = \{t: (t, n) \in B\}$ and

$$\tilde{B}_1 = \{t: (t, n) \in B_1\}.$$

Thus $\tilde{B}$ and $\tilde{B}_1$ are one-dimensional closed sets. The complement of $\tilde{B}$, being open, can be represented as a countable union of disjoint open intervals I_1, I_2, $\cdots$. Call the right hand end points of these intervals a_1, a_2, $\cdots$. Then $a_i \in \tilde{B}$ for $i = 1, 2, \cdots$. Let $J = \{i: a_i \in \tilde{B}_1\}$. Then B_1^* is the set of paths such that either $t_n \in \tilde{B}_1$ or $t_n \in I_i$ and $t_{n+1} \geqq a_i$ for some $i \in J$. Therefore, B_1^* is Borel. But the set of paths which *do* reach B_1 is the intersection of B_1^* with the set of paths which do *not* stop with $y < n$; by the induction hypothesis, treating in turn $B \cap \{y = k\}$ for $k = y(p), \cdots, n - 1$, the set of paths which *do* stop with $y < n$ is Borel; and so the result is proved.

The set of paths which reach a particular point q in the plane is just the intersection of the Borel set $\{t_{y(q)} \leqq t(q) \leqq t_{y(q)+1}\}$, the set of paths which do not stop in the set $B \cap \{y < y(q)\}$, and the set of paths which do not stop in the set $B \cap \{y = y(q), t < t(q)\}$; since all three are Borel sets, we see that for any point q, Pr (reach q) is defined. In fact, the obvious extension, treating a Borel set Q (instead of a point q) just as we treated B_1, we see that for any Borel set Q, Pr (reach Q) is defined. If we take $B_1 = B$, we find the probability of hitting B. If this probability is not unity, then there is a non-zero probability of never hitting B, so that $E(t) = \infty$, and thus B is a very unwise boundary indeed.

We now define the notion of interior point of B. (The notion is dependent on the starting point p.) Given p and B, we say:

DEFINITION 1. q is an interior point if $q \not\in B$, and there exists a path to q from p which does not first meet B.

DEFINITION 2. q is an interior point if $q \not\in B$ and for $\lambda > 0$, Pr (reach q from $p \mid \lambda) > 0$.

Denote the set of interior points by B_{int}. These two definitions are equivalent: If Pr (reach q from $p \mid \lambda) > 0$, there are paths to q from p which do not first hit B, so that $B_{\text{int}}(\text{def. } 1) \supset B_{\text{int}}\,(\text{def. } 2)$. On the other hand, assume there exists a path to $q \not\in B$, from p, which does not first meet B. Consider one such path, with the jump from $y = i$ to $y = i + 1$ occurring at $t_{i+1}(y(p) \leqq i < y(q))$. Neither (t_{i+1}, i) nor $(t_{i+1}, i + 1)$ is a point of B because the path does not meet B in this region. Since B is closed, there exist open t-intervals I_{i+1} on the line $y = i$, and I'_{i+1} on the line $y = i + 1$, each containing t_{i+1}, which do not intersect B. Therefore, any path with the jump from $y = i$ to $y = i + 1$ occurring at a t-value in the non-empty open interval $I_{i+1} \cap I'_{i+1}$, for all i with $y(p) \leqq i < y(q)$, will reach q before it reaches B. The measure of this set of paths is positive, proving that $B_{\text{int}}\,(\text{def. } 1) \subset B_{\text{int}}\,(\text{def. } 2)$.

We now prove that on the line $y = n$, B_{int} is an open (one-dimensional) set, for any positive integer n. Suppose $q \in B_{\text{int}}$ and $y(q) = n$. Then $q \not\in B$, so there is an open interval containing q which does not intersect B. Also, Pr (reach q from $p \mid \lambda > 0) > 0$. Since Pr (reach q from $(t(q), n - 1)$) is zero, there is an interval having q as right end-point and consisting entirely of points of B_{int} (otherwise, Pr (reach q from $p \mid \lambda) = 0$, since then no path which reaches the line $y = n$ at $t < t(q)$ will ever reach q). On the other hand, Pr (reach q from $p \mid \lambda > 0) > 0$, and there is an interval I_1 with q as left end point which is disjoint from B, so Pr (reach q_1 from $p \mid \lambda > 0) > 0$, for all $q_1 \in I_1$. This demonstrates that $q \in B_{\text{int}} \Rightarrow q \in I \subset B_{\text{int}}$, where I is an open interval on the line $y = y(q)$, and therefore that $B_{\text{int}} \cap \{y = n\}$ is an open (one-dimensional) set.

It is easy to see that there exists a probability distribution on B from any interior point q. The probability of reaching $B_1 \subset B$ from q is just the measure of the paths which reach B_1 by going through q, divided by the measure of the paths which go through q. The former set is a Borel set of paths, as we saw previously, while the latter set has positive probability because $q \in B_{\text{int}}$.

2.3. *A lemma.* (For a note on generality, see the end of this section.) Let us review our notation, and add a bit more.

As before, $t(q)$ and $y(q)$ denote the t- and y-coordinates respectively, of the point q.

(2.4) p denotes a starting point; $t(p) > 0$, $y(p)$ a positive integer.

(2.5) $B = \{b\}$, $t(b) \geqq t(p)$, $y(b) \geqq y(p)$; a boundary set, topologically closed, such that Pr (reach B from $p \mid \lambda > 0) = 1$.

