On Sequential Detection of a Shift in the Probability of a Rare Event

RON KENETT and MOSHE POLLAK*

Suppose one is monitoring a sequence of observations for a possible increase in the probability of a rare event, and that it is not possible to immediately stop the process under observation or influence it to return to its normal state. One would then desire a scheme which takes advantage of observations occurring after a detection of a change is proclaimed. A modification of Page's CUSUM Procedure is developed, taking account of these additional observations. A table is given enabling one to select a modified Page procedure with a specified rate of false alarms. A comparison is made between the modified Page procedure and a procedure proposed by Chen (1978).

KEY WORDS: Page procedure; Optimal stopping; Cumulative sum; Surveillance.

1. INTRODUCTION AND SUMMARY

The occurrence of rare, often undesired events, is common to many situations. At any given time a computer system may break down, a baby may be born with a congenital malformation, an industrial process may produce a defective item. Under normal circumstances, such events occur at a constant low frequency. Should this frequency change, it would be of interest to ascertain as quickly as possible that a change has taken place.

The problem we address in this article is how to detect such a change when it is not possible to immediately stop the process under observation or to influence it to return to its normal state. Under such circumstances the behavior of the process will not change after the shift is detected. One example in which such considerations arose is the thalidomide episode in the 1960's. Another example would be a situation where there is an increase in the breakdown rate of a computer system which may, for instance, be due to aging hardware, a lapse in maintenance, or a change in the machine load. Identification of the source of a problem may require a specific investigation. Monitoring continues while such an investigation is carried out. See Morton and Lindsten (1976), Weatherall and Haskey (1976) and Hook (1978) among others for more examples.

Typically the characteristics of interest in a surveillance scheme are the time elapsed between an increase and its detection, and the rate of false alarms. While most schemes are determined by the expected time until a (first) false alarm, we follow Chen (1978) in also taking into account succeeding false alarms.

Page (1954) introduced a procedure for detecting changes in frequencies of failures. In his set-up the mechanism generating the failures can be stopped and the failure frequency can be reset to its nominal level (by overhauling a machine for example). The main tool of the Page procedure is a cumulative sum (CUSUM) of the times between failures. When the CUSUM climbs above a prespecified detection level the Page procedure sounds an alarm and recycles by deleting all past observations.

Briefly, the monitoring scheme we propose here is a modified Page procedure: A Page (1954) procedure is employed until the first alarm. Subsequently, another alarm is sounded whenever a failure occurs such that the appropriate CUSUM is above the detection line crossed at the first alarm. The rate of such alarms which are false will determine the characteristics of the procedure. Should the CUSUM fall below a "recycling" level, the Page procedure is started anew.

In Section 2 the Page procedure and the modified Page procedure are presented in detail and we derive a table which enables the user to select a modified Page procedure with a specified rate of false alarms. Our method for deriving this table relies on transforming binomial observations with small p to exponential observations (Anscombe and Page 1954), using results of Lorden and Eisenberger (1973) for the Page procedure on exponential variables and employing renewal theory (Feller 1971) for what happens after the first alarm.

In Section 3 we compare the procedure proposed by Chen (1978) with the modified Page procedure. We find that in terms of the expected time elapsed between a change and its detection, the specific procedures proposed in Chen (1978) take 4–32 percent more time to detect a true change, given equal false alarm rates.

2. THE PAGE AND THE MODIFIED PAGE PROCEDURES

The process we observe consists of a sequence of Bernoulli random variables with a probability of failure π_0.

* Ron Kenett is Reliability Engineering Manager, Tadiran, Telecommunications division, Petah Tikva, Israel. Moshe Pollak is Senior Lecturer at the Department of Statistics, The Hebrew University of Jerusalem, Jerusalem, Israel. This work was initiated while both authors were at the Department of Statistics, University of Wisconsin, Madison, WI, and was completed while the first author was in Bell Laboratories, Piscataway, NJ. The authors would like to thank Alan M. Gross for statistical and editorial advice and Steve Brzezinski for writing and thoroughly testing the computer programs used in this study.

