A SURVEILLANCE SYSTEM BASED ON A SHORT MEMORY SCHEME

by

David Lee Shore

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1850T

April 1988

A SURVEILLANCE SYSTEM BASED ON A SHORT MEMORY SCHEME

by

David Lee Shore

A dissertation submitted to the faculty of the
University of North Carolina at Chapel Hill
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Biostatistics

Chapel Hill, 1988

Approved by:

Dana Quade
Adviser

Reader

Lawrence L. Kupper
Reader

ABSTRACT

DAVID LEE SHORE. A Surveillance System Based on a Short Memory Scheme. (Under the direction of Dana Quade.)

A surveillance system is proposed to detect an increase in the mean of a Poisson distribution of cases of a disease. This system is based on a Short Memory (SM) scheme in which conditional binomial tests are performed sequentially at fixed time intervals, say weekly. One parameter of the scheme, say s, determines how far back to look for old cases. In general s will be small. After the new data are used in a test, they become part of the memory replacing the oldest data. The scheme has an sth order Markov property which results in dependencies between tests.

Let R be the random variable which represents the week where an increase in the baseline rate is signalled. R is the run length of the scheme, and the probability of rejection at each week defines its probability distribution. The pdf of R is shown to have a geometric tail. E(R) is derived under null and non-null conditions. Under H_0 and conditional on $R>s$, R follows a geometric pdf with parameter $p=Pr(R>s+1)/Pr(R>s)$. Under H_1 and conditional on $R>2s$, R follows a geometric pdf with parameter $p=Pr(R>2s+1)/Pr(R>2s)$.

When $s=1$ it is shown that a negative correlation between the first two tests causes $E(R|H_0)$ to be smaller than if R followed a geometric

distribution with parameter $p = Pr(R = 1)$. The SM scheme is compared to a scheme with sequential but independent (I) tests. It is shown that $E(R|H_1,SM) < E(R|H_1,I)$ when $E(R|H_0,SM) = E(R|H_0,I)$.

The SM scheme for $s = 2$ is compared to other testing schemes that use the available information in different ways. For the examples chosen the SM scheme is usually superior as measured by $E(R|H_1)$. The limiting test, as $s \rightarrow \infty$, of the likelihood ratio test of the ratio of two Poisson means is derived.

The performance of the SM scheme is compared to the CUSUM (Page, 1954). For small means the CUSUM outperforms the SM scheme for small s. The CUSUM scheme is simulated with a misspecified baseline mean. E(R) for the CUSUM is quite large for misspecifications of 50—60%.

An epidemic in 1970 caused by contaminated intravenous fluid is used for illustration.

ACKNOWLEDGEMENTS

First and foremost I would like to thank Dana Quade. This dissertation would not have been possible without his patience, his guidance and, most of all, his sharing. Next I would like to thank my wife, Bercedis, who supported me in every way throughout this time and is the only person happier than I am to see this project completed. The next round of applause is for Rocky Feuer without whom I would have lost my sense of humor and my way through the days of coursework. Many thanks go to my committee, Larry Kupper, Dick Shachtman, Bert Kaplan and especially P. K. Sen. For all of the others from whom I have drawn inspiration, especially Bob McConnoughey who showed me that the light at the end of the tunnel was not a train coming the other way, I thank you. Finally, I want to thank my parents, Jack and Barbara, who obviously made all of this possible and who now have a son, the doctor.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES (Continued)

Chapter I

Control Charts and Other Epidemic Detection Schemes

1.1. Introduction.

Suppose cases of some disease occur sporadically over time, with the incidence in the t-th time period being some random variable, X_t. Suppose also that the random variables $X_0, X_1, X_2, \cdots$ are mutually independent, and are identically distributed until some unknown time T when a change in the distribution occurs. For example, the common distribution might be Poisson with parameter λ for $t \leq T$ and parameter $\mu = \gamma\lambda$ thereafter, where $\gamma > 1$. The time T may be thought of as the beginning of an 'epidemic' when there was a γ-fold increase. Then the problem, in general terms, is to find a scheme for detecting the change as soon as possible after it occurs, while yet not sounding a false alarm before the change.

Three aspects to consider when constructing and comparing epidemic detection schemes are as follows.

(1) Memory — How far backwards should the scheme look? The value of an accurate baseline measurement has to be balanced against the bias which may occur from using out of date data.

(2) Underlying Process — Assumptions are made concerning the distribution properties of X. How sensitive is the scheme to these assumptions?

(3) Type of Change — As formulated above, the change at time T is visualized as a step function. Other alternatives which could be investigated may be that the change is continual, increasing in expectation in a linear or an exponential fashion.

1.2. Control Charts.

Although we have used the terminology of epidemics, it should be clear that the general mathematical description would also apply to many other situations. One such is the occurrence over time of defectives in a manufacturing process. Control charts have traditionally been used in such settings for detecting changes in the quality of the output from a continuously producing process. The quality of the output is summarized by some parameter θ. If a change in θ is detected, then some rectifying action is taken.

Shewhart (1931) proposed a process inspection scheme in which samples of a fixed size are taken at regular intervals and the observed values of some statistic such as the mean or the range are plotted in temporal order on a 'control chart.' On this chart are drawn 'action lines,' usually at a distance of $\pm 3\sigma$ from θ_0, the value specifying acceptable quality of output. Any point which falls outside these lines signals a change in quality and necessitates rectification. This scheme is referred to as a 'single-sample' scheme since the decision to take action is based on one sample's point on the chart.

One improvement on this scheme is to add 'warning lines' at, say, $\pm 2\sigma$ from θ_0 and add the decision rule that if k of the last n points are outside of the 2σ lines then action is necessary. This allows for detection of smaller

steady changes in the quality of the output.

Even if the quality remains at θ_0, either of the two inspection schemes explained above will eventually signal for action to be taken. That is, there is a non-zero probability that some point will fall outside of the action lines or that k of the last n points will be between the warning and action lines. The probability that one of these will eventually happen is unity. Page (1954) defined the average run length (ARL) of a process inspection scheme as the expected number of articles sampled before action is taken when the quality remains constant. When the quality of the output is satisfactory the ARL is a measure of the run length to be expected before needlessly interfering with the process because of a Type I error. For constant unsatisfactory quality, the ARL is a measure of the amount of scrap produced before intervention, analogous to a Type II error.

Consider the scheme with action lines only. Since m, the number of samples examined before action is taken, is a geometric random variable with $p[\theta]$ equal to the probability that a sample point falls within the action lines, it follows that

$$Pr(m=k) = p^{k-1}(1-p)$$

for $p[\theta] = p$, and

$$E(m) = (1-p)^{-1}.$$

So, for samples of size n the ARL function is

$$ARL = n\,(1-p)^{-1}.$$

For the scheme with both action and warning lines, the ARL is calculated by listing the possible combinations of the positions of the last $(m-1)$ points. Page (1954) cites Bartlett (1953) as giving methods for finding the expected number of samples before a combination demanding action

occurs.

1.3. Decision Interval Cumulative Sum Tests.

The next improvement was to plot cumulative sums of the sample statistic rather than single sample values on the chart. Page (1954) introduced this idea and called the result a cumulative sum (CUSUM) control chart. For detecting one sided deviations from θ, Page suggested using a sequential scheme where observations are made singly at regular intervals and a decision is weighed after each observation. In this scheme a score, sometimes $(y_i - \theta_0) = x_i$, is assigned to each observation y_i and the cumulative score $S_n = \sum_{i=1}^{n} x_i$ is plotted. Action is taken if the present sum S_n exceeds the previous minimum, $\min_{0 \leq i < n} S_i$, by more than say h; that is, if

$$S_n - \min_{0 \leq i < n} S_i \geq h.$$

This rule is equivalent to one where

$$S'_n = \max(S'_{n-1} + x_n, 0) \qquad n \geq 1$$

and

$$S'_0 = z.$$

Now, $S'_n = 0$ whenever $S'_n < \min_{0 \leq i < n} S_i$ and action is taken after the nth observation if $S'_n \geq h$. The initial score S'_0 can be zero but in general z is used. Once the first test ends at zero, all subsequent tests have $S'_0 = 0$.

This rule generates a sequence of Wald sequential tests with boundaries at (0,h) and initial score zero. When the previous test ends at the lower boundary, the test is reapplied. When the upper boundary is reached then action is taken. There is not, in a strict sense, any time at which a null hypothesis is accepted since the test is always restarted if zero

is reached before h.

In deriving the ARL of this scheme there is a similarity to the single sample schemes since the number of sequential tests, R, until action is taken is a geometric random variable. If P[z] is the probability that a test which begins at z will end at the lower boundary, then for P[0] we have

$$\Pr(R=k) = P[0]^{k-1}(1-P[0]), \quad k=1, 2, 3, \ldots .$$

Thus on the average there are $(1-P[0])^{-1}$ sequential tests in one run of a sequential scheme, of which $P[0](1-P[0])^{-1}$ end on the lower boundary and one ends on the upper boundary. If $E(N;0|S_n' \leq 0)$ and $E(N;0|S_n' \geq h)$ are the average sample numbers conditional upon tests ending on the lower and upper boundaries, respectively, given S_0', then in general

$$E(N;z) = P(z)E(N;z|S_n' \leq 0) + [1-P(z)]E(N;z|S_n' \geq h).$$

So,

$$\begin{aligned} E(R;0) &= E(N;0|S_n' \leq 0)P[0](1-P[0])^{-1} + E(N;0|S_n' \geq h) \\ &= [E(N;0|S_n' \leq 0)P[0] + (1-P[0])\, E(N;0|S_n' \geq h)]\,(1-P[0])^{-1} \\ &= E(N;0)(1-P[0])^{-1} . \end{aligned}$$

This expression is similar to the ARL function for single-sample schemes except that the sample size of each test is a random variable with a finite expected value.

Page (1954) and Ewan and Kemp (1960) derived integral equations and equivalent equations for discrete variables for the ARL function, the probability P(z) of a test ending at or below zero, and the average sample number (ASN). First, some definitions. Let f(x) denote the probability of obtaining the value x. This value may be some sample statistic. A reference value, k, will be subtracted from each x; some rules for choosing k will be stated later. The decision value, h, will be the critical value for determining

that action should be taken. For acceptable quality of output, $S_n = \sum_{i=1}^{n}(x_i - k)$ will be less than or equal to zero so only from the first x such that $(x-k) > 0$ is it necessary to accumulate the scores. Also, let z be the value at which the first test begins. Then following Wetherill (1977), the CUSUM scheme proceeds as follows. The current score of a test is z, and the new observation contributes x, so the new score is $z+x-k$, provided this is in the open interval (0,h). If $z+x-k \leq 0$ then the test ends and a new test is started from zero, while if $z+x-k \geq h$ then the test ends and appropriate action is taken; otherwise the test continues.