(2.6) $C = \{c\}$, another closed set, such that $C \subset B \cup B_{\text{int}}$, and no path from p hits B before hitting C.

(2.7) q, the generic point of the plane.

(2.8) β, a measure on B; we take β so that $\beta(B_1) =$ Pr (reach B_1 from $p \mid \lambda = 1$), for any Borel subset B_1 of B. (We have proved that such a measure exists for any λ.)

(2.9) γ, a measure on C, chosen so that $\gamma(C_1) =$ Pr (first coincidence with C occurs in the set C_1, starting from $p \mid \lambda = 1$), for any Borel subset C_1 of C. (γ exists, since C is closed, for the same reason that β exists.)

(2.10) $h_p(c, \lambda)$, a non-negative function of λ and $c \, \varepsilon \, C$, measurable with respect to γ, such that:

$$\text{Pr (first reach } C \text{ in } C_1 \text{ from } p \mid \lambda) = \int_{C_1} h_p(c, \lambda) \, d\gamma,$$

for any Borel $C_1 \subset C$ and all λ.

(Note that if $\gamma(C_1) = 0$, then Pr (first coincidence with C occurs in $C_1 \mid \lambda = \lambda_0) = 0$, for any value λ_0. Thus the measure for first coincidence with C, for any λ, is γ-continuous, and thus by the Radon-Nikodym Theorem ([8] p. 196) such a function exists.)

(2.11) $k_q(b, \lambda)$, a non-negative function of λ, $b \, \varepsilon \, B$, and $q \, \varepsilon \, B_{\text{int}}$, measurable with respect to β, such that Pr (reach $B_1 \subset B$ from $q \, \varepsilon \, B_{\text{int}} \mid \lambda) = \int_{B_1} k_q(b, \lambda) \, d\beta$, for any Borel B_1.

(Note that if Pr (reach $B_1 \subset B$ from $p \mid \lambda = 1) = 0$, then Pr (reach $B_1 \subset B$ from $p \mid \lambda = \lambda_0) = 0$, and in fact, Pr (reach $B_1 \subset B$ from $q \, \varepsilon \, B_{\text{int}} \mid \lambda = \lambda_0) = 0$. Therefore, by the Radon-Nikodym Theorem, such a function exists for $q \, \varepsilon \, B_{\text{int}}$. In general, there does *not* exist such a function for $q \, \varepsilon \, B$, since, for $q \, \varepsilon \, B$, Pr (reach $B_1 \subset B$ from $q \mid \lambda) = 1$ or 0 according as $q \, \varepsilon \, B_1$ or $q \not\varepsilon \, B_1$.)

(2.12) $L(q, \lambda)$, the loss incurred, being at the point q, by taking an immediate decision, if λ is the true parameter value. We have assumed only that $L(q, \lambda)$ is continuous from the right in $t(q)$ and continuous in λ; the acceptance points will be determined later. $L(q, \lambda)$ is non-negative, since we measure loss from an immediate correct decision.

(2.13) $L_S(q)$, the expected loss, being at q, if we take an immediate decision.

(2.14) $L_B(q)$, the expected loss, being at q, if we use the boundary B.

(2.15) $f_q(\lambda)$, the Bayes density for λ at the point $q = (t, y)$, given by $f_q(\lambda) = f_{t,y}(\lambda) = e^{-\lambda t} \lambda^{y-1} t^y / \Gamma(y)$.

(2.16) $L_B(p, \lambda) = E(L(b, \lambda) + t(b) - t(p) | p, \lambda)$, which is simply $L(\lambda)$ with the dependence on p and B emphasized.

We will also assume $L_B(p) < \infty$. Since $L_S(p) < \infty$, we can always do better than an infinite loss.

We now state the LEMMA. *With B, C, p, λ, $h_p(c, \lambda)$, and $f_p(\lambda)$ defined as above,*

$$L_B(p) = \int_0^\infty \left\{ \int_C L_B(c) h_p(c, \lambda) \, d\gamma \right\} f_p(\lambda) \, d\lambda + \int_0^\infty \left\{ \int_C (t(c) - t(p)) h_p(c, \lambda) \, d\gamma \right\} f_p(\lambda) \, d\lambda. \tag{2.17}$$

This expresses $L_B(p)$ in terms of $L_B(c)$ and the expected cost of reaching C.

The PROOF is as follows. We have $L_B(p) = \int_0^\infty L_B(p, \lambda) f_p(\lambda) \, d\lambda$. Clearly,

$$L_B(p, \lambda) = \int_C L_B(c, \lambda) h_p(c, \lambda) \, d\gamma + \int_C [t(c) - t(p)] \, h_p(c, \lambda) \, d\gamma.$$

Hence,

$$L_B(p) = \int_0^\infty \int_C L_B(c, \lambda) h_p(c, \lambda) \, d\gamma \, f_p(\lambda) \, d\lambda + \int_0^\infty \int_C [t(c) - t(p)] \, h_p(c, \lambda) \, d\gamma \, f_p(\lambda) \, d\lambda.$$

It must be shown that $L_B(c, \lambda)$ can be replaced by $L_B(c)$ in the first iterated integral. But this integral can be thought of as $E(L_B(c, \lambda))$, where c and λ are random variables with joint density $h_p(c, \lambda) f_p(\lambda)$. Hence, since $f_c(\lambda)$ is the conditional density of λ given c, we have

$$\begin{aligned} E\,[L_B(c, \lambda)] &= E\{E\,[L_B(c, \lambda) \mid c]\} \\ &= E\left[\int_0^\infty L_B(c, \lambda) f_c(\lambda) \, d\lambda\right] = E\,[L_B(c)] \\ &= \int_0^\infty \int_C L_B(c) h_p(c, \lambda) \, d\gamma \, f_p(\lambda) \, d\lambda. \end{aligned}$$

This completes the proof.