June 1983, Volume 78, Number 382
Theory and Methods Section

(A "failure" may be a child born with a congenital malformation or a computer system breaking down, for example.) The problem is to determine on the basis of the observations whether π_0 has increased at an unknown past point in time to a given value $\pi_1 = \gamma\pi_0(\gamma > 1)$. A good procedure detects the change as soon as possible after its occurrence, subject to a controlled false alarm rate.

We assume the observations are made at discrete times (e.g., births, or 15 minute intervals). Let B_i be the number of Bernoulli variables observed between the $(i - 1)$th failure and the ith failure (including the latter).

Under the null hypothesis (H_0) of no increase, the distribution of B_i is Geometric(π_0). Set $X_i = (B_i - 1)\log((1 - \pi_0)/(1 - \pi_1))$. Asymptotically, as $\pi_0 \to 0$, X_i is Exponential(θ_0), where $\theta_0 = 1/(\gamma - 1)$. (See Anscombe and Page 1954). Under the alternate hypothesis (H_1) of an increase from π_0 to $\pi_1 = \gamma\pi_0$, B_i is Geometric(π_1) for observations following the first failure after the increase and, asymptotically as $\pi_0 \to 0$, X_i is Exponential(θ_1), where

$$\theta_1 = \frac{\gamma}{\gamma - 1} = \gamma\theta_0.$$

Note that $\theta_1 - \theta_0 = 1$.

In mathematical terms the problem is to find, subject to certain conditions, a statistical procedure for detecting changes in a failure probability at an unknown point in time. This requires the identification of a stopping time N for an observed sequence of random variables $X_1, X_2, \ldots, X_N$. (N is an integer-valued random variable which depends on present and past observations only.)

2.1 The Page Procedure

A general procedure for detecting an increase in failure rates was proposed by Page (1954). His procedure consists of repeated applications of a sequential probability ratio test (SPRT) which in the present context is definable by the inequalities

$$0 < L_n = n \log \gamma - S_n < a \qquad (1)$$

where $S_n = X_1 + X_2 + \cdots + X_n$, $a > 0$. In the Page procedure the applications stop as soon as the inequality on the right side is violated, with the proviso that if the left inequality is violated first, all observations up to that point are discarded and the procedure is recycled with $S_1, S_2, \ldots$ denoting cumulative sums of the new observations. (See Page 1961 for an overview of the Page procedure and references cited therein for further analyses of its properties.)

The characteristics of a Page procedure are determined by α, the probability under H_0 that for some $n \geq 1$, $a > L_i > 0$ for $i = 1, \ldots, n - 1$ and $L_n \geq a$, and by $1 - \beta$, the probability of the same event under the special case of H_1 in which the increase is already in effect at the start of the surveillance. For given values of α, β, and γ, one can evaluate the detection level a in (1). (See Lorden and Eisenberger 1973). The application of a Page procedure to the monitoring of failure rates typically stipulates that after $L_n \geq a$, all observations are discarded and the procedure is started anew. (See Morton and Lindsten 1976 for such an application of the Page procedure to the surveillance of Down's Syndrome.) The implication of recycling the procedure, once H_0 is rejected in favor of H_1, is that if H_1 is true the recycled procedure will not redetect it until the detection level a is crossed again.

2.2 The Modified Page Procedure

The procedure we advocate here for monitoring failure rates is identical to the Page procedure until the right inequality in (1) is violated for the first time, resulting in an alarm. The proposed modification is that no recycling is carried out at that point and that following the first alarm further alarms are declared whenever the right inequality in (1) holds. Such action should be continued until (if ever) the next time that the left inequality in (1) is violated, in which case one recycles and starts the procedure anew. The implication of this modified Page procedure is that these subsequent alarms sounded after the first alarm prior to any recycling are further evidence for the rejection of H_0 in favor of H_1. The rate of alarms, including these, which are false determines the detection level α. A similar concept is used in the procedure proposed by Chen (1978).