First a formula for P(z) will be found. For a discrete variate the probability of the new score being zero is equal to $F(k-z) = \sum_{x=0}^{k-z} f(x)$. If the new score is $y = z+x-k$, $0<y<h$, then there is a further probability, P(y), of ending at the lower boundary. If the new score is h then this outcome is not relevant to P(z). Thus, the equation which results is

$$P(z) = F(k-z) + \sum_{y=1}^{h-1} P(y)f(y+k-z) .$$

The ASN and the ARL are determined in a similar fashion. To find the formula for E(N;z) note that for all z, $0 \leq z < h$ there is the new observation in the test; to this is added the probability that the new observation will keep the CUSUM between 0 and h, multiplied by the respective expected sample number from the new value of the CUSUM. Thus,

$$E(N;z) = 1 + \sum_{y=1}^{h-1} E(N;y)f(y+k-z) .$$

The ARL function has the formula

$$E(R;z) = 1 + E(R;0)F(k-z) + \sum_{y=1}^{h-1} E(R;y)f(y+k-z) .$$

The description of the components of the formula is that the next

observation is always there for $0 \leq z < h$, hence the 1. Added to this is the probability that the new observation returns the CUSUM to zero multiplied by the ARL from zero, plus the probability that the new observation lands the CUSUM between 0 and h again multiplied by the respective ARL function from the new value of the CUSUM.

Entirely equivalent formulas are used for the case where x is a continuous variable, so

$$P[z] = F(k-z) + \int_0^h P[y]\, f(y+k-z)\, dy$$

where $F(k-z) = \int_0^{k-z} f(t)\, dt$

and

$$E(N;z) = 1 + \int_0^h E(N;z)\, f(y+k-z)\, dy$$

and

$$E(R;z) = 1 + E(R;0)\, F(k-z) + \int_0^h E(N;z)\, f(y+k-z)\, dy\ .$$

Whenever P[0] and E(N;0) are known, then E(R;0) can be found, since $E(R;0) = E(N;0)(1-P[0])^{-1}$.

1.4. Two-sided Schemes.

It is often desirable to detect changes of parameter in either direction. For convenience suppose that the test is symmetric. Page (1954) suggested a rule that takes action after the nth article sampled if either

$$S_n - \min_{0 \leq i < n} S_i \geq h \quad \text{or} \quad \max_{0 \leq i < n} S_i - S_n \geq h:$$

that is, if the cumulative sum score either rises a distance h above its previous minimum or falls a distance h below its previous maximum. Nadler

and Robbins (1971) showed that this is equivalent to the a procedure based on the range of the sequence of partial sums. Let

$$V_n = \max_{0 \le i < n} S_i \ , \quad U_n = \min_{0 \le i < n} S_i \ .$$

Then, the equivalent rule states that action is necessary when

$$V_n - U_n \ge h \ .$$

To see this suppose that Page's inequality holds for the first time when N=n. Then, $S_N - U_N \ge h$ or $V_N - S_N \ge h$ and certainly $S_N - U_N + V_N - S_N \ge h$. Conversely, suppose N′ is the first observation for which $V_{N'} - U_{N'} \ge h$. Then either a new minimum or a new maximum was just reached, so either $S_{N'} = U_{N'}$ or $S_{N'} = V_{N'}$. Thus, either $S_{N'} - U_{N'} \ge h$ or $V_{N'} - S_{N'} \ge h$. Thus, N′= N, and the two procedures are equivalent.

Page wanted to consider an alternative scheme because of the difficulty of evaluating characteristics of the above procedure. This was to consider the simultaneous application of two Wald tests. Suppose that the two Wald tests have boundaries (−h,0) and (0,h), respectively, and that the first test only accumulates scores when $\sum_{i=1}^{n}(x_i + k) < 0$ until either the lower boundary, −h, is crossed or this CUSUM becomes non-negative, at which point the cumulation ceases for this test. The second test is simultaneously accumulating the scores as long as $\sum_{i=1}^{n}(x_i - k) > 0$ until either the upper boundary, h, is crossed or this CUSUM becomes non-positive. There is a complete dependence between these two one-sided tests, as shown by Kemp (1961). That is, if the CUSUM for one test has absolute value greater than h, then the other test has stopped cumulation of values. Thus, rejection of acceptable quality of output by one test will not interfere with or prematurely terminate the other test from a cumulation which was headed for

rejection. The other test will have ceased cumulation at that observation.

Consider the two sums

$$S_n(r_1) = \sum_{i=n-r_1}^{n} (x_i - \mu - k) \quad \text{and} \quad s_n(r_2) = \sum_{i=n-r_2}^{n} (x_i - \mu + k)$$

where

$$0 < S_n(j) < h \text{ for } 0 \le j \le r_1$$

and

$$-h < s_n(l) < 0 \quad \text{for} \quad 0 \le l \le r_2 \quad .$$

If $S_{n+1}(r_1+1) \ge h$, then

$$(x_{n+1} - \mu) \ge h + k - S_n(r_1)$$

so that

$$\begin{aligned} s_{n+1}(r_2+1) = s_n(r_2) + (x_{n+1} - \mu) + k \ge s_n(r_2) - S_n(r_1) + h + 2k \\ s_{n-r_1-1}(r_2 - r_1 - 1) + 2(r_1+2) + h \quad , \quad r_2 > r_1 \\ = \quad h + 2(r_2+2) \quad , \quad r_2 = r_1 \\ h + 2(r_2+2)k - S_{n-r_2-1}(r_1 - r_2 - 1) \quad , \quad r_1 > r_2 \end{aligned}$$

so that $s_{n+1}(r_2+1) > 0$. Similarly, if $s_{n+1}(r_2+1) \le -h$ then $S_{n+1}(r_1+1) < 0$.

The probability that the boundary at h is crossed is equal to $[E(R|s_n(r_2) > -h, \text{ for all } n, r_2)]^{-1}$. If T samples are taken then the number of occasions when the boundary might be crossed is $T[E(R|s_n(r_2) > -h, \text{ for all } n, r_2)]^{-1}$. Similarly, if the other component CUSUM is considered simultaneously this neither affects nor is affected by the first CUSUM, so that in T samples the boundary $-h$ is expected to be crossed T $[E(R|S_n(r_1) < h, \text{ for all } n, r_1)]^{-1}$. It follows then that the expected run length of the two-sided scheme for any T samples is T divided by the number of times that either boundary is crossed. So,

$$E(R) = T / \left\{ T[E(R|s_n(r_1) > -h)]^{-1} + T[E(R|S_n(r_1) < h)]^{-1} \right\}$$

and

$$[E(R)]^{-1} = [E(R|s_n(r_1) > -h)]^{-1} + [E(R|S_n(r_1) < h)]^{-1} .$$

Van Dobben de Bruyn (1968) proved that this relationship will also be true for asymetrical two-sided CUSUM schemes, if and only if

$$k^+ + k^- \geq | h^+ - h^- | .$$

1.5. Graphical Procedures for CUSUMs.

Barnard (1959) suggested using a V-mask placed in a particular way on the CUSUM chart. If any point is obscured by the mask then this is taken as an indication of a need for action. The blanked out portion of (n,S_n)-plane has a vertex at the point $(n+d, s_n)$, a distance d from the last point plotted, with two limbs inclined at angle ϕ to the horizontal. The cumulative sum $S_n = \sum_{i=1}^{n}(x_i - \mu)$ is plotted.

Kemp (1961) discussed a situation where the mean value of x is to be kept at the target value μ. Let the horizontal distance between successive points be w when measured in terms of the unit distance on the vertical scale. The lower limb of the V-mask is crossed when

$$\sum_{i=n-r+1}^{n} (x_i - \mu - w\tan\phi) \geq d\tan\phi \quad .$$

Similarly, the upper limb is crossed when

$$\sum_{i=n-r+1}^{n} (x_i - \mu + w\tan\phi) \leq -d\tan\phi \quad .$$

The use of a V-mask on a CUSUM chart is seen to be equivalent to simultaneously operating two single-sided decision interval schemes that are symmetric with $h = d\tan\phi$ and $k = w\tan\phi$.

Johnson (1961) drew a parallel between the use of a V-mask and a sequential ratio probability test (SPRT) applied 'in reverse'. This connection, though imprecise, is useful in determining the values of the parameters of

this scheme: the lead distance d, and the angle ϕ. The limbs of the V-mask may be regarded as the lower and upper limits of a chart of the type used by Armitage (1950) to discern between

$$H_{-1}: \mu=-\delta \qquad H_0: \mu=0 \qquad H_1: \mu=\delta$$

applied to the sequence of adjusted values

$$\{y_i\}_{i=1}^{n} \text{ where } y_1=\frac{\bar{x}_n-\mu}{\hat{\sigma}_{\bar{x}}}, \quad y_2=\frac{\bar{x}_{n-1}-\mu}{\hat{\sigma}_{\bar{x}}}, \text{ etc.}$$

That is, the observations are viewed in reverse, from the last to the first. The boundaries for the SPRT will be $(\beta/(1-\alpha),(1-\beta)/\alpha)$ where α and β are the Type I and Type II errors, respectively.

Assuming normality, so that $\mathcal{L}\mathcal{R}=\log\frac{f(y,\delta)}{f(y,0)} = \frac{1}{2}[2\delta\sum_{j=1}^{n}y_j-n\delta^2]$, leads to the following rule. Decide on the least change in the mean which it is necessary to detect, say D; then $\delta=D/\hat{\sigma}_{\bar{x}}$. Decide on a false alarm rate 2α. Then $\phi=\tan^{-1}(D/2\hat{\sigma}_{\bar{x}})$ and $d=2\delta^{-2}\log((1-\beta)/\alpha)$. It is fair to assume that β is small since the null hypothesis will never be accepted; the acceptance and continuation regions are melded together. Therefore, $\Pr(H_0|H_1)$ is meaningless. Thus, $d \doteq -2\delta^{-2}\log\alpha$.

Van Dobben de Bruyn (1968) pointed out two assumptions which the above parameterization depends on. The first is that although α will not exactly be the proportion of false alarms, tests with the same α have approximately the same proportion of false alarms. That is, these tests have approximately the same ARL when the null hypothesis is true. The second is that if the true parameter represents a shift of δ standard deviations from the null parameter, (showing a bias of $\delta\hat{\sigma}_{\bar{x}}$), then the ARL is equal to $-2\delta^{-2}\log\alpha$ which is equal to $-\log\alpha[E(\mathcal{L}\mathcal{R}|H_1)]^{-1}$. To see this, consider

$$\begin{aligned} E_{H_1}(N) &= [(1-\beta)\log(B)+\beta\log A]/E_{H_1}(R) \\ &= [(1-\beta)\log((1-\beta)/\alpha)+\beta\log(\beta/(1-\alpha))][\tfrac{1}{2}\delta^2]^{-1} \end{aligned}$$

$$= \log((1-\beta)/\alpha)[\tfrac{1}{2}\delta^2]^{-1}$$

$$\doteq -\log\alpha[\tfrac{1}{2}\delta^2]^{-1} .$$

Since the sums are normalized here, take $w=1$, $h=d\tan\phi=-\delta^{-1}\log\alpha$ and $k=\tan\phi=\frac{1}{2}\delta$. Then, this ARL is equal to $d=h/k$. This points out that the detection of a shift of δ standard deviations is best done by using $k=\frac{1}{2}\delta$ independently of what the other parameters are. This is referred to as the 'central reference value' becuase it places k halfway between the 'acceptable quality level' (AQL) represented by μ and the 'rejectable quality level' (RQL) represented by $\mu+\delta$ or $\mu-\delta$.