2.4. *The decision; remarks.* By this lemma, we see that $L_B(p)$ is minimized, for given B and p, by minimizing $L_S(b)$ for all $b \in B$. But we have shown that $L_S(q)$ is minimized if we choose $L(q, \lambda) = c_1(\lambda)$ whenever

$$\int_0^\infty [c_2(\lambda) - c_1(\lambda)] f_q(\lambda) \, d\lambda > 0, \quad \text{and} \quad L(q, \lambda) = c_2(\lambda)$$

whenever $\int_0^\infty [c_1(\lambda) - c_2(\lambda)] f_q(\lambda) \, d\lambda > 0$; if this integral has the value 0, we choose $L(q, \lambda)$ to fulfill the assumption of right-continuity. For our example, we thus have

$$L(q, \lambda) = c_1(\lambda) \quad \text{if} \quad y(q) \leqq t(q);$$

$$L(q, \lambda) = c_2(\lambda) \quad \text{if} \quad y(q) > t(q).$$

For a Poisson process, we lose nothing in taking B and C closed and assuming $L_B(p) < \infty$. We have used our particular cost function $t(b) - t(p)$, but the result obviously holds true as long as the cost of experimentation is a non-decreasing linear function of the sample point, i.e., as long as the cost of a sample path going from p to c to b is equal to the cost of going from p to c plus the cost of going from c to b, both $\geqq 0$. These are fairly general conditions, but they do not cover the case where the cost of experimentation is $C + k(t - t_0)$ for $t > t_0$, and 0 for $t = t_0$, where C is a fixed constant, the so-called setup cost for the experiment. A note covering this case will be found near the end of Section 2.

In general, for a continuous-time stationary process with independent increments, if B is a boundary such that $L_B(p)$ is finite and C is a curve such that Pr (hit B before C) $= 0$, and B and C are such that γ, β, $h_p(c, \lambda)$, and $k_q(b, \lambda)$ exist; if the prior distribution is of the form of the likelihood function; if the cost of experimentation is linear in the sample point; if the coordinates of the sample point are sufficient statistics for λ; then the lemma is true. This is a long list of "if's", but not altogether unreasonable. For example, the lemma could be applied to the Wiener process with a normal prior distribution for μ, say with σ known, and with cost of observation $= k_1 t + k_2 y$ (where one of the k_i might be zero), and the cost of error proportional to $|\mu|^k$ for some $k > 0$. We pursue the matter no further here.

2.5. *Inequalities involving* $L_S(t, y)$. The purpose of this section is to prove the following statement needed for succeeding sections: There exists an $M > 0$ such that if max $(t, y) > M$, $(y/t)L_S(t, y + 1) < \frac{1}{2}$. As a preliminary result, it will be shown that $L_S(t, y) \leqq L_S(u, u)$ where $u = \max(t, y)$.

For $0 < t \leqq y$, $L_S(t, y) = \phi \int_0^1 (1 - \lambda)[e^{-\lambda t}\lambda^{y-1}t^y/\Gamma(y)]\, d\lambda$. Integrating by parts, we have $L_S(t, y) = \phi \int_0^1 [1 + t(1 - \lambda)][e^{-\lambda t}\lambda^y t^y/\Gamma(y + 1)]\, d\lambda$. Now $e^{-\lambda t}(\lambda t)^y$ and $[1 + t(1 - \lambda)]$ are each nondecreasing in t for $\lambda \leqq 1$, $t \leqq y$. Replacing t by y above, we see that for $0 < t \leqq y$, $L_S(t, y) \leqq L_S(y, y)$. For $0 < y \leqq t$, $L_S(t, y) = \phi \int_1^\infty (\lambda - 1)[e^{-\lambda t}\lambda^{y-1}t^y/\Gamma(y)]\, d\lambda$. Since $(\lambda t)^{y-1}/\Gamma(y)$ is monotone increasing in y for $y \leqq \lambda t$, and over the range of integration $y \leqq t \leqq \lambda t$, we have for $0 < y \leqq t$, $L_S(t, y) \leqq L_S(t, t)$. These establish the preliminary result.

Furthermore,

$$\begin{aligned} L_S(t, t) &= \phi \int_0^1 (1 - \lambda)[e^{-\lambda t}\lambda^{t-1}t^t/\Gamma(t)]\, d\lambda \\ &= \phi e^{-t}t^t/\Gamma(t + 1). \end{aligned}$$

Applying Stirling's formula to $\Gamma(t + 1)$, we see that $L_S(t, t)$ is a monotone decreasing function of t, approaching zero as t increases. Choose M so that $L_S(M, M) < (6\phi)^{-1}$; $M = 6\phi^4$ will do. Then for (t, y) such that $y < 3\phi t$ and $\max(t, y) > M$, $(y/t)L_S(t, y + 1) < (3\phi)(6\phi)^{-1} = \frac{1}{2}$. But for $y > 3\phi t$, $y + 1 > y > t$, and

$$(y/t)L_S(t, y + 1) = (\phi y/t) \int_0^1 (1 - \lambda)[e^{-\lambda t}\lambda^y t^{y+1}/\Gamma(y + 1)] \, d\lambda$$

$$< [\phi/\Gamma(y)]t^y \int_0^1 \lambda^y \, d\lambda < [\phi/\Gamma(y + 1)](y/3\phi)^y < (2\pi y)^{-\frac{1}{2}}(e/3)^y < \tfrac{1}{2}.$$