Formally, let $Y_i = \log \gamma - X_i$. Let $L_n = \sum_{i=1}^{n} Y_i$ for all n until the first n for which the following case (ii) applies; subsequently define

$$Y_{i,n} = \begin{cases} Y_i & i > \text{index of last recycling (before } n) \\ 0 & \text{otherwise,} \end{cases}$$

$$L_n = \sum_{i=1}^{n} Y_{i,n} \quad \text{(where recycling is defined in case (ii))}$$

The modified Page procedure is implemented by tracking the path of L_n as the number of failures, n, increases. Three cases are to be considered:

(i) $0 < L_n < a$, no action is taken,
(ii) $L_n \leq 0$, the procedure is recycled by deleting all observations up to n (inclusive),
(iii) $L_n \geq a$, an alarm is declared.

This is in contrast to the regular Page procedure which is recycled after every n such that $L_n \geq a$. The rate Λ of false alarms when applying the modified page procedure is given by

$$\Lambda = \frac{E_{H_0}(\text{number of } n\text{'s up to } M_2 \text{ such that } L_n \geq a)}{E_{H_0}(M_2)}, \qquad (3)$$

where M_2 = first n such that $L_n \leq 0$ after $L_m \geq a$ for some $m < n$. In Table 1 we present values of Λ for different detection levels a and values of γ. For specified

Table 1. Values of 10^5 False Alarm Rate = $10^5\ \Lambda$

a	γ: 1.1	1.3	1.5	1.75	2	3	5	7	10	100
.5	58756	55577	52988	50379	48175	42131	35812	32280	28996	15848
1.0	35638	33709	32143	30542	29220	25605	21714	19550	17554	9608
1.5	21615	20445	19495	18524	17724	15519	13186	11870	10650	5826
2.0	13110	12401	11825	11235	10750	9414	7993	7200	6464	3534
2.5	7952	7521	7172	6815	6520	5710	4848	4366	3920	2143
3.0	4823	4562	4350	4133	3955	3463	2941	2648	2377	1300
3.5	2925	2767	2638	2507	2399	2101	1784	1606	1442	788
4.0	1774	1678	1600	1521	1455	1274	1082	974	875	478
4.5	1076	1018	971	922	882	773	656	591	530	290
5.0	653	617	589	559	535	469	398	358	322	176
5.5	396	374	357	339	325	284	241	217	195	107
6.0	240	227	217	206	197	172	146	132	118	65

NOTE: $\pi_0 = P_{H_0}$ (failure); a = detection level for sounding an alarm; γ = multiple of π_0 being monitored.

values of Λ, π_0 and π_1 one can determine the detection level a by interpolation. Computations on which Table 1 is based are derived in the Appendix. Since the modified Page procedure is a regular Page procedure until the first n for which $L_n \geq a$, the values of α and β of both procedures are identical (see Lorden and Eisenberger 1973 for a derivation).

3. COMPARISON OF CHEN'S PROCEDURE WITH THE MODIFIED PAGE PROCEDURE

The surveillance system proposed by Chen (1978) is based on the number of consecutive Bernoulli observations between failures, $O_i = B_i - 1$. Such a group of consecutive observations is defined as a set. An alarm is signaled if a sequence of m sets appears such that each set is below a certain size. Let

$$\mu_0 = E_{H_0}(O_i) = \frac{1-\pi_0}{\pi_0}, \quad \pi_1 = \gamma\pi_0.$$

Chen shows that under the special case of H_1 where the change was already in effect at the time surveillance started

$$P_{H_1}(O_{\max} \leq K\,\mu_0) = (1 - e^{-K\gamma})^m$$

where $K > 0$ and $O_{\max}$ is the size of the maximal set in the sequence.

We base our comparison of Chen's procedure with the modified Page procedure on the expected number of observations it takes to detect a true increase in failure rate after matching the rate of false alarms.

We first derive under the previously mentioned alternative more accurate results for Chen's procedure. Probabilities and expectations for this case will be denoted by P_{H_1} and E_{H_1}, respectively. From Wald's lemma, we have

$$A_C = E_{H_1}\left(\sum_{i=1}^{N} B_i\right) = E_{H_1}(N)\,E_{H_1}(B_i) = E_{H_1}(N)\,\frac{1}{\pi_1}, \tag{4}$$

where $\sum_{i=1}^{N} B_i$ is the number of observations until a first alarm is signaled with Chen's procedure, A_C its expected value, and N is a random variable counting the number of sets observed in that period of time.