1.6. The Run Length Distribution of CUSUM Tests.

Let p(r;z) be the probability that a test starting at z will have run length R equal to r. Following an argument which leads to the equation for P(z) produces the equation

$$p(r;z) = p(r-1;0)F(k-z) + \int_0^h p(r-1;y)f(y+k-z)dy$$

when x is a continuous variate and

$$p(r;z) = p(r-1;0)F(k-z) + \sum_{y=0}^{h-1} p(r-1;y)f(y+k-z)dy$$

when x is a discrete variate.

Let $\chi_R(z,t)=\sum_{r=1}^{\infty} p(r;z)e^{rt}$ be the moment generating function of the run length distribution. Denote by N the value of the sample number n when either the lower or the upper boundary is crossed. Then S_n is the value of the CUSUM whenever $0<S_n<h$, and S_N whenever this inequality is violated. Also, $P(z)=Pr(S_N\leq 0)$. Finally let $\chi_N(t,z)$ be the moment generating function of the sample number distribution.

It is possible to find the moments of the run length R directly by using

$$e^{-t}\chi_R(z,t) = p(1,z) + \sum_{r=2}^{\infty} p(r,z)e^{(r-1)t}$$
$$= 1 - F(h+k-z) + \chi_R(0,t)F(k-z) + \int_0^h \chi_R(y,t)fy+k-z)dy$$

for continous x. (The discrete case is not derived here but follows the same reasoning). Differentiating and setting $t=0$ yields the moments of R. As derived before,

$$E(R;z) = 1 + E(R;0)F(k-z) + \int_0^h E(R;y)f(y+k-z)dy \ .$$

Also,

$$E(R^2;z) = 2E(R;z) - 1 + E(R^2;0)F(k-z) + \int_0^h E(R^2;y)f(y+k-z)dy \ .$$

The moment generating function of R can also be expressed as a function of the moment generating functions of the number of samples in a single acceptance test and a single rejection test, $\chi_N(t|S_N \le 0)$ and $\chi_N(t|S_N \ge h)$, respectively. The number of tests in an inspection scheme is distributed as a geometric random variable with parameter P(z). Assume for now that all tests begin at the acceptance boundary, and drop the value z from the notation. Following Feller's (1959) theorem on a compound function,

$$\chi_R(t|N) = [\chi_N(t|S_N \le 0)]^{R-1}\ \chi_N(t|S_N \ge h)$$

so,

$$\chi_R(t) = \sum_{r=1}^{\infty} \chi_R(t|N) = [\chi_N(t|S_N \le 0)]^{r-1}\ \chi_N(t|S_N \ge h)(P[0])^{r-1}(1-P[0])$$

$$= \frac{(1-P[0])\chi_N(t|S_N \ge h)}{1-P[0]\chi_N(t|S_N \le 0)} \ .$$

Differentiating and setting t equal to zero yields

$$E(R;0) = \chi'_R(0) = (1-P[0])^{-1}E(N;0|S_N \le 0) + E(N;0|S_N \ge h)$$
$$= (1-P[0])^{-1}E(N;0)$$

as derived by Page.

Let V(N;0) and V(R;0) represent the variance of the sample number and

the run length, respectively, where all tests have initial score zero. Then,

$$V(R;0) = V(N;0)[1-P[0]]^{-1} + E(R;0)P[0]\left\{\frac{2E(N;0|S_N \le 0)}{1-P[0]} - E(R;0)\right\} .$$

At AQL, $1-P[0]$ is usually very small. That is, large values of ARL are due to many repetitions of the test rather than large ASN. So,

$$E(N;0|S_N \le 0)P[0][1-P[0]]^{-1} = E(R;0) - E(N;0|S_N \ge h) \doteq E(R;0)$$

and

$$V(R;0) \doteq V(N;0)[1-P[0]]^{-1} + E^2(R;0) .$$

At AQL the standard deviation of R can be approximated by E(R;0). It seems reasonable, then, to expect that p(r,0) will have approximately an exponential distribution, except for small values of R. Ewan and Kemp (1960) suggested this and derived the expression

$$p(r;0) \doteq [E(R;0)]^{-1} \exp\left\{-(r-1)[E(R;0)]^{-1}\right\} .$$

The probability of obtaining a run length less than or equal to cE(R;0) is then

$$\sum_{r=1}^{cE(R;0)} p(r;0) \doteq \int_1^{cE(R;0)} [E(R;0)]^{-1} \exp\left\{-(r-1)[E(R;0)]^{-1}\right\} dr$$

$$= 1 - \exp\left\{[E(R;0)]^{-1} - c\right\} .$$

Kemp (1971) investigated the distribution of R for large values of R and found the expression

$$p(r;0) \doteq [E(R;0)]^{-1} \exp\left\{-r[E(R;0)]^{-1}\right\} .$$

This leads to the probability of obtaining a run length R less than or equal to cE(R;0) being equal to

$$\sum_{r=1}^{cE(R;0)} p(r;0) \doteq 1-\exp(-c)$$

when cE(R;0) is a positive integer. This expression is convenient since the cumulative probability depends only on c.

For a particular discrete situation, Brook and Evans (1972) introduced

Markov chains as a means of examining the actual probability distribution of run length. This was originally proposed by Page (1954), but he did not present further work along these lines. The case Brook and Evans consider is that of a discrete random variable where h and k have integer values. They define

$$S_n = \sum_{i=1}^{n} (D_i - k)$$

where D_i is the number of defectives in the ith sample. Plotting of (n, S_n) commences when D_j exceeds k for the first time and continues until either h is exceeded, signaling a need for action, or S_n reaches zero again, signaling that the process is in control. The CUSUM S_n can only take on one of the values 0,1,2, ... , h. The Markov chain is in the state ϵ_i if $S_n = i$. Each realization of the inspection scheme can then be regarded as a random walk over the states $\epsilon_1, \epsilon_2, \ldots, \epsilon_h$, where ϵ_h is an absorbing state.

The probability distribution completely determines the transition probabilities from state ϵ_i, $i = 0,1,2, \ldots, h-1$: Let

$$p_{i0} = Pr(\epsilon_i \to \epsilon_0) = Pr(D \leq k - i)$$

$$p_{ij} = Pr(\epsilon_i \to \epsilon_j) = Pr(D = k + j - i) \quad j = 1, 2, \ldots, h-1$$

$$p_{ih} = Pr(\epsilon_i \to \epsilon_h) = Pr(D \geq k + h - i) \quad .$$

These equations form a Markov chain and its transition matrix, ($\underset{\sim}{P}$, say), can be completely constructed from the above equations. Let $\underset{\sim}{T}$ be the matrix obtained from $\underset{\sim}{P}$ by deleting its last row and column. Also, let R_i be the number of steps to go from state ϵ_i to the absorbing state ϵ_h for the first time and let $\underset{\sim}{\mu}^{(s)}$ be the vector of sth factorial moments for the random variables $R_0, R_1, \ldots, R_{h-1}$. In matrix form this is $(\underset{\sim}{I} - \underset{\sim}{T})\mu^{(s)} = s\underset{\sim}{T}\underset{\sim}{\mu}^{(s)}$, s=2,3,... . In particular, for s=1, $\underset{\sim}{\mu}(\underset{\sim}{I} - \underset{\sim}{T})^{-1}\underset{\sim}{1}$. In general, the ith element of $\underset{\sim}{\mu}$ is equal to E(R;i) for $i = 0, 1, 2, \ldots, h-1$. Successive substitution leads to the solution

$$\underset{\sim}{\mu}^{(s)} = s!\underset{\sim}{T}^{s-1}(\underset{\sim}{I}-\underset{\sim}{T})^{-s}\underset{\sim}{1} \quad .$$

Thus, the probability distribution of $\underset{\sim}{R}$ is a multidimensional generalization of a geometric distribution.

To help derive the probability distribution of the run length, define the r-step transition matrix as $\underset{\sim}{P}^r$ and $P_h^{(r)}$ as the vector of probabilities of going from, say, state ϵ_i $i=0, 1, 2, \ldots, h-1$ to state ϵ_h in r steps. Thus, the ith element of $P_h^{(r)}$ equals $Pr(R_i \leq i)$, $i=0, 1, 2, \ldots, h-1$. Then ,

$$P_h^{(r)} = (\underset{\sim}{I}-\underset{\sim}{T}^r)\underset{\sim}{1}, \quad r=1, 2, \ldots \quad .$$

Define

$$\underset{\sim}{L}_r = (Pr(R_o=r), Pr(R_1=r), \ldots, Pr(R_{h-1}=r))', \quad r=1,2, \ldots$$

Then

$$\underset{\sim}{L}_1 = \underset{\sim}{P}_h = (\underset{\sim}{I}-\underset{\sim}{T})\underset{\sim}{1}$$

and

$$\underset{\sim}{L}_r = \underset{\sim}{T}^{r-1}\underset{\sim}{L}_1 = \underset{\sim}{T}^{r-1}(\underset{\sim}{I}-\underset{\sim}{T})1 \quad \text{for } r=1,2, \ldots \quad .$$

Brook and Evans also found a limiting form of the probability distribution of run length. The formula is analogous to the ones found by Ewan and Kemp and by Kemp.

Another result for large R is as follows. Let λ be the largest real eigenvalue $\underset{\sim}{T}$. It is shown to exist and to be smaller than one. Then for any initial starting state, the conditional probability $Pr(R_i=r|R_i \geq r)$ tends to $(1-\lambda)$ as r tends to infinity and is thus invariant. Then, for large R, $E(R;i) \doteq 1/(1-\lambda)$.

The advantages of this approach can be seen not only in the specification of the probability distribution of the run length but also in the ease with which the ARL can be found directly. Instead of having to depend

on knowing two of the three values in the equation

$$E(R;0) = E(N;0)(1-P[0])^{-1}$$

or the moment generating functions of the conditional sample numbers, the individual tests are suppresed and the first passage time to state ϵ_h is elicited. Also, inspection schemes can be compared in another fashion by looking at their respective conditional probability distributions given they are still inspecting, namely $(1-\lambda)$.

1.7. Approximations to the ARL using Brownian Motion.

Generally, the discrete time process of a CUSUM can be approximated by a Brownian motion process and correspondingly its ARL can be approximated by the expected first passage time for the continuous time process. This includes work done by Nadler and Robbins (1971), Reynolds (1972, 1975), Johnson and Bagshaw (1974) and Bagshaw and Johnson (1975).