We have shown:

$$(y/t)L_S(t, y + 1) < \tfrac{1}{2} \quad \text{if} \quad \max(t, y) > M = 6\phi^4. \tag{2.18}$$

2.6. *A differential equation for $L_B(t, y)$.* Let u' denote the derivative of u with respect to t. Then since

$$L_B(t, y) = \int_0^\infty L_B(t, y; \lambda) f_{t,y}(\lambda) \, d\lambda,$$

$$L_B'(t, y) = \int_0^\infty [L_B'(t, y; \lambda) f_{t,y}(\lambda) + L_B(t, y; \lambda) f_{t,y}'(\lambda)] \, d\lambda.$$

But $L_B(t - h, y; \lambda) = h + (1 - \lambda h)L_B(t, y; \lambda) + \lambda h L_B(t, y + 1; \lambda) + o(h)$, whenever $(t - h, y) \in B_{\text{int}}$ for all sufficiently small positive h, so that

$$L_B'(t, y; \lambda) = -1 - \lambda[L_B(t, y + 1; \lambda) - L_B(t, y; \lambda)];$$

and $f_{t,y}'(\lambda) = [(y/t) - \lambda]f_{t,y}(\lambda)$.
Therefore $L_B'(t, y) = -1 - (y/t)[L_B(t, y + 1) - L_B(t, y)]$, which we will write as

$$(\partial/\partial t)L_B(t, y) = -1 - (y/t)\Delta_y L_B(t, y). \tag{2.19}$$

This equation of course refers to derivatives from the left, and holds for $(t, y) \in B_{\text{int}}$ and for (t, y) the right hand end point of an interval of B_{int}. The same equation can be shown to hold for derivatives from the right, *except* where $(t, y + 1) \in B$ and $(t, y + 1)$ is the left end point of an interval of B_{int}. For these points, $L_B(t, y + 1)$ must be replaced by $\lim_{h \to 0} L_B(t + h, y + 1)$.

2.7. *Boundedness of the set of "go" points.* Again, let u' denote $\partial u/\partial t$. Divide the points q into two sets: those q from which there exists a plan B of the type under consideration with $L_B(q) < L_S(q)$, and those points from which no such plan exists. They will be referred to as "go" points and "stop" points, respectively. Let G denote the set of go points. Then the main result of this section is a proof that G is contained in the square given by $t(q) \leqq M$, $y(q) \leqq M$ (which we call the M-square), for some M.

From Section 2.5 there exists a value M such that $(y/t)L_S(t, y + 1) < \frac{1}{2}$ outside the M-square. Now $L_S(t, y)$ is continuous in t, and has one-sided derivatives at every point; they are equal except at $t = y$. Straightforward calculation shows that, using derivatives from the left and denoting $F(t, y + 1) - F(t, y)$ by $\Delta_y F(t, y)$ or simply $\Delta_y F$, $L_S'(t, y) = -(y/t)\Delta_y L_S$ for $t \leqq y$ and for $t > y + 1$, while for $y < t < y + 1$, $L_S'(t, y) = -(y/t)\Delta_y L_S + (\phi y/t)[1 - (y + 1)/t] > -(y/t)\Delta_y L_S - \phi y/t^2 > -(y/t)\Delta_y L_S - (\phi/t) > -(y/t)\Delta_y L_S - (\phi/y)$. Take $M > 2\phi$; this gives $L_S'(t, y) > -(y/t)\Delta_y L_S - \frac{1}{2} > -1 + (y/t)L_S(t, y)$ outside the M-square. We will use this result to prove that G is contained in the M-square.

Suppose there is a go-point q outside the M-square. Then there exists a plan B with $L_B(q) < L_S(q)$. Then either (Case A) there exist one or more points b of B to the right of q (with $y(b) = y(q)$), or (Case B) there does not exist a point of B to the right of q. We treat the cases in turn.

In Case A, let b be the first of these points of B. (Such a first point exists since B is closed.) $L_B(q) < L_S(q)$, and $L_B(b) = L_S(b)$; and both L_B and L_S are continuous on the closed interval from q to b. But then there exists at least one point q_1 with $t(q) < t(q_1) \leqq t(b)$, $y(q_1) = y(b)$, $L_B(q_1) = L_S(q_1)$, and $L_B'(q_1) \geqq L_S'(q_1)$, using derivatives from the left. We have just shown that $L_S'(q) > -1 + [y(q)/t(q)]L_S(q)$. It was shown in Section 2.6 that $L_B'(q) \leqq -1 + [y(q)/t(q)]L_B(q)$. At q_1, $L_B(q) = L_S(q)$; so $L_B'(q_1) < L_S'(q_1)$, a contradiction. This proves Case A impossible; we turn to Case B.