To compute $E_{H_1}(N)$ let

$$g = 1 - e^{-K\gamma}. \tag{5}$$

Then

$$P_{H_1}(N < m) = 0,$$

$$P_{H_1}(N = m) = g^m,$$

$$P_{H_1}(N = i) = g^m(1-g)P_{H_1}(N \geq i-m) \quad \text{for } i > m.$$

Hence

$$P_{H_1}(N \geq j) = \frac{P_{H_1}(N = j+m)}{g^m(1-g)} \quad \text{for } j \geq 1.$$

Note that $P_{H_1}(N = j + m) = g^m(1 - g)$ for $j = 1, \ldots, m$.

Now

$$\begin{aligned} E_{H_1}(N) &= \sum_{j=1}^{\infty} P(N \geq j) \\ &= m + \sum_{j=m+1}^{\infty} \frac{P_{H_1}(N = j+m)}{g^m(1-g)} \\ &= m + \frac{1-(g^m + mg^m(1-g))}{g^m(1-g)} \\ &= \frac{1-g^m}{g^m(1-g)}. \end{aligned} \tag{6}$$

Hence

$$A_C = \frac{1-g^m}{g^m(1-g)}\frac{1}{\pi_1}.$$

Values of A_C for various values of π_0 and γ with the size of sequences computed by Chen (1978) are presented in column 3 of Table 2. Let $P_0 = (1 - e^{-K})^m$ denote the probability under H_0 that a sequence of length m signals

Table 2. Comparison of Chen's procedure with the Modified Page Procedure (Δ = 96000 in 20 years)

	$\pi_0 \times 10^4$	m(Chen 1978)	A_c	P_0	λ_Δ	a	A_P	A_c/A_P
$\gamma = 6$	1	3	5100	.1548	1.486	1.282	4897	1.04
$K = .77$	2	4	3417	.0831	1.596	1.904	2979	1.15
	3	5	2862	.0446	1.286	2.526	2342	1.22
	4	6	2589	.0240	.921	3.147	2026	1.28
	5	6	2071	.0240	1.151	3.147	1620	1.28
	6	6	1726	.0240	1.381	3.147	1350	1.28
	7	7	1734	.0129	.865	3.769	1312	1.32
	8	7	1518	.0129	.989	3.769	1148	1.32
	9	7	1349	.0129	1.112	3.769	1020	1.32
	10	7	1214	.0129	1.236	3.769	918	1.32
$\gamma = 7$	1	3	4372	.1128	1.083	1.551	4151	1.05
$K = .66$	2	4	2929	.0545	1.046	2.279	2547	1.15
	3	4	1953	.0545	1.569	2.279	1698	1.15
	4	5	1840	.0263	1.011	3.006	1511	1.22
	5	5	1472	.0263	1.264	3.006	1209	1.22
	6	5	1226	.0263	1.516	3.006	1007	1.22
	7	6	1268	.0127	.855	3.734	999	1.27
	8	6	1109	.0127	.977	3.734	874	1.27
	9	6	986	.0127	1.099	3.734	777	1.27
	10	6	887	.0127	1.221	3.734	700	1.27

NOTE: $\pi_0 = P_{H_0}$ (failure); $P_0 = P_{H_0}$ (sequence of length m, in Chen's procedure, signals an alarm); A_C = expected number of observations until a first true alarm using Chen's procedure, λ_Δ = expected number of false alarms in 20 years (i.e. the expected number of false alarms per Δ = 96000 Bernoulli observations) using Chen's procedure; a = detection level with modified Page procedure resulting in an expected number λ_Δ of false alarms in 20 years; A_P = expected number of observations until a first true alarm using the modified Page procedure.