Consider the two-sided procedure. Allowing for a general value for k but assuming that $z=0$ when $t=0$, it is found that if h^- is violated first then

$$E^-(R;0) = \begin{cases} -\dfrac{1}{\mu+k}\left[h - \dfrac{\sigma^2}{2(\mu+k)}\left(\exp\left(\dfrac{2(\mu+k)h}{\sigma^2}\right)-1\right)\right] & , \ \mu+k \neq 0 \\ h^2/\sigma^2 & , \ \mu+k=0 \ . \end{cases}$$

If h^+ is violated first, then

$$E^+(R;0) = \begin{cases} \dfrac{1}{\mu+k}\left[h + \dfrac{\sigma^2}{2(\mu+k)}\left(\exp\left(\dfrac{-2(\mu-k)h}{\sigma^2}\right)-1\right)\right] & , \ \mu-k \neq 0 \\ h^2/\sigma^2 & , \ \mu-k=0 \ . \end{cases}$$

These expressions hold whether the observations are normally distributed or not. Under the hypothesis of no change in μ, Donsker's invariance theorem states that the distribution of the test statistic does not depend on the particular underlying distributions of the observations so long as their variance is finite.

The ARL for the two-sided scheme is the combination of the individual one-sided schemes as given before. That is $\frac{1}{E(R;0)} = \frac{1}{E^{+}(R;0)} + \frac{1}{E^{-}(R;0)}$.

1.8. Evaluating the Approximations.

Reynolds (1975) compared the Brownian motion approximation to the ARL with the tables in van Dobben de Bruyn (1968). The approximation consistently underestimated the tabled values, quite seriously for small values of h. He showed that if $h+\Delta$ is used instead of h the approximation is better, where Δ is chosen such that the approximate ARL^{+} is equal to the exact ARL^{+} at $\mu-k=0$. It is not an ideal solution since, for instance, for $\mu-k=-.2$ the modified approximation overestimates the ARL^{+} by 20–25% for $h=10.0$.

Johnson and Bagshaw (1974) compared the asymptotic distribution of run length with a histogram based on 5000 simulations for one set of values of h and k. The distribution of first passage time has the same shape as the histogram but also has a negative location bias.

1.9. Effects of Correlation and Variance Estimation on CUSUMS.

The effects of non-independence of observations on the ARL were first studied by Goldsmith and Whitfield (1961) using a simulation based on a symmetric CUSUM procedure. For the values of the parameters studied, at

AQL the effect of correlation is dramatic: the ARL is very much larger for large negative correlations, decreasing steadily as the correlation becomes positive until a steep rise towards infinity near perfect positive correlation. At RQL, however, the correlation does not have much impact on the ARL except when it is near unity; then the ARL is infinite. A lack of robustness of the CUSUM procedure seems to exist with regard to correlated observations. This is borne out by later research.

A theoretical investigation was carried out by Johnson and Bagshaw (1975). They found that at AQL the effect of negative correlation was to decrease the true Type I error rate, thus increasing the ARL for a truncated version of Page's (1955) test. An investigation of Page's one-sided test showed that relatively small correlations lead to widely varying shapes of the asymptotic run length distribution. This effect on the ARL is similar to what was reported above except that correlations near unity were not investigated. The asymptotic distributions were again checked against a histogram based on 5000 simulations of the run length distribution and showed, again, good agreement but a slight negative location bias.

Van Dobben de Bruyn (1968) stated that the standard deviation is usually estimated from the same observations used in the CUSUM procedure and that errors of the order of 10—20% are common. The effect is to multiply h and k by the same factor as the wrongful estimation. Bagshaw and Johnson (1975) used the asymptotic ARL of the CUSUM test and the fact that $s \sim \sigma^2\chi^2(n-1)$ to find the expected asymptotic ARL with respect to s. This expression converges to the asymptotic ARL as $n \rightarrow \infty$. For finite samples they will not necessarily be equal.

1.10. The SETS Detection Scheme.

A surveillance scheme was proposed by Chen (1978) for the purpose of monitoring congenital malformations (CMs) in newborns. This system is based on the number of consecutive births which occur between the birth of an infant with the CM and the next birth of an infant with a CM. This group of births, excluding the two births with the CM, is defined as a 'set'. Assuming CMs occur at random, set size is a geometric variable. If there occurs a sequence of sets in which each set is below a certain size, then an alarm is sounded. Chen suggested that while the CUSUM technique may have a smaller ARL, this technique can have a quicker detection time if the change is large enough. The reason for this is that the CUSUM operates on a fixed schedule while under her scheme (SETS) a new analysis begins every time an infant is born with the CM under study.

Chen's scheme depends on three parameters, which are defined in the notation of this thesis as follows:

γ = the proportional increase in the rate of the monitored CM

n = the number of consecutive sets which are considered together for alarm purposes.

c = a multiplying factor used when expressing the size of the set.

Let X_i be the size of set i, and let λ_0 be the normal rate of the CM. Under H_0,

$$E(X_i) = \mu_0 = (1-\lambda_0)/\lambda_0 \qquad i=1,2,\ldots$$

and

$$\begin{aligned} Pr(X_i < \mu_0) &= 1-(1-\lambda_0)^{\mu_0} \\ &= 1-(1-\lambda_0)^{\frac{1}{\lambda_0}-1}. \end{aligned}$$

Since $1/\lambda_0$ is very large,

$$\Pr(X_i < \mu_0) \doteq 1 - e^{-1}$$

and in general

$$\Pr(X_i < c\mu_0) = 1 - (1-\lambda_0)^{c\mu_0}$$
$$\doteq 1 - e^{-c} \qquad i=1,2, \ldots .$$

Under H_1,

$$E(X_i) = \mu_1 = (1-\gamma\lambda_0)/\gamma\lambda_0 \doteq \mu_0/\gamma \qquad i=1,2, \ldots$$

and

$$\Pr(X_i < c\mu_0) = 1 - (1-\lambda_1)^{c\mu_0}$$
$$= 1 - (1-\lambda_1)^{\frac{c\gamma}{\lambda_1}-1} \qquad i=1,2, \ldots .$$

Let P_i be the probability under H_i that a given sequence signals an alarm, $i=0,1$. Then,

$$\Pr(X_{max} \leq c\mu_0) = \begin{cases} (1-e^{-c})^n \doteq P_0 \\ (1-e^{-\gamma c})^n \doteq P_1 \end{cases}$$

where X_{max} is the size of the largest set in the sequence of length n. Since $1-e^{-c\gamma}$ is specified, (usually large so that P_i will be large), c can be found.

Since $P_0 \doteq (1-e^{-c})^n$, it follows that $n \doteq \log(P_0)/\log(1-e^{c})$. First P_0 is found from the number of false alarms to be expected in a given interval of time, say l. If there are r newborns in this period of time then it is expected that $r\lambda_0$ of them will have the CM. Since each newborn with the CM starts a set, and each set starts a sequence, there are $r\lambda_0-(n-1)$ sequences of length n expected. Then,

$$l = [r\lambda_0-(n-1)]P_0$$

so

$$P_0 = l/[r\lambda_0-(n-1)]$$

and n is found.

The expected number of newborns included in a sequence which signals

an alarm after an increase, say r_1, is equal to

$$\begin{aligned} r_1 &= nE(X_i \mid X_i < c\mu_0) \\ &= n[(1-\lambda_1)/\lambda_1 - c\mu_0((1-\lambda_1)^{c\mu_0}/(1-(1-\lambda_1)^{c\mu_0}))] \quad , \quad i=1,2,\ldots \\ &\doteq n\mu_1(1-c\gamma/(e^{c\gamma}-1)) \end{aligned}$$

The time to detection under epidemic conditions is found from r_1 and the birth rate.

Kenett and Pollak (1983) used a sequential sampling approach to find a more accurate formula for the expected number of newborns in a sequence which correctly signals an alarm. With X_i defined as above, let

$$B_i \equiv X_i + 1 ,$$

where

$\sum_{i=1}^{N} B_i$ is the number of observations until a first alarm is signalled

and

N is the number of sets observed in that period of time.

Then

$$r_1 = E\Big[\sum_{i=1}^{N} B_i \mid H_1\Big] = E(N|H_1)\, E(B_i|H_1)$$

$$= E(N|H_1)\, \frac{1}{\lambda_1} .$$

Letting $g = 1 - e^{-c\gamma}$ they found that $\Pr(N \geq j|H_1) = \dfrac{\Pr(N = j+n|H_1)}{g^n(1-g)}$, for $j \geq 1$ and that $\Pr(N = j+n|H_1) = g^n(1-g)$ for $j = 1, 2, \ldots, n$.

Now,

$$E(N|H_1) = \sum_{j=1}^{\infty} \Pr(N \geq j|H_1)$$

$$= m + \sum_{j=m+1}^{\infty} \frac{\Pr(N = j+m|H_1)}{g^n(1-g)}$$

$$= \frac{1-g^n}{g^n(1-g)} .$$

Thus,

$$r_1 = \frac{1-(1-e^{-c\gamma})^n}{(1-e^{-c\gamma})^n e^{-c\gamma}} \frac{1}{\lambda_1} \quad .$$

This yields a larger value for r_1 than that derived by Chen (1979).

Chen also suggested a procedure for combining the data from several different hospitals. In this scheme, the testing is done at fixed intervals of time. The size of each completed set is recorded along with the total number of sets. If all m of the sets are smaller than $c_m\mu_0$ then an alarm is sounded. The value of c_m is chosen so that under H_0, the probability of a false alarm given that at least one set has been completed is q_0. Thus,

$$q_0 = (1-e^{-c_m})^m$$

so

$$c_m = -\log(1-q_0^{1/m}) \ .$$

Let M be the number of sets completed in the given time interval. Then,

$$P_0 = \sum_{m=1}^{\infty} \Pr(X_{max} < c_m\mu_0 | M=m) \Pr(M=m)$$

$$= q_0 \sum_{m=1}^{\infty} \Pr(M=m)$$

$$= q_0(1-(1-\lambda_0)^N)$$

where N is the number of newborns observed in that period. Thus

$$q_0 = P_0/(1-(1-\lambda_0)^N)$$

For large N, $q_0 \doteq P_0$, and under epidemic conditions

$$P_1 = \sum_{m=1}^{\infty} (1-e^{-c\gamma})^m \Pr(M=m) \quad .$$

Chen, Mantel, Connelly and Isacson (1981) proposed a scheme which is close in nature to the one proposed by Chen (1978). The major difference is

that the underlying distribution of the cases (or new diagnoses) in a geographic area is assumed to be Poisson. Then, the time between consecutive new cases has an exponential distribution. The advantage of the distributional assumptions here is that the probabilities are in an easily manageable form without approximations.