Either (Case B_1) there exists a point q_1 to the right of q with $L_S(q_1) \leqq L_B(q_1)$, or (Case B_2)$L_B(q_1) < L_S(q_1)$ for all q_1 to the right of q. In Case B_1, add the point q_1 to B. This of course will not increase $L_B(q)$, but will reduce the situation to Case A which has already been disposed of. In Case B_2, pick a point q_1 with $t(q_1) > y(q_1) = y(q)$; write t for $t(q_1)$, and y for $y(q_1)$. By assumption, $L_B(q_1) < L_S(q_1)$; by Section 2.5, $L_S(q_1) = L_S(t, y) < t/(y - 1)$. Now $L_B(q_1) \geqq$ expected cost of observation, so that

$$L_B(q_1) \geqq \int_0^\infty \left\{\int_0^\infty ue^{-\lambda u}\lambda\, du\right\} e^{-\lambda t}\frac{\lambda^{y-1}t^y}{\Gamma(y)}\, d\lambda$$
$$= \int_0^\infty e^{-\lambda t}\frac{\lambda^{y-2}t^y}{\Gamma(y)}\, d\lambda = \frac{t}{y-1} > L_S(q_1).$$

This contradiction completes the proof.

2.8. *Characterization of the optimum boundary.* Since $L_B(q)$ is continuous in $t(q)$ for q in B_{int}, we see that G, the set of go points, is open (on each line $y = n$), and therefore $G \cap B_{\text{int}}$ is likewise open for any B, hence Borel. We remark that if a plan B has $B_{\text{int}} - G$ nonempty, then defining $B' = B \cup (B_{\text{int}} - G)$, we see by the lemma with $C = B'$ that for any point q, $L_{B'}(q) \leqq L_B(q)$. Therefore we need only consider plans wherein escape from G means immediate stoppage, i.e., plans such that any point not in G may be considered a point of B.

We now restrict ourselves to a given y_0; and for convenience we take y_0 an integer. This will give all points on B having integral y-values; to get points in between, we would need to take $y_0 = n + h$ for values of h in the interval $(0, 1)$, instead of $y_0 = n$, and repeat the procedure now to be outlined.

For any p not in G, there is an obvious optimum plan: take an immediate decision. Let $y = n$ be the maximum integral value of y taken by points of G. Consider a point (t_0, n) in G. As pointed out above, we may assume that any point not in G is in B, and we ask whether there is an optimum among such plans. But $G \cap \{q \mid y(q) = n + 1\}$ is empty. Therefore, any such plan is determined, as far as the point (t_0, n) is concerned, by the infimum t_1 of t-values $> t_0$ assumed on $B \cap \{q:y = n\}$. But the point (t_1, n) is in B. Since (t_1, n) may be restricted to the set of points in G (and the end point thereof), it follows that the continuous function $L_B(t_0, n)$ of t_1 has a minimum for some t_1. It is easily seen (by the lemma) that t_1 cannot be an interior point of G, since then we could get a plan with loss less than the minimum. Therefore, (t_1, n) is the first point to the right of t_0 which is *not* in G, i.e., the first stop point to the right of t_0. Thus for any (t_0, n) in G, the optimum plan is to continue until a stop point is reached.

We now know the optimum for any point, on the line $y = n$; and for any point not in G, on the line $y = n - 1$. Then for q in G and $y(q) = n - 1$, by similar reasoning, we need only consider plans where if an event occurs, i.e., if the line $y = n$ is reached, observation continues until a stop point is reached. But these plans are again determined by the first stop point to the right of q, and again its t-coordinate is in the closed interval $t \geqq t(q)$, $(t, n - 1)$ in $\bar{G}$; and $L_B(q)$ is continuous in this t-coordinate, so the optimum from any point on the line $y = n - 1$ exists, and consists of observing until a stop point is reached. Similarly, we can work all the way back to the line $y = 1$, giving this main result.

THEOREM.

(1) *Assumptions.* (a) *The cost of observation is proportional to the length of time the process is observed.* (b) *The cost of a wrong decision is proportional to the value* $|\lambda - c|$ *for some known* c, *where the decision is either* "$\lambda < c$" *or* "$\lambda > c$". (c) *The prior distribution is given as*

$$e^{-\lambda t_0}(\lambda^{y_0-1}t_0^{y_0}/\Gamma(y_0)), \qquad y_0 \geqq 1,\ t_0 > 0.$$

(2) *Restrictions on the decision rule. Consider all decision rules corresponding to a fixed boundary in the* (t, y)*-plane, that is, all decision rules which are given as: "Let* $y - y_0$ *be the number of events that have occurred during the time* $t - t_0$. *Observe till* (t, y) *first reaches the set* B, *and then stop and make a decision (which is given in terms of the values* t, y, t_0, *and* y_0*)", subject only to these two restrictions:* (a) *The expected loss* $L_B(t_0, y_0)$ *is defined.* (b) *The decision cost,* $L(t, y, \lambda)$, *is continuous* (t) *from the right.*

(3) *Conclusions. Then replacing the variable* t *by* t'/c, *we can, for a given ratio of cost constants, find one boundary in the* (t', y)*-plane which will give minimum expected loss for any prior distribution of the type assumed; we simply start the sample path at* (ct_0, y_0). *This boundary set* B, *say, includes all points outside some square* $y \leqq M$, $t' \leqq M$, *and is a closed set in the plane.*

As a matter of fact, the boundary could be calculated analytically (in theory only, since the equations are hopeless); but a more satisfactory method seems to be the use of an electronic computer.

A setup cost S can now be taken care of, as follows: We determine an optimum boundary as above, ignoring the setup cost; and on it, we mark contours of $L_S(q) - L_B(q)$. The contour $L_S(q) - L_B(q) = S$ then becomes an auxiliary curve; the given plan is optimum for any starting point p having $L_S(p) - L_B(p) > S$, while an immediate decision is optimum for all other points.