an alarm. The expected number of failures (under H_0) among Δ Bernoulli observations is $\Delta\pi_0$. In the long run a proportion π_0 of the observations will be failures and a proportion P_0 of sequences of length m signal false alarms; hence by the law of large numbers it follows that the expected number λ_Δ of false alarms per Δ (Bernoulli) observations with Chen's procedure is given by

$$\lambda_\Delta = \Delta\ \pi_0 P_0. \quad (7)$$

Using Table 1 we determine the detection level α of the modified Page procedure that results in a false alarm rate dictated by λ_Δ. In order to derive A_P, The expected number of observations until the first alarm under the previously mentioned special case of H_1 with the modified Page procedure, we use results of Lorden and Eisenberger (1973, equations (11) and (14)). These are directly applicable here since until the first alarm the modified Page procedure is identical to the regular Page procedure. (However, we replace Lorden and Eisenberger's upper bound to the expected overshoot by a renewal theoretic approximation of the expected overshoot.)

To make the comparison meaningful we specify the detection level a to match the rate of false alarms of both procedures. The results of these computations are presented in Table 2. The comparison is particularly remarkable for π_0 of the order of 10^{-3} where Chen's procedure requires almost 33 percent more observations than the modified Page procedure to detect a six-fold increase in failure rate.

Perhaps an explanation of this performance of Chen's procedure lies in that it can only signal an alarm after completion of a full sequence of m sets. In fact, scrutiny of A_c in Table 2 shows that under H_1 the expected number of observations is close to m/π_1 for the alternatives considered.

The Chen procedure has an inherent disadvantage with respect to the modified Page procedure: should H_1 be the case, the number of times after the first alarm that an alarm is not sounded (upon the occurrence of another failure) is finite if one uses the modified Page procedure. This number is infinite for Chen's procedure. In other words, if H_1 be the case, an alarm sounded by the modified Page procedure will be more insistent than one sounded by a Chen procedure. Even if this be disregarded, Table 2 indicates that the specific procedures proposed by Chen (1978) seem to do worse than the modified Page procedure. However, if one is willing to consider a Chen-type procedure with $m = 2$ or $m = 1$, a comparison as done above will not show the modified Page procedure to be uniformly better. In fact, for the alternatives considered in Chen (1978), when $m = 1$ the modified Page procedure will be uniformly worse. (This comparison isn't fair: when $m = 1$ the Chen procedure recycles right after each alarm; a fairer comparison would therefore be to the original Page procedure.)

4. REMARKS AND DISCUSSION

We consider here the problem of monitoring shifts in failure rates of continuing processes. One particular feature of the modified Page procedure is that following a first alarm, subsequent alarms carry weight in the decision to continue rejecting the null hypothesis of no shift in the failure rate. These subsequent alarms are assigned weights equal to the first alarm. A possible extension to the modified Page procedure is to assign different weights to subsequent alarms. For example one could give a unit weight to the first alarm and a weight ϵ to the following alarms. Our calculations can be modified to account for such a differential weighting of alarms. In the modified Page procedure and Chen's procedure, $\epsilon = 1$, in the regular Page procedure, $\epsilon = 0$.

Following Page (1954), Pollak and Siegmund (1975), and Chen (1978), among others, we model the shift by a sudden jump at an unknown point in time from a given failure rate to one which is higher. More realistic shifts are gradual and not necessarily monotone so that in the actual implementation of a surveillance scheme one must be aware of possible departures from the sudden shift assumption.

In principle one could apply, instead of a Page procedure, a procedure for detecting a change in distribution such as the procedure described in Pollak and Siegmund (1975). The difficulty would be in determining the appropriate detection level a: inequality (33) in Pollak and Siegmund (1975) is in our case (π_0 small) very far from being an equality (in terms of the number of binomial observations).

APPENDIX: DERIVATION OF THE RATE OF FALSE ALARMS FOR THE MODIFIED PAGE PROCEDURE

In order to compute Λ for given values of a and γ, we proceed in two steps. First we compute the denominator of expression (3) and then we compute its numerator.