For the large scale scheme which combines data from several sources, analysis is done at fixed intervals. In this scheme, each source has its own critical duration and total number of intervals and neither is fixed in advance. However, P_0 is made to be equal for all the separate sets of parameters. Then, considering the number of sources which signal an alarm as a binomial variate, with parameters P_0 and n equal to the total number of sources, an alarm is sounded if the observed number is large.

Chen (1979) compared the CUSUM scheme with the SETS scheme and with a scheme in which the baseline rate is used to find an expected value and a simple test of hypothesis of observed against the expected number of cases is done. In all three of the schemes a Poisson distribution of new cases was assumed. Baseline rates from 1.8/10,000 to 13.0/10,000 and a variety of increases were used based on the tables given in Ewan and Kemp (1960) for the CUSUM scheme. The comparison was in terms of the expected time to detection or ARL and used the estimate rather than the exact formula for r_1. It is easy to reexpress this in terms of number of births until detection, etc. The results show that the SETS technique outperforms the simple Poisson scheme for 400 births/month but only does as well as or slightly worse than the Poisson scheme for 1200 births/month. More importantly, the CUSUM scheme is superior to the SETS scheme for the range of baseline rates and the fold increases exhibited. Barbujani and Calzolari (1984) performed a

Monte Carlo simulation that compared the SETS scheme CUSUM scheme for four different CMs with baseline rates of 1/435, 1/578, 1/660 and 1/1193 and three different increases for an epidemic (2, 3.5 and 5). For this range of values, the CUSUM scheme had greater sensitivity, specificity and accuracy (defined as the ratio of correct outcomes to all outcomes) than the SETS scheme and, on average, had a shorter time to detection. Chen (1984) said that this study was not carried out correctly since CMs with different rates were lumped together and that the proper comparison was not made. This last point is in regard to the fact that the SETS scheme was applied to data generated as if from individual 'hospitals' but the CUSUM scheme was applied to the data from all of the 'hospitals' lumped together. The authors stated in a rejoinder that this comparison may not be the most benevolent to the SETS procedure but explained that it is a typical way that monitoring would be done and therefore the results are valid. Chen (1987) did a Monte Carlo simulation which shows that for CMs with small expected number of cases in a year, (≤ 5), the SETS technique has a smaller expected time to detection than the CUSUM technique but that the relationship is reversed for higher expected number of cases in a year for a 1.5 increase under epidemic conditions. This is the situation that the SETS scheme was introduced for, namely, a rare disease and a necessity for detecting a small increase in the number of cases. The SETS scheme also outperforms the CUSUM scheme when the fold increase is large even for small baseline rates. The simulation also shows that lengthening the time between CUSUM tests can result in the CUSUM scheme being more efficient than the SETS scheme. This makes it difficult to recommend one technique over the other at present.

1.11. Modified CUSUM Scheme.

Kenett and Pollak (1983) suggested a modified CUSUM procedure as the basis of a detection scheme. They were interested in detecting a shift in the failure probability in a sequence of Bernoulli random variables, similar to Chen.

The modification to the usual CUSUM technique is that once an alarm is sounded, that is $S_n - \min_{0 \le i < n} S_i \ge h$, then the CUSUM does not reset but continues to add observations until $S_n - \min_{0 \le i < n} S_i \le 0$ at which point the modified procedure will reset by discarding all observations. The modified CUSUM scheme will sound alarms between the first alarm and resetting so long as $S_n - \min_{0 \le i < n} S_i \ge h$. This is similar to how the unmodified SETS procedure works since the sets which set off an alarm are not immediately discarded but continue to be used for a while in conjunction with new sets to check for increased rates. The overall false alarm rate takes these subsequent alarms into account, thus leading to a different value for the upper boundary, h, and a different ARL than the usual CUSUM scheme. Since the modified CUSUM scheme is a regular CUSUM scheme until the first n for which $S_n - \min_{0 \le i < n} S_i \ge h$ the values of α and β of both schemes up to that point are identical. Under the modified CUSUM scheme, the rate of false alarms is given by

$$\Lambda = \frac{E\left\{\#\text{ false alarms until } 1^{st} \text{ reset such that } S_n - \min_{0 \le i < n} S_i \ge h \mid H_0\right\}}{E(\#\text{ CMs until } 1^{st} \text{reset after } 1^{st} \text{ alarm} \mid H_0)} .$$

Chen (1978), the modified CUSUM scheme has an ARL which is anywhere from 4% to 32% smaller than the SETS scheme under H_1 when the ARL is set to be equal under H_0.

1.12. Modified SETS Scheme.

The SETS technique was amended by Gallus, Mandelli, Marchi and Radaelli (1986) to utilize a more realistic alarm definition. They pointed out that once the first $(n-1)$ sets have been observed then every new set could cause an alarm. Thus, the requirement for the number of false alarms should be changed so that there is, on average, one every $r\lambda_t$ newborns which leads to $P_t=(r\lambda_t)^{-1}$. They also required that consecutive alarm sets be disjoint so that once n sets have sounded an alarm, a further n sets are required for a new alarm to sound. This is similar to how the CUSUM procedure works since the CUSUM is reset to zero after an alarm is sounded and completely new data are needed to work towards sounding a new alarm. This alarm definition says that a sequence of n sets can only sound an alarm if it is preceded by a 'large' set or 2n sets preceded by a 'large' set, etc. Let $p_t=\Pr(X_{max}\leq c\mu_0|H_t)$. Then

$$P_t = p_t^n(1-p_t)[1+p_t^n+(p_t^n)^2+(p_t^n)^3+\ \dots\ ]$$

$$= \frac{p_t^n(1-p_t)}{1-p_t^n}\ .$$

Thus,

$$r\lambda_t = \frac{1-p_t^n}{p_t^n(1-p_t)}\ .$$

They define a system that has optimal efficiency if it minimizes the expected delay to an alarm after an increase in the baseline rate of the CM under study. This says to minimize $r\lambda_t$, the number of CMs before the first alarm.

In the examples shown, the modified SETS scheme always outperforms the original SETS scheme, sometimes reducing the expected delay by as much as 79%. The CUSUM scheme is also inferior to the modified SETS scheme if

the increase in the infection rate is large enough ($\gamma > 4$). Gallus et al. pointed out that this is the type of increase they were interested in detecting, namely a large single jump in the rate of CMs.

Chapter II

Overview of Short Memory Schemes

2.1. Motivation for Surveillance Model.

The motivation for a model which is useful for the surveillance of infection rates stemmed from a problem posed by the Centers for Disease Control (CDC) for the National Nosocomial Infections Study (NNIS) (Haley and Garner, 1985). They were interested in monitoring site by pathogen infection rates for the purpose of detecting significant increases which represent epidemics. At the time, NNIS was using a 'threshold program' to identify infection problems in individual hospitals. Under this program, an investigation is instituted when three or more infections of a particular site-by-pathogen combination (SPC) have occurred in a particular calendar month. All the data reported by the hospital since July 1974 for that SPC are used to establish a baseline rate and an expected number of infections. This expected number is set equal to the mean of a Poisson distribution, from which it is determined if the probability of the observed number of infections is less than 0.005. If a large number of hospitals have reported significant increases in the infection rate then NNIS considers the possibility that an epidemic is occurring. This system leaves something to be desired since there is no firm probabilistic underpinning for the decisions to be made.

Chen (1979) stated three requirements for a statistical technique to be useful in surveillance systems. They are as follows. (1) It should be

powerful in detecting increases of rare events. Even under epidemic conditions the event may still be rare. (2) Increases observed in consecutive data collecting periods should be self-reinforcing in leading to a significant result. Thus, a small increase in consecutive time periods should lead to a significant result with higher probability than a similar increase observed in just one time period. (3) The framework should be adaptable to surveillance of the population at risk as a whole or when broken down into subsets.

2.2. Short Description of Competing Models.

Most of the literature in surveillance systems was born out of the quality control problem in continuous industrial production. In the industrial setting the problem is to detect a systematic departure in the quality of the output of a machine from the specified target value. In the public health setting, reliable baseline rates may not be as readily available and the events under study are subject to more sources of variation. Nevertheless, these techniques have been adapted and used.

Three main techniques are currently being used for surveillance systems. These include systems in use in England and Wales (Hill, Spicer and Weatherall 1968; Weatherall and Haskey 1976), Norway (Bjerkedal and Bakketeig 1975) and Israel (Klinberg, Chen, Chemke and Levin 1971). The first two are adaptations from quality control systems and the third was specially devised for health monitoring.

The first technique has many variations but is based on the control chart developed by Shewhart in 1931 for industrial processes. As described in Chapter I, if the number of observed infections exceeds the number expected in a fixed time period by, say, three standard deviations then a

significant increase is declared. One modification is to add the proviso that if the observed exceeds the expected by, say, two standard deviations for, say, three consecutive time periods then this, too, will signal an increase.

The second technique represents an advancement over the control chart in quality control of industrial items. Originally proposed by Page in 1954, it is called the cumulative sum or CUSUM technique. Here, a reference value is subtracted from each group of observations and this difference is added to the sum of the previous groups' differences. For instance, the known baseline rate of a SPC and an acceptable amount of variation in the infection rate are subtracted from the grouped data. The CUSUM continues to build until either an upper boundary is crossed signalling an increase or until a lower boundary is crossed which restarts the sum building process. This is similar in concept to a Sequential Probability Ratio Test (SPRT) except that there is never an acceptance of the hypothesis of no increase in the SPC infection rate since monitoring never ceases.

It is this technique which has received the most attention from statisticians and has been most commonly incorporated in monitoring schemes. The extensions of the original work which include adaptations for nonindependent observations, studies of the optimal reference value to use in construction of the sums, an extension to a two sided test and an investigation of the effects of wrongly estimating nuisance parameters are discussed in Chapter I in general.

The third technique was developed by Chen in 1978 to monitor the occurence of congenital malformations in newborns. Since some health workers were put off by the tedious calculations necessary to carry out the CUSUM technique, the SETS technique was offered as a substitute.

Surveillance does not depend on fixed time intervals to sample the 'output' but rather the occurrence of the event under study defines the end of one time period and beginning of another. The healthy babies born in the time interval between malformed babies form a set. A fixed number of consecutive sets form a sequence. If all sets in a sequence fall below a certain size then an increase is signalled. Each set begins a new sequence so that the data are analyzed frequently and repeatedly. Systems can be operated concurrently on exclusive subsets of the population. This system depends on knowing the exact time of the occurrence of the event so that the set is the correct size. Typically, this is not the case when monitoring infections in large populations from a central source since the data come in already grouped. One modification is to combine all subsets into a system which periodically checks all of the sets which are formed in the relevant time period. This operates like a fixed time technique and if each set is below a certain size then an increase is signalled.

2.3. Definition of Short Memory Scheme.

This dissertation studies a scheme which will detect a sudden increase in the rate of occurrence of rare events and depends on using only the most recently acquired data. It can thus be called a short memory (SM) scheme. The concept is general but it will be convenient to continue to use the terminology of 'cases' and 'epidemics'. The detection procedure involves a comparison of the number of cases occurring in fixed time periods which are mutually independent. For convenience, the cases which occur in a week are considered together. In a typical surveillance scheme the data are grouped and then reported so that building this structure into the scheme is realistic.