If one wishes to avoid a prior distribution, one could try to make the posterior distribution equal to the (normalized) likelihood function of the observations, on the grounds that the experiment is all that counts. This corresponds to starting observation at (0, 1). However, it will be shown in the next paragraph that there is an interval of nonzero length, with (0, 1) as left end point, which is not in G; therefore, from (0, 1), the optimum plan is to stop immediately. This result is not surprising, viewed from the "prior distribution" point of view; but it does indicate that the attitude "Let the data speak for themselves" deserves the response, "How?"

Above the main diagonal,

$$t < y, L_S' = -(y/t)\Delta_y L_S, \quad \text{and} \quad L_B' = -1 - (y/t)\Delta_y L_B .$$

Within one unit (vertically) of the upper boundary, i.e., the boundary above the diagonal,

$$L_B' = -1 - (y/t)[L_S(t, y + 1) - L_B(t, y)].$$

Let $D(t, y) = L_B(t, y) - L_S(t, y)$; then $D' = -1 + (y/t)D(t, y)$ near the upper boundary. The solution of this equation is $D = t(ct^{y-1} + 1)(y - 1)^{-1}$, where c is a function of y alone. Hence $D(0, y) = 0$ and $D'(0, y) = (y - 1)^{-1} > 0$ (for $y > 1$). Therefore, since G is bounded, the upper boundary steps down (as t decreases) at *positive* values of t. Therefore, there exists an interval from $t = 0$ to some $t_0 > 0$ wherein $L_B(t, 2) = L_S(t, 2)$; but then for $y = 1$, $D'(t, 1) = -1 + t^{-1}D(t, 1)$; or $D(t, 1) = t \log t - ct, c > 0; D(0, 1) = 0; D'(0, 1) = +\infty$. This proves that G does not extend all the way to the point (0, 1).

Finally it bears mentioning that this starting point is (0, 1) and not (0, 0) simply because $f_{t,y}(\lambda)$ was taken as $e^{-\lambda t}\lambda^{y-1}t^y/\Gamma(y)$ and not $e^{-\lambda t}\lambda^y t^{y+1}/\Gamma(y + 1)$. This was done so that the acceptance set would be the "triangle" $t \geqq y$, instead of $t \geqq y + 1$.

3. Calculation of the optimum boundary B.

3.1. *General appearance.* If the point $q = (t, y)$ is in G, then there must exist a path from (t, y) to the "other side" of the diagonal $y = t$, and in fact there must be a positive probability of reaching or crossing the line. (If not, then being at q, we have $L_B(q) < L_S(q)$; but the probability of changing the decision by going on is zero, while the cost of experimentation is positive, so by (1.1) and (1.2) we see that $L_B(q) > L_S(q)$.) This means, loosely speaking, that G has no "bulges" to the right, below the diagonal, and no bulges upward above the diagonal. Thus if $y = n$ contains part of G but $y = n + 1$ does not, then if $(n - u_0, n)$ is in G, $(n - u, n)$ is in G for all u in $(0, u_0)$. Also, there does not exist a point (t, n) in G with $t \geqq n + 1$, by the previous section. Therefore, there is a point p_0 in $\bar{G}$, $p_0 = (t_0, n)$, $n \leqq t_0 < n + 1$, with no points of G to its right and an interval of G to its immediate left. Let (t_1, n) in $\bar{G} - G$ represent the left end point of the largest such interval of G. If $t_1 < n$, then there are no points (t, n) of G with $t < t_1$. In Sections 2.6 and 2.7, it was shown that for $y < t \leqq y + 1$,

$$L_S' = -(y/t)[L_S(t, y + 1) - L_S(t, y)] + (y\phi/t)[1 - (y + 1)/t];$$

otherwise, $L_S' = -(y/t)[L_S(t, y + 1) - L_S(t, y)]$; while for (t, y) any interior point, $L_B' = -1 - (y/t)[L_B(t, y + 1) - L_B(t, y)]$. Now if (t, y) is in B and is the right hand end point of an interval of B_{int}, this formula becomes $L_B' = -1 - (y/t)[L_B(t, y + 1) - L_S(t, y)]$. With these results we can show the following:

(a) If (t, n) is a point of B and a right hand end point of an interval of B_{int}, then $L_B'(t, n) = L_S'(t, n)$. *Proof:* If $L_B' < L_S'$, then there are points in B_{int} which should be in B; therefore B is not optimum, a contradiction. If $L_B' > L_S'$, then because the difference is a continuous function of t, there is a $u_0 > 0$ such that the inequality holds at $(t + u, n)$ for $0 \leqq u \leqq u_0$ if $(t, n + 1)$ is in B_{int}, while if $(t, n + 1)$ is in B, we note that $L_B(t + u, n + 1) \leqq L_S(t + u, n + 1)$ ensures the existence of such a u_0. If we then let $(t + u, n)$ be points of B_{int} and $(t + u_0, n)$ be a point of B, we see that this altered B has $L_B(t, n) < L_S(t, n)$, and therefore (t, n) should not have been in B. This contradiction completes the proof.

(b) If (t, n) is a point of B (i.e., a stop point) and $(t + u, n)$ is in B_{int} for $0 < u \leqq u_0$, for some u_0, then $L \equiv \lim_{u \to 0} L_B(t + u, n) = L_S(t, n)$. *Proof:* If $L > L_S(t, n)$, then by continuity of L_S, there exist points $(t + u, n)$ in B_{int} for which $L_S(t + u, n) < L_B(t + u, n)$, a contradiction. If $L < L_S(t, n)$, then (treating (t, n) as an interior point) it can easily be shown that $L_B(t, n) < L_S(t, n)$, and so (t, n) should not be in B. This completes the proof.