A.1 The Denominator

Recall that we transformed our Bernoulli observations to independent exponentially distributed variables X_i such that

$$\text{Under } H_j\ X_i \text{ is Exponential } (\theta_j)\ j = 0, 1 \quad \text{(A.1)}$$

Let M_1 be the number of exponential observations until the first alarm is sounded. Let $M_2 - M_1$ be the additional number of X_i's until $L_n \le 0$. We compute $E(M_2)$ by separately calculating $E(M_1)$ and $E(M_2 - M_1)$.

Using Wald's lemma, we have

$$E_{H_0}\left(\sum_{i=1}^{M_1} X_i\right) = \frac{1}{\theta_0} E_{H_0}(M_1). \quad \text{(A.2)}$$

Let N_1 be the number of exponential variables in a single SPRT. Then, again by Wald's lemma,

$$E_{H_0}(M_1) = \frac{1}{\alpha} E_{H_0}(N_1), \quad \text{(A.3)}$$

since $1/\alpha$ is the expected number of SPRT's in a Page procedure. From Lorden and Eisenberger (1973) we have (in our notation)

$$\alpha = \frac{(\gamma - 1)\, G^*(\gamma)}{e^a \gamma - G^*(\gamma)}, \quad \text{(A.4)}$$

and

$$E_{H_0}(N_1) = \frac{\alpha(a + E_{H_0}(Q_1)) + (1 - \alpha)(1 - \gamma)}{\log \gamma + 1 - \gamma}, \quad \text{(A.5)}$$

where

$$G^*(\gamma) = \frac{\log \gamma + 1 - \gamma}{\gamma - 1 - \gamma \log \gamma}$$

and $E_{H_0}(Q_1)$ is computed and given in (A.12). This completes the computation of $E_{H_0}(M_1)$. We are left with computing $E_{H_0}(M_2 - M_1)$. Let $Q_1 = L_{M_1} - a$, $Q_2 = L_{M_2}$. Define $Y_i = \log \gamma - X_i$ so that $L_n = \sum_{i=1}^n Y_i$. Then clearly

$$E_{H_0}\left(\sum_{i=M_1+1}^{M_2} Y_i\right) = E_{H_0}[-(a + Q_1 + Q_2)]. \quad \text{(A.6)}$$

On the other hand, using Wald's lemma, we have

$$E_{H_0}\left(\sum_{i=M_1+1}^{M_2} Y_i\right) = E_{H_0}(M_2 - M_1)\, E_{H_0} Y_1. \quad \text{(A.7)}$$

Therefore

$$E_{H_0}(M_2 - M_1) = \frac{-E_{H_0}(a + Q_1 + Q_2)}{\log \gamma - \dfrac{1}{\theta_0}}. \quad \text{(A.8)}$$

In order to compute the numerator of (A.8) we first derive the distributions of Q_1 and Q_2. Clearly, under H_0,

$$Q_2 \text{ is Exponential}(\theta_0) \text{ and therefore } E_{H_0}(Q_2) = \frac{1}{\theta_0}. \quad \text{(A.9)}$$

To derive the distribution of Q_1 we have to go through several steps.

Let $Y = \log \gamma - X$ where X is Exponential (θ_0). Then

$$F_Y(y) = P(X > \log \gamma - y)$$

and

$$f_Y(y) = \theta_0 \exp\{\theta_0(y - \log \gamma)\}\, 1_{(y < \log \gamma)}.$$

Since $E(Y_i) < 0$, the ladder distribution of $L_n = \sum_{i=1}^n Y_i$ is defective. To put things into the framework of Feller (1971, p. 405), we consider $-Y$ instead of Y; thus,

$$F_{-Y}(x) = 1 - \exp\{-\theta_0(\log \gamma + x)\} = 1 - \gamma^{-\theta_0} e^{-\theta_0 x}$$

and

$$\mu = E(-Y) = \frac{1}{\theta_0} - \log \gamma > 0,$$

and for $x < 0$,

$$\rho(x) = F_{-Y}(x) + \theta_0 \int_{-\infty}^{x} F_{-Y}(s)\,ds = 1_{(x > -\log \gamma)}\, \theta_0(x + \log \gamma). \quad \text{(A.10)}$$

After returning $-Y$ to Y we obtain the ladder distribution L which is defective uniform on $[0, \log \gamma]$ with total mass $\rho(0) = 1 - \theta_0 \mu = (\log \gamma)/(\gamma - 1)$.