The number of cases occurring in week k is distributed as a random variable which follows the Poisson probability distribution.

The number of cases in the current week is compared to the total number of cases in the previous s weeks. The parameter s represents the memory of the scheme. If the non-epidemic condition persists throughout the s+1 weeks which make up a testing interval then the total number of cases should be distributed uniformly among the s+1 weeks. The proportion of cases which occur in the present week should be approximately $1/(s+1)$ with the remaining proportion $s/(s+1)$ of the cases having occurred in the previous s weeks. Let the random variables X and Y represent the number of cases in nonoverlapping time periods and suppose X and Y are distributed as Poisson variates. Cox and Lewis (1966) discuss the situation where two Poisson means are tested for equality in a simple test of hypothesis. The means are from nonoverlapping time periods, which makes the probability distributions independent. By the additivity of the Poisson distribution, if

$$\Pr(X=x)=e^{-\lambda_x}\lambda_x{}^x/x! \text{ and } \Pr(Y=y)=e^{-\lambda_y}\lambda_y{}^y/y!$$

then

$$\Pr(X+Y=x+y)=e^{-(\lambda_x+\lambda_y)}(\lambda_x+\lambda_y)^{x+y}/(x+y)! \ .$$

Then

$$\Pr(Y=y|X+Y=x+y) = \frac{\Pr(Y=y,\ X+Y=x+y)}{\Pr(X+Y=x+y)}$$

$$= \frac{\Pr(Y=y,X=x)}{\Pr(X+Y=x+y)}$$

$$= \frac{(e^{-\lambda_x}\lambda_x{}^x/x!)\ (e^{-\lambda_y}\lambda_y{}^y/y!)}{e^{-(\lambda_x+\lambda_y)}(\lambda_x+\lambda_y)^{x+y}/(x+y)!}$$

$$= \frac{(x+y)!}{y!\ x!} \left(\frac{\lambda_y}{\lambda_x+\lambda_y}\right)^y \left(\frac{\lambda_x}{\lambda_x+\lambda_y}\right)^x .$$

Suppose that $\lambda_x = t_1\lambda$ and $\lambda_y = t_2\lambda$, which is to say that the (independent) random variables X and Y are sums of Poisson random variables with mean λ and differ only in the length of their time periods, t_1 and t_2 respectively. Then, the above probability becomes

$$= \frac{(x+y)!}{y!\ x!} \left(\frac{t_2\lambda}{t_1\lambda+t_2\lambda}\right)^y \left(\frac{t_1\lambda}{t_1\lambda+t_2\lambda}\right)^x$$

$$= \frac{(x+y)!}{y!\ x!} \left(\frac{t_2}{t_1+t_2}\right)^y \left(\frac{t_1}{t_1+t_2}\right)^x .$$

Letting $Y = X_{s+1}$ and $X = \sum_{i=1}^{s} X_i$, the conditional distribution of X_{s+1} given $\sum_{i=1}^{s+1} X_i$ is found by substituting in the equation directly above which yields:

$$\Pr\left(X_{s+1} = m \mid \sum_{i=1}^{s+1} X_i = n\right) = \binom{n}{m} \left(\frac{1}{s+1}\right)^m \left(\frac{s}{s+1}\right)^{n-m} \quad \text{for } 0 \le m \le n$$

If X_{s+1} is relatively large then this is evidence that an increase in the infection rate may have occurred during the last week. Formally, this is written as the rule below. If

$$\sum_{j=x_{s+1}}^{n} \binom{n}{j} \left(\frac{1}{s+1}\right)^j \left(\frac{s}{s+1}\right)^{n-j} \le \alpha$$

where $\sum_{i=1}^{s+1} X_i = n$, then reject the hypothesis that no change has occurred from the baseline rate. The actual level of this test may be smaller than α because of its discrete nature.

Note that the conditional probability distribution of X_{s+1} depends on

the underlying Poisson probability distribution of the X_t's, and not only on the fact that they are independent and identically distributed.

2.4. Markovian Property and Alarm States.

Each test like the one described above that is performed looks back for a length of s weeks and no further. Thus, the SM scheme is Markovian of order s, the length of the memory. The s+1 dimensional sample space of ordered sets generated by this week's and the last s weeks' number of infections defines the state space S with states $S(x_1, x_2, \ldots, x_s, x_{s+1})$ where x_k is the number of cases in week k. If p_{x_k}=Pr(i cases in week k under non-epidemic conditions) then the transition probabilities between states are as follows:

$$\Pr\left[S(x_1, x_2, \ldots, x_s, x_{s+1}) \rightarrow S(x_2, x_3, \ldots, x_{s+1}, x_{s+2})\right] = p_{x_{s+2}} .$$

The limiting probability of state $S(x_1, x_2, \ldots, x_s, x_{s+1})$ is $\prod_{k=1}^{s+1} p_{x_k}$, under non-epidemic conditions.

The state space can be partitioned into two non-overlapping sets W and W^c where

$S(x_1, x_2, \ldots, x_s, x_{s+1}) \in W$ if and only if $(x_1, x_2, \ldots, x_s, x_{s+1})$ causes an alarm, and

$S(x_1, x_2, \ldots, x_s, x_{s+1}) \in W^c$ iff $(x_1, x_2, \ldots, x_s, x_{s+1})$ does not cause an alarm.

Of course, $S(x_1, x_2, \ldots, x_s, x_{s+1}) \in W \Rightarrow S(x_1, x_2, \ldots, x_s, x_{s+1}) \notin W^c$. The elements of the sets W and W^c are described by the (s+1)-dimensional indicator 'array' $\underset{\sim}{A}$ where

$$A_{x_1, x_2, \ldots, x_s, x_{s+1}} = \begin{cases} 1 \text{ if } S(x_1, x_2, \ldots, x_s, x_{s+1}) \in W \\ 0 \text{ if } S(x_1, x_2, \ldots, x_s, x_{s+1}) \notin W \end{cases} .$$

For a given α, $\underset{\sim}{A}$ depends only on s and can be constructed directly from the test given above. The critical value of the test is defined as the smallest value of m such that inequality (2) holds. Each value of $\sum_{i=1}^{s} X_i = n-m$ generates a critical value. Formally define

$$C_s(n-m) = \min\left\{ j \,\middle|\, \sum_{j=x_{s+1}}^{n} \binom{n}{j} \left(\frac{1}{s+1}\right)^j \left(\frac{s}{s+1}\right)^{n-j} \leq \alpha \right\}$$

as the critical value for a scheme with a memory of s where $n-m$ ranges over the values of $\sum_{i=1}^{s} X_i$. Then $\underset{\sim}{A}$ is built from the definition

$$A_{x_1, x_2, \ldots, x_s, x_{s+1}} = \begin{cases} 1 \text{ if } x_{s+1} \geq C_s(x_1, x_2, \ldots, x_s) \\ 0 \text{ if } x_{s+1} < C_s(x_1, x_2, \ldots, x_s) \end{cases} .$$

The problem with this definition is that the level of the test can only be controlled from an upper bound. Specification of the exact level of the test *a priori* is impossible. One way around this dilemma is to randomize the test to be equal to α. In so doing, the elements of the array defined above are no longer interpreted as indicating whether a combination of states signals an alarm or not but as representing the probability that the combination of states will sound an alarm. For most of the elements of $\underset{\sim}{A}$ there is no change. For a single test for a scheme of memory s, $\underset{\sim}{A}$ need only be indexed by $m = \sum_{i=1}^{s} X_i$ and $k = x_{s+1}$. Thus the array for any single test can be represented by a 2-dimensional matrix. This will be called the 'alarm matrix' and the lines of $\underset{\sim}{A}$ which run perpendicular to its diagonals will be called the counter-diagonals of $\underset{\sim}{A}$. The counter-diagonals of the matrix $\underset{\sim}{A}$ represent the family of curves defined by the binomial distribution with parameters $N = \sum_{i=1}^{s+1} X_i$, $p = 1/(s+1)$. Each counter-diagonal will have one element changed

to a value that is between zero and one so that Pr(Test Rejects | N=n)=α. That is, find the value of w_n, $0 \le w_n \le 1$ such that

$$\sum_{j=c+1}^{n} \binom{n}{j}\left(\frac{s}{s+1}\right)^{j}\left(\frac{1}{s+1}\right)^{n-j} + w_n \binom{n}{c}\left(\frac{s}{s+1}\right)^{c}\left(\frac{1}{s+1}\right)^{n-c} = \alpha \ .$$

Define the matrix $\underset{\sim}{A}^{(s)} = [\![\alpha_{ck}^{(s)}]\!]$ where $\alpha_{ck}^{(s)}$ is the probability of sounding an alarm if we observe k cases in the current week given that c cases have been observed in the s weeks of memory, for c=0, 1, 2, ... k=0, 1, 2, ... Then,

$$\alpha_{00}^{(s)} \equiv \alpha$$

$$\alpha_{0k}^{(s)} = \min\left(1,\ \alpha(s+1)^{k}\right) \ .$$

For $m \ge 0$ and $k > 0$, let

$$\phi_{ck}^{(s)} = \frac{\alpha(s+1)^{c+k} - \sum_{j=0}^{k-1}\binom{c+k}{j} s^{c+k-j}}{s^{c}\binom{c+k}{k}}$$

then

$$\alpha_{ck}^{(s)} = \max\left(0,\ \min(1, \phi_{ck}^{(s)})\right).$$

Regardless of what the mean of the underlying Poisson distribution is, Pr(Test Rejects|H_0)=α. That is, each $\alpha_{ck}^{(s)}$ is independent of λ. For n=x+y, see that

$$\sum_{x=0}^{\infty}\sum_{y=0}^{\infty} \Pr(X=x)\ \Pr(Y=y)\ \alpha_{x,y}^{(s)} = \sum_{x=0}^{\infty}\sum_{y=0}^{\infty} \frac{e^{-s\lambda}(s\lambda)^{x}}{x!}\ \frac{e^{-\lambda}\lambda^{y}}{y!}\ \alpha_{x,y}^{(s)}$$

and letting x=k, y=n−k and rearranging the terms yields

$$= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{e^{-(s+1)\lambda} s^k \lambda^n}{k!\,(n-k)!} \alpha_{k,n-k}^{(s)}$$

$$= \sum_{n=0}^{\infty} \frac{e^{-(s+1)\lambda} ((s+1)\lambda)^n}{n!} \sum_{k=0}^{n} \binom{n}{k} \left(\frac{1}{s+1}\right)^{n-k} \left(\frac{s}{s+1}\right)^k \alpha_{k,n-k}^{(s)} .$$

Now, independent of λ and n,

$$\sum_{k=0}^{n} \binom{n}{k} \left(\frac{1}{s+1}\right)^{n-k} \left(\frac{s}{s+1}\right)^k \alpha_{k,n-k}^{(s)} = \alpha, \text{ by definition of } \alpha_{k,n-k}^{(s)}$$

and independent of λ

$$\sum_{n=0}^{\infty} \frac{e^{-(s+1)\lambda} ((s+1)\lambda)^n}{n!} = 1 .$$

Thus, independent of λ,

$$\sum_{n=0}^{\infty} \Pr(N=n)\Pr(\text{Test Rejects}|N=n) = \sum_{n=0}^{\infty} \Pr(N=n)\alpha = \alpha.$$

The expected value of the test statistic is α and the level of the test can be specified in advance independently of λ.