(c) By (a), there exist go points on every line $y = n \leqq \phi$, and no others; there exist no go points with t-coordinate $> \phi + 1$. *Proof:* The points of G having maximum y-coordinate have as right end point the point $q = (t, y)$ such that $L_S' = -(y/t)[L_S(t, y + 1) - L_S(t, y)] + (y\phi/t)[1 - (y + 1)/t] = -(y/t)[L_S(t, y + 1) - L_S(t, y)] - 1 = L_B'$, or $y\phi(t - y - 1) + t^2 = 0$; and $y \leqq t < y + 1$. The quadratic formula gives

$$2t = -\phi y + [\phi^2 y^2 + 4\phi y(y + 1)]^{\frac{1}{2}} < 2(y + 1), \tag{3.1}$$

so that any solution t_0 will be less than $y + 1$; it remains to discover under what conditions there will be a solution $t_0 \geqq y$. Rewriting (3.1) as $2t = -\phi y + [(\phi y)^2 + 4(\phi y)y + 4y^2(\phi/y)]^{\frac{1}{2}}$, we discover that $t_0 \geqq y$ if and only if $\phi \geqq y$. We thus state: Let y_0 be the largest integer value of y appearing in the set G;

and let $t_0 = \sup\{t:(t, y_0) \in G\}$. Then

$$(3.2)\qquad \begin{aligned} y_0 &= [\phi], \text{ the greatest integer} \leqq \phi; \\ 2t_0 &= -\phi y_0 + [\phi^2 y_0^2 + 4\phi y_0(y_0 + 1)]^{\frac{1}{2}}. \end{aligned}$$

We have derived a formula for the upper right limiting point of G.

Let (t_1, y_0) represent the first stop point to the left of (t_0, y_0). Then by (b),

$$\begin{aligned} L_S(t_1, y_0) &= L_B(t_1, y_0) \\ &= \int_0^\infty \left\{ \int_0^{t_0 - t_1} [L_S(t + u, y + 1) + u] e^{-\lambda u} \lambda \, du \right\} e^{-\lambda t_1} \frac{\lambda^{y_0 - 1} t_1^{y_0}}{\Gamma(y_0)} \, d\lambda. \end{aligned}$$

After the algebra, we have

$$(3.3)\qquad (t_1/t_0)^{y_0 - 1} = \{1 + [\phi(y_0 - 1)/t_0^2](t_0 - y_0)\}^{-1}.$$

Let $(t_2, y_0 - 1)$ be the right limit of G on the line $y_0 - 1$. Then the same sort of method gives

$$(3.4)\qquad (t_2/t_0)^{y_0+1} = [\phi(y_0 - 1)(t_2 - y_0) + 2t_2^2]/[\phi(y_0 - 1)(t_0 - y_0) - t_0^2].$$

These are the first three of the increasingly complicated formulas which specify the boundary. The first two are useful to determine where the boundary first becomes accessible, i.e., to determine the maximum value of y that could be reached from (say) the point (1, 1). For $y_0 = \phi$, the boundary overlaps; that is, $t_1 > t_2$, which means that the go points on the line y_0 cannot be reached from a point (t, y) with $t < t_2$, because the stop points on the line $y = y_0$ and the point $(t_2, y_0 - 1)$ form a closed corner. This will happen as long as $t_1(y) > t_2(y)$. But when this happens, then $t_2(y_0) = t_0(y_0 - 1)$, and we calculate the boundary on and below the line $y_0 - 1$ by ignoring the points of G on the line y_0.

The largest value of y for which this does not happen is given as $[y']$ where y' is the solution of the equation

$$(3.5)\qquad \left[\frac{(\phi + 1)y - 1}{(\phi + 1)y + \phi}\right]^{y-1} = \left[1 + \frac{\phi(y - 1)\dfrac{\phi - y}{\phi + 2}}{\left(y + \dfrac{\phi - y}{\phi + 2}\right)^2}\right]^{-1}$$

which expresses approximately the statement $t_0(y - 1) = t_1(y)$, using two terms of the approximation

$$(3.6)\qquad t_0(y) = y + \frac{\phi - y}{\phi + 2} - \frac{(\phi - y)^2}{y(\phi + 2)^3} + - \cdots$$

where terms go up like $(\phi - y)^k/[k!\, y^{k-1}(\phi + 2)^{2k-1}]$. We invert each side of (3.5) and further approximate, to obtain

$$(3.7)\qquad \exp\left(\frac{y - 1}{y + 1}\right) = 1 + \frac{(y - 1)(\phi - y)}{(y + 1)^2}.$$

Examples of the limit of accessible boundary are:

ϕ	$[y']$ (approx.)
15	5
45	17
95	35.

3.2. *An example.* A program was written for the Datatron 205 computer, and used to produce the example shown in Figure 2, for $\phi = 50$, for which $y' = 18$.

When $17 \leqq y < 50$, the optimum plan is to sample for a fixed time or until an event occurs, whichever is shorter. The *right* hand (lower) boundary points follow essentially a straight line from $t = 17.634$, $y = 17$ to $t = 50$, $y = 50$, while the left hand (upper) points have a slight curvature, from (16.599,17) thru (33.937,34) to (50,50); see Table 1.