The associated distribution to this distribution (Feller

1971, p. 376) has density

$$L^*(dy) = e^y \theta_0 1_{(0<y<\log\gamma)} dy.$$

Let $Z_x(t) = P$ (there will be an overshoot over t which does not exceed x), and $z_x(t) = P$ (the first ladder variable will exceed t, but won't exceed $t + x$). Then

$$Z_x(t) = z_x(t) + \int_0^t Z_x(t-y)\, L(dy). \quad \text{(A.11)}$$

Invoking Feller (1971 p. 376) we have (denoting $\mu^* = \int_0^\infty y L^*(dy)$) for $0 \le x \le \log\gamma$

$$\lim_{t\to\infty} e^{-t} Z_x(t) = \frac{1}{\mu^*}\int_0^\infty e^s z_x(s)ds = \frac{\theta_0}{\mu^*}(\gamma - \gamma e^{-x} - x)$$

Now we want the distribution of the overshoot Q_1 conditional on there being an overshoot so that

$$P_{H_0}(Q_1 \le x) = \frac{Z_x(t)}{Z_{\log\gamma}(t)},$$

and for $0 \le x \le \log\gamma$,

$$\lim_{t\to\infty} P_{H_0}(Q_1 \le x) = \frac{\gamma - \gamma e^{-x} - x}{\gamma - 1 - \log\gamma}, \quad \text{(A.12)}$$

so that

$$E_{H_0}(Q_1) = \frac{1}{\gamma - 1 - \log\gamma} \times \left[\gamma - \log\gamma - 1 - \frac{1}{2}(\log\gamma)^2\right], \quad \text{(A.13)}$$

and therefore (A.8) yields

$$E_{H_0}(M_2 - M_1) = -\frac{a + \dfrac{1}{\theta_0} + \dfrac{\gamma - \log\gamma - 1 - \frac{1}{2}(\log\gamma)^2}{\gamma - 1 - \log\gamma}}{\log\gamma - (1/\theta_0)} \quad \text{(A.14)}$$

Combining (A.3), (A.4), (A.5), and (A.14), we have the denominator of (3), which is

E_{H_0} (Total number failures until

$L_n \le 0$ for the first time

n after the first crossing of the detection line)

$$= E_{H_0}(M_1) + E_{H_0}(M_2 - M_1)$$

A.2 The Numerator

The numerator of expression (3) is

E_{H_0} (Number of false alarms until

$L_n \le 0$ for the first time

n after the first crossing of the detection line)

$$= 1 + \sum_{m=1}^\infty P_{H_0}\left(a + Q_1 + \sum_{i=1}^m Y_i > a\right) - \sum_{m=1}^\infty P_{H_0}\left(-Q_2 + \sum_{i=1}^m Y_i > a\right) \quad \text{(A.15)}$$

where we recall that Q_1 is the overshoot of L_i over the detection line a and Q_2 is the "undershoot" after L_i returns below the origin (at $i = n$). Now $\sum_{i=1}^m Y_i = m\log\gamma - \sum_{i=1}^m X_i$ where under H_0, X_i are exponential(θ_0) iid. Thus

$$\text{(A.15)} = 1 + \sum_{m=1}^\infty E_{H_0} P_{H_0}\left(Q_1 > -m\log\gamma + \sum_{i=1}^m X_i \,\middle|\, \sum_{i=1}^m X_i\right) - \sum_{m=1}^\infty P_{H_0}\left(-Q_2 + \sum_{i=1}^m Y_i > a\right)$$

Asymptotically (see (A.12)),

$$P\left(Q_1 > -m\log\gamma + \sum_{i=1}^m X_i \,\middle|\, \sum_{i=1}^m X_i\right) = 1 - \left[\gamma - \gamma\exp\left(-\left(-m\log\gamma + \sum_{i=1}^m X_i\right)\right) - \left(-m\log\gamma + \sum_{i=1}^m X_i\right)\right] / [\gamma - 1 - \log\gamma]$$

when $\sum_{i=1}^m X_i - m\log\gamma \in [0, \log\gamma]$. Otherwise it is obviously 0 or 1. Clearly, under H_0, $\sum_{i=1}^m X_i$ is Gamma(θ_0, m).