Lehmann (1986) wrote the joint distribution of X and Y in exponential form as

$$\Pr(X=x,\ Y=y) = \frac{e^{-(\lambda_x+\lambda_y)}}{x!y!} \exp\left[y \log \frac{\lambda_y}{\lambda_x} + (x+y) \log\lambda_x\right]$$

and finds the uniformly most powerful (UMP) unbiased test of the hypothesis $\lambda_y=\lambda_x$ against the alternative $\lambda_y>\lambda_x$, using the parameter $\log \frac{\lambda_y}{\lambda_x}$ or equivalently $\frac{\lambda_y}{\lambda_x}$. This test is done using Y conditionally on X+Y. Sufficient statistics for λ_x and λ_y are X and Y, respectively, so it is not suprising that the UMP test of the hypothesis $\lambda_y=\lambda_x$ against the alternative $\lambda_y>\lambda_x$ is done conditionally on X+Y.

2.5. Multiple Alarms in a Continuing Scheme.

The random variable to be studied is the time to the first alarm. This will remain the only random variable of interest in light of the foreseen arrangement of events described below. Incorporating multiple alarms into the scheme is not difficult but their usefulness is questionable. It is assumed that an investigation phase takes place after an alarm is sounded. During this investigation phase the cause of the alarm, if it is a true alarm, is found and corrected. Following this begins the re-collection of memory for a new non-epidemic period. The infection rate may not be the same as it was before the investigation, since intervention can cause a host of changes that may mean a lower non-epidemic infection rate than before. The only requirement is that the scheme begins operating again with a 'clean' memory. Testing is resumed when the memory is collected and in place. Of course, it may prove to be too much of a temptation to sneak a peek during this phase of the process since it is possible to start testing the data on a scheme while its memory is being collected and not dropping old data until the memory is full.

The probability that an investigation is triggered, then, can be viewed as the Type I error of the scheme under non-epidemic conditions and as the power of the scheme under epidemic conditions.

2.6. The Time to the First Alarm Under Non-Epidemic Conditions.

As was stated in the previous section, the random variable of interest is the number of weeks until an alarm occurs, which triggers an investigation. Define R as the integer-valued random variable that counts the number of weeks until the first alarm. The distribution of R, known as the run length

distribution, can be written down in terms of the sequential testing which occurs each week. Recall that

$$p(x_j) = \text{probability of i cases in non-epidemic week j.}$$

Now some definitions. Let

$$\varsigma_s(m,k) = 1-\alpha_{mk}^{(s)} \quad m=0, 1, 2, \ldots, \quad k=0, 1, 2, \ldots$$

which is the probability of not sounding an alarm given k cases this week and m cases over the last s weeks.

Let

$$C_{ks} = \varsigma_s\Big(\sum_{j=k-s}^{k-1} x_j \, , \, x_k\Big) \, p(x_k)$$

which is the probability that week k has x_k cases, and that an alarm is not sounded, given that the memory of s weeks has a total of $\sum_{j=k-s}^{k-1} x_j$ cases.

Let $P_s(R=r)$ be the probability that the time to the first alarm, R, equals r when the memory is s weeks. When the scheme is operating under non-epidemic conditions, this probability can be written as an (r+s)-fold sum of the above probabilities. Then,

$$P_s(R=r) = \sum_{x_{1-s}=0}^{\infty} \cdots \sum_{x_r=0}^{\infty} \Big(\prod_{j=1-s}^{0} p(x_j)\Big)\Big(\prod_{k=1}^{r-1} C_{ks}\Big)[p(x_r)-C_{rs}]$$

where

$j = -s+1, -s+2, \ldots, 0$ index weeks of memory,

$j = 1, 2, \ldots, r-1$ index weeks with tests but no alarm,

$j = r$ indexes the week when the alarm occurred.

Implicit in this notation are two points. First, the subscripts indicate that the dependency connects s variables together. Second, the first test occurs at week 1 regardless of the memory of the scheme. This is more than just

arbitrary indexing of the schemes. The schemes are to be compared from this identical starting point. When schemes are compared under epidemic conditions, it will be assumed that the change in the infection rate occurred at the conclusion of week 0.

The run length distribution is nonincreasing and thus has a single mode at R=1. To prove that $Pr(R=r)>Pr(R=r+1)$ consider that

$$P_s(R=r) = \sum_{x_{1-s}=0}^{\infty} \cdots \sum_{x_r=0}^{\infty} \left(\prod_{j=1-s}^{0} p(x_j)\right)\left(\prod_{k=1}^{r-1} C_{ks}\right)[p(x_r)-C_{rs}]$$

and

$$P_s(R=r+1) = \sum_{x_{1-s}=0}^{\infty} \cdots \sum_{x_r=0}^{\infty} \left(\prod_{j=1-s}^{0} p(x_j)\right)\left(\prod_{k=1}^{r} C_{ks}\right)[p(x_{r+1})-C_{(r+1)s}].$$

$P_s(R=r)$ can be rewritten by renaming the indices to start at week $2-s$ and to end at week $r+1$. Thus,

$$P_s(R=r) = \sum_{x_{2-s}=0}^{\infty} \cdots \sum_{x_{r+1}=0}^{\infty} \left(\prod_{j=2-s}^{1} p(x_j)\right)\left(\prod_{k=2}^{r} C_{ks}\right)[p(x_{r+1})-C_{(r+1)s}].$$

Then,

$$P_s(R=r)-P_s(R=r+1) = \sum_{x_{2-s}=0}^{\infty} \cdots \sum_{x_{r+1}=0}^{\infty} \left(\prod_{j=2-s}^{1} p(x_j)\right)\left(\prod_{k=2}^{r} C_{ks}\right)[p(x_{r+1})-C_{(r+1)s}]$$

$$\times \left(1-\sum_{x_{1-s}=0}^{\infty} C_{1s}\right).$$

And since

$$0 \le \left(1-\sum_{x_{1-s}=0}^{\infty} C_{1k}\right) < 1,$$

and all other terms are non-negative, $P_s(R=r)-P_s(R=r+1)>0$.

The run length distribution can also be written in terms of the probability of the individual tcsts performed each week. The outcome of the test performed at week j is a Bernoulli random variable, T_j, where

$P_s(T_j=1)=\alpha$ and $P_s(T_j=0) = 1-P_s(T_j=1)$. This is the unconditional probability that a test will reject or accept, respectively, the hypothesis that the non-epidemic situation exists. It is unconditional because it does not depend on the outcome of the tests which were performed previously. This formulation will explicitly take advantage of the Markovian property of the SM scheme, as follows. In general,

$$P_s(R=r) = P_s(T_r=1,\ T_{r-1}= \cdots =T_1=0) = P_s\left(T_r=1,\ \left\{T_j=0\right\}_{j=1}^{r-1}\right).$$

Rewriting $P_s(R=r)$ and using the Markov property yields:

for $r \leq s$,

$$= P_s\left(T_r=1 \middle| \left\{T_j=0\right\}_{j=1}^{r-1}\right) \prod_{u=1}^{r-1} P_s\left(T_u=0 \mid \left\{T_j=0\right\}_{j=1}^{r-1}\right)$$

and for $r \geq s+1$

$$= P_s\left(T_{s+t}=1 \middle| \left\{T_j=0\right\}_{j=t}^{s+t-1}\right) \left(\prod_{v=1}^{t-1} P_s\left(T_{s+t-v}=0 \middle| \left\{T_j=0\right\}_{j=t-v}^{s+t-v-1}\right)\right)$$

$$\times \prod_{u=1}^{s} P_s\left(T_u=0 \middle| \left\{T_j=0\right\}_{j=1}^{u-1}\right)$$

Under non-epidemic conditions, let $\theta_0 \equiv 1$ and for $r \geq 1$ let

$$\theta_r \equiv Pr(R>r) = Pr(T_r=0,\ T_{r-1}=0,\ \ldots ,T_1=0).$$

Then,

for $1 \leq r \leq s$ $\quad Pr(R=r) = \theta_1 \dfrac{\theta_2}{\theta_1} \cdots \dfrac{\theta_{r-1}}{\theta_{r-2}} \left(1 - \dfrac{\theta_r}{\theta_{r-1}}\right)$

$$= \theta_{r-1} - \theta_r$$

and for $r \geq s+1$

$$Pr(R=r) = \theta_1 \frac{\theta_2}{\theta_1} \frac{\theta_3}{\theta_2} \cdots \frac{\theta_s}{\theta_{s-1}} \left(\frac{\theta_{s+1}}{\theta_s}\right)^{r-s-1} \left(1 - \frac{\theta_{s+1}}{\theta_s}\right)$$

$$= \theta_s \left(\frac{\theta_{s+1}}{\theta_s}\right)^{r-s-1} \left(1 - \frac{\theta_{s+1}}{\theta_s}\right)$$

$$= (\theta_s - \theta_{s+1}) \left(\frac{\theta_{s+1}}{\theta_s}\right)^{r-s-1}.$$

The distribution of R conditioning on R > s is geometric since

$$\Pr(R=r|R>s) = \left(\frac{\theta_{s+1}}{\theta_s}\right)^{r-s-1} \left(1-\frac{\theta_{s+1}}{\theta_s}\right).$$

The general form of θ_{s+t} for $t \geq 1$ is $\theta_{s+t} = \frac{\theta_{s+1}^t}{\theta_s^{t-1}}$.