TABLE I

Horizontal Distance of Boundary Points from the Diagonal. ($\times 10$)

y	Upper Boundary	Lower Boundary
50	0	0
49	0.002	0.192
48	0.008	0.384
47	0.018	0.577
46	0.033	0.769
45	0.052	0.961
44	0.075	1.154
43	0.104	1.346
42	0.138	1.538
41	0.177	1.730
40	0.222	1.923
39	0.274	2.115
38	0.331	2.307
37	0.395	2.500
36	0.466	2.692
35	0.545	2.884
34	0.631	3.076
33	0.726	3.269
32	0.829	3.461
31	0.942	3.653
30	1.065	3.845
29	1.198	4.037
28	1.343	4.230
27	1.500	4.422
26	1.670	4.614
25	1.854	4.806
24	2.054	4.998
23	2.269	5.190
22	2.503	5.382
21	2.757	5.574
20	3.031	5.766
19	3.329	5.958
18	3.653	6.150
17	4.008	6.342

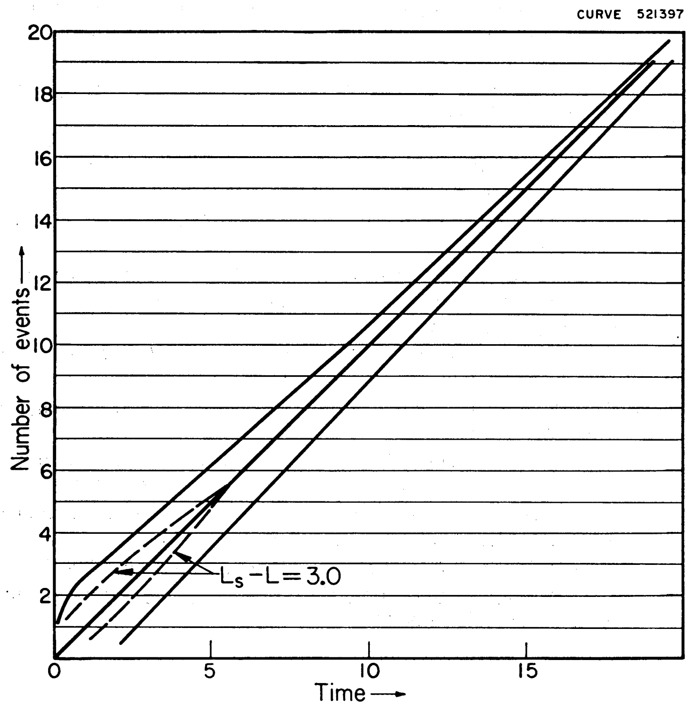

Fig. 2. The optimum boundary for $\phi = 50$

Below $y = 17$, both boundaries look almost straight for a while, slowly curving away from the diagonal; then each curves more sharply back toward the origin.

The inside (dashed) curve on Figure 2 is the locus of (t, y) such that $L_S - L = 3$. This then is an auxiliary curve when there is a setup cost equal to 3; for (t_0, y_0) outside this curve, the expected saving from using the optimum plan instead of making an immediate decision is less than the setup cost, so that no experimentation should be done, while for (t_0, y_0) inside this curve, the optimum plan should be used.

An illustration of the saving due to experimentation is the following:

$$L_S(1, 1) = 18.4; L(1, 1) = 4.17.$$

4. Acknowledgments. The author wishes to express his sincere gratitude to Prof. F. J. Anscombe, under whose guidance this work was done. Also, thanks are due the referee, whose comments were of considerable help in shortening the paper and in simplifying some of the derivations. In particular, the proof used herein of the lemma was proposed by him.

[1] Breakwell, John V. (1956). Economically optimum acceptance tests. *J. Amer. Statist. Assoc.* **51** 243–256.

[2] Chernoff, Herman (1959). Asymptotically optimal stopping rules in sequential analysis (preliminary report). *Ann. Math. Statist.* **30** 1273 (Abstract).

[3] CHERNOFF, HERMAN (August 26, 1960). Sequential Tests for the Mean of a Normal Distribution. Technical Report No. 59, Applied Math. and Stat. Labs., Stanford Univ.

[4] DE GROOT, MORRIS H. (1960). Minimax sequential tests of some composite hypotheses. *Ann. Math. Statist.* **31** 1193–1200.

[5] GRUNDY, P. M., HEALY, M. J. R., and REES, D. H. (1956). Economic choice of the amount of experimentation. *J. Roy. Statist. Soc. Ser. B.* **18** 32–55.

[6] DVORETZKY, A., KIEFER, J., and WOLFOWITZ, J. (1953). Sequential decision problems for processes with continuous time parameters. Testing hypotheses. *Ann. Math. Statist.* **24** 254–264.

[7] KIEFER, J., and WOLFOWITZ, J. (1956). Sequential tests of hypotheses about the mean occurrence time of a continuous parameter Poisson process. *Naval Research Logistics Quart.* **3** 205–219.

[8] MUNROE, M. E. (1953). *Introduction to Measure and Integration.* Addison-Wesley, Cambridge.

[9] SCHWARZ, G. (1962). Asymptotic shapes of Bayes sequential testing regions. *Ann. Math. Statist.* **33** 224–236.

[10] SOBEL, MILTON (1953). An essentially complete class of decision functions for certain standard sequential problems. *Ann. Math. Statist.* **24** 319–337.

[11] WALD, ABRAHAM (1939). Contributions to the theory of statistical estimation and testing hypotheses. *Ann. Math. Statist.* **10** 299–326.

[12] WALD, ABRAHAM (1950). *Statistical Decision Functions.* Wiley, New York.

[13] YATES, F. (1952). Principles governing the amount of experimentation in developmental work. *Nature* **170** 138–140.