Dividing $P_{H_0}(Q_1 > -m\log\gamma + \sum_{i=1}^m X_i \mid \sum_{i=1}^m X_i)$ into three cases, such that $-m\log\gamma + \sum_{i=1}^m X_i \in (-\infty, 0), [0, \log\gamma], (\log\gamma, \infty)$, we get

$$E_{H_0} P_{H_0}\left(Q_1 > -m\log\gamma + \sum_{i=1}^m X_i \,\middle|\, \sum_{i=1}^m X_i\right) = R + P(\text{Gamma}(\theta_0, m) < m\log\gamma) = R + P(\text{Gamma}(1, m) < \theta_0 m\log\gamma),$$

where

$$R = S + \left[1 - \frac{\gamma + m\log\gamma}{\gamma - 1 - \log\gamma}\right] P(I),$$

$$S = \frac{\gamma^{m+1}}{\gamma - 1 - \log\gamma} T + \frac{1}{\gamma - 1 - \log\gamma} U,$$

$$T = \left(\frac{1}{\gamma}\right)^m P(II), \quad U = \frac{m}{\theta_0} P(III),$$

$$P(I) = P(\theta_0 m\log\gamma < \text{Gamma}(1, m) < \theta_0(m+1)\log\gamma),$$

$$P(II) = P[(m \log \gamma)\,\theta_1 < \text{Gamma}\,(1, m) < [(m + 1) \log \gamma - a]\theta_1],$$

and

$$P(III) = P[\theta_0 m \log \gamma) < \text{Gamma}\,(1, m + 1) < \theta_0\,(m + 1) \log \gamma].$$

An evaluation of the Q_2 term completes the computations of the numerator:

$$P_{H_0}\left(- Q_2 + \sum_{i=1}^{m} Y_i > a \right) = P_{H_0}\left(\theta_0 \sum_{i=1}^{m} X_i + \theta_0\, Q_2 < \theta_0\,(m \log \gamma - a) \right) = P[\text{Gamma}\,(1, m + 1) < \theta_0\,(m \log \gamma - a)].$$

[*Received April 1981. Revised November 1982.*]

REFERENCES

ANSCOMBE, F.J., and PAGE, E.S. (1954), "Sequential Tests for Binomial and Exponential Populations," *Biometrika*, 41, 252–253.

CHEN, R. (1978), "A Surveillance System for Congenital Malformations," *Journal of the American Statistical Association*, 73, 323–327.

FELLER, W. (1971), *An Introduction to Probability Theory and its Applications*, Vol. II, New York: John Wiley.

HOOK, E.B. (1978), "Monitoring Human Mutations and Consideration of a Dilemma Posed by an Apparent Increase in One Type of Mutation Rate," in *Genetic Epidemiology*, eds. N.E. Morton and C.S. Chung, New York: Academic Press, 481–528.

LORDEN, G., and EISENBERGER, I. (1973), "Detection of Failure Rate Increases," *Technometrics*, 15, 167–175.

MORTON, N.E., and LINDSTEN, J. (1976), "Surveillance of Down's Syndrome as a Paradigm of Population Monitoring," *Human Heredity*, 26, 360–371.

PAGE, E.S. (1954), "Continuous Inspection Schemes," *Biometrika*, 41, 100–115.

——— (1961), "Cumulative Sum Charts," *Technometrics*, 3, 1–9.

POLLAK, M., and SIEGMUND, D. (1975), "Approximations to the Expected Sample Size of Certain Sequential Tests," *The Annals of Statistics*, 3, 1267–1282.

WEATHERALL, M.B., and HASKEY, J.C. (1976), "Surveillance of Malformations," *British Medical Bulletin*, 32, 39–44.