Proof:

$t=1$: $$\Pr(R>s+2) = \Pr(R>s+1) - \Pr(R=s+2)$$

so,

$$\theta_{s+2} = \theta_{s+1} - (\theta_s - \theta_{s+1})\left(\frac{\theta_{s+1}}{\theta_s}\right)$$

$$= \frac{\theta_{s+1}^2}{\theta_s}.$$

$P(t) \rightarrow P(t+1)$:

Assume $$\theta_{s+t} = \frac{\theta_{s+1}^t}{\theta_s^{t-1}}$$

then,

$$\theta_{s+t+1} = \frac{\theta_{s+1}^t}{\theta_s^{t-1}} - (\theta_s - \theta_{s+1})\left(\frac{\theta_{s+1}}{\theta_s}\right)^t$$

$$= \frac{\theta_{s+1}^{t+1}}{\theta_s^t}.$$

Thus,

$$\theta_{s+t} = \frac{\theta_{s+1}^t}{\theta_s^{t-1}}, \quad t \geq 1. \qquad \text{QED}$$

The expected value of R can be found easily from the results above.

$$E(R|H_0) = \sum_{r=0}^{\infty} \Pr(R>r|H_0) = \sum_{r=0}^{\infty} \theta_r = \sum_{r=0}^{s} \theta_r + \sum_{r=s+1}^{\infty} \theta_r$$

and

$$\sum_{r=s+1}^{\infty} \theta_r = \sum_{t=1}^{\infty} \frac{\theta_{s+1}^t}{\theta_s^{t-1}}$$

$$= \theta_s \sum_{t=1}^{\infty} \left(\frac{\theta_{s+1}}{\theta_s}\right)^t$$

$$= \theta_s\left[\sum_{t=0}^{\infty}\left(\frac{\theta_{s+1}}{\theta_s}\right)^t - 1\right]$$

$$= \theta_s\left[\frac{\theta_{s+1}}{\theta_s - \theta_{s+1}}\right] .$$

Thus,

$$E(R|H_0) = \sum_{r=0}^{s}\theta_r + \theta_s\left[\frac{\theta_{s+1}}{\theta_s - \theta_{s+1}}\right] .$$

This result can alternatively be established from the relationship which says that

$$E(R|H_0) = Pr(R \leq s)[E(R|R \leq s,H_0)] + Pr(R \geq s+1)[E(R|R \geq s+1,H_0) - (s+1)].$$

First,

$$E(R|R \leq s,H_0) = \sum_{r=1}^{s} r(\theta_{r-1} - \theta_r)/(1-\theta_s)$$

$$= \left[\sum_{r=0}^{s-1}\theta_{s-1} - s\theta_s\right]/(1-\theta_s).$$

Next

$$E(R|R \geq s+1,H_0) = \sum_{r=s+1}^{\infty} r(\theta_{r-1} - \theta_r)/\theta_s$$

$$= \sum_{r=s+1}^{\infty} \theta_r/\theta_s + (s+1)$$

$$= \frac{\theta_{s+1}}{\theta_s - \theta_{s+1}} + (s+1) .$$

Thus,

$$E(R|H_0) = (1-\theta_s)\left[\sum_{r=0}^{s-1}\theta_{s-1} - s\theta_s\right]/(1-\theta_s) + \theta_s\left[\frac{\theta_{s+1}}{\theta_s - \theta_{s+1}} + (s+1)\right]$$

$$= \sum_{r=0}^{s}\theta_r + \theta_s\left[\frac{\theta_{s+1}}{\theta_s - \theta_{s+1}}\right] .$$

The moment generating function of the run length distribution can be found in a similar fashion using the conditional distribution of R. In general,

$$E(e^{tr}) = \sum_{r=0}^{\infty} e^{tr}\Pr(R=r) = \sum_{r=0}^{s} e^{tr}\Pr(R=r) + \sum_{r=s+1}^{\infty} e^{tr}\Pr(R=r),$$

where

$$\sum_{r=s+1}^{\infty} e^{tr}\Pr(R=r) = \theta_s \sum_{r=s+1}^{\infty} e^{tr}\left[1-\frac{\theta_{s+1}}{\theta_s}\right]\left[\frac{\theta_{s+1}}{\theta_s}\right]^{r-(s+1)}$$

$$= \theta_s \frac{(\theta_s-\theta_{s+1})e^{t(s+1)}}{\theta_s-\theta_{s+1}e^t}, \qquad t<-\log\left[\frac{\theta_{s+1}}{\theta_s}\right]$$

and

$$\sum_{r=1}^{s} e^{tr}\Pr(R=r) = \sum_{r=1}^{s-1}\theta_r(e^{t(r+1)}-e^{tr}) + e^t - \theta_s e^{ts}$$

$$= (e^t-1)\sum_{r=1}^{s-1}\theta_r e^{tr} + e^t - \theta_s e^{ts}.$$

Thus,

$$E(e^{tr}) = 1 + (e^t-1)\left[\sum_{r=1}^{s-1}\theta_r e^{tr} + 1\right] + \theta_s e^{ts}\left[\frac{(\theta_s-\theta_{s+1})e^t}{\theta_s-\theta_{s+1}e^t} - 1\right].$$

2.7. Time to the First Alarm under Epidemic Conditions.

The epidemic situation will be assumed to be a one time increase in the mean of the underlying distribution from which cases are generated. Under epidemic conditions the run length distribution can be written down with the additional definitions below. Let

$$p'(x_j) = \text{probability of } i \text{ cases in epidemic week } j.$$

Also, as above,

$$\varsigma_s(m,k) = 1-\alpha_{mk}^{(s)} \qquad m=0, 1, 2, \ldots, \qquad k=0, 1, 2, \ldots$$

which is the probability of not sounding an alarm given k cases this week and m cases over the last s weeks. And let

$$C'_{ks} = \varsigma_s\left[\sum_{j=k-s}^{k-1} x_j, x_k\right] p'(x_k)$$

which is the probability that week k has x_k cases, and that an alarm is not

sounded, given that the memory of s weeks has a total of $\sum_{j=k-s}^{k-1} x_j$ cases are defined as above. When the scheme is operating under epidemic conditions, then,

$$P_s(R=r) = \sum_{x_{1-s}=0}^{\infty} \cdots \sum_{x_r=0}^{\infty} \Big(\prod_{j=1-s}^{0} p(x_j)\Big)\Big(\prod_{k=1}^{r-1} C'_{ks}\Big)[p'(x_r)-C_{rs}]$$

where

$j = -s+1, -s+2, \ldots, 0$ index (non-epidemic) weeks of memory,

$j = 1, 2, \ldots, r-1$ index (epidemic) weeks with tests but no alarm,

$j = r$ indexes the (epidemic) week when the alarm occurred.

The situation under epidemic conditions is slightly different from that under non-epidemic conditions when trying to write the run length distribution in terms of the individual tests. A test still only uses the past s weeks of memory and the joint distribution of $s+1$ tests will be using a total of 2s weeks of data, as before. The difference now is that the joint distributions contain mixtures of epidemic and non-epidemic weeks that do not settle down to a geometric distribution until beyond the $2s^{th}$ week instead of the sth week as in the non-epidemic situation. After the sth week the test is processing only data generated by the alternative distribution and the proportion of cases in the latest week is again expected to be $1/(s+1)$ of the total number of cases in $s+1$ weeks. The joint distribution of the tests extends back s tests and hence the entire set of data is not constant again until after week 2s. Under epidemic conditions, let $\beta_0 \equiv 1$; and for $r \geq 1$ let

$$\beta_r \equiv \Pr(R>r) = \Pr(T_r=0,\ T_{r-1}=0,\ \ldots, T_1=0).$$

Then for $1 \leq r \leq 2s$

$$\Pr(R=r) = \beta_{r-1} - \beta_r$$

and for $r > 2s$

$$\Pr(R=r) = \beta_{2s}\Big(\frac{\beta_{2s+1}}{\beta_{2s}}\Big)^{r-(2s+1)}\Big(1-\frac{\beta_{2s+1}}{\beta_{2s}}\Big)$$

$$= (\beta_{2s} - \beta_{2s+1}) \left(\frac{\beta_{2s+1}}{\beta_{2s}}\right)^{r-(2s+1)}.$$

The expected value of R under epidemic conditions is given by

$$E(R|H_1) = \sum_{r=0}^{\infty} \beta_r$$

$$= \sum_{r=0}^{2s} \beta_r + \beta_{2s}\left\{\frac{\beta_{2s+1}}{\beta_{2s} - \beta_{2s+1}}\right\}.$$

As stated above, from week 2s+1 onward, the underlying distribution of cases is identical from week to week but at the increased 'epidemic' rate. For an γ-fold increase in the baseline rate under epidemic conditions,

$$\Pr(R > 2s+1 \mid R > 2s,\ H_1\colon \lambda_1 = \gamma\lambda_0) = \Pr(R > s+1 \mid R > s,\ H_0\colon \gamma\lambda_0).$$

Let $\theta_r[\lambda_0]$ and $\beta_r[\lambda_1] = \beta_r[\gamma\lambda_0]$ represent the tail probabilities as defined before but with the underlying rate explicitly shown. Then,

$$\frac{\beta_{2s+1}[\gamma\lambda_0]}{\beta_{2s}[\gamma\lambda_0]} = \frac{\theta_{s+1}[\gamma\lambda_0]}{\theta_s[\gamma\lambda_0]}$$

and for $r > 2s$,

$$\Pr(R = r) = \beta_{2s}[\gamma\lambda_0]\left(\frac{\theta_{s+1}[\gamma\lambda_0]}{\theta_s[\gamma\lambda_0]}\right)^{r-(2s+1)}\left(1 - \frac{\theta_{s+1}[\gamma\lambda_0]}{\theta_s[\gamma\lambda_0]}\right).$$

The power of the scheme depends on λ_0 as well as on γ and θ_1. A simple demonstration of this involves looking at the first conditional binomial test. The randomization weights will be ignored for the sake of notational clarity. The first test under non-epidemic conditions is equal to

$$\theta_1 = \sum_{n=0}^{\infty} \Pr(N = n)\Pr(\text{Test Accepts} | N = n)$$

$$= \sum_{n=0}^{\infty} \frac{e^{-\lambda(s+1)}(\lambda(s+1))^n}{n!} \sum_{j=0}^{c-1} \binom{n}{j}\left(\frac{1}{s+1}\right)^j \left(\frac{s}{s+1}\right)^{n-j}.$$

Under epidemic conditions this becomes, for an γ-fold increase in λ,

$$\beta_1 = \sum_{n=0}^{\infty} \frac{e^{-\lambda(s+\gamma)}(\lambda(s+\gamma))^n}{n!} \sum_{j=0}^{c-1}\binom{n}{j}\left(\frac{\gamma}{s+\gamma}\right)^j\left(\frac{s}{s+\gamma}\right)^{n-j}.$$

$$= \sum_{n=0}^{\infty} \frac{e^{-\lambda(s+1)}(\lambda(s+1))^n}{n!}\, e^{-\lambda(\gamma-1)} \left(\frac{s+\gamma}{s+1}\right)^n$$

$$\times \sum_{j=0}^{c-1}\binom{n}{j} \frac{s^{n-j}}{(s+1)^n} \left(\frac{s+1}{s+\gamma}\right)^n \gamma^j.$$

$$= e^{-\lambda(\gamma-1)} \sum_{n=0}^{\infty} \frac{e^{-\lambda(s+1)}(\lambda(s+1))^n}{n!} \sum_{j=0}^{c-1}\binom{n}{j}\left(\frac{1}{s+1}\right)^j\left(\frac{s}{s+1}\right)^{n-j}\gamma^j.$$

Lehmann (1986) showed that for each fixed value of $\frac{1}{s+1}$, the power of an unconditional test, $1-\beta_1$, is an increasing function of λ. The argument starts by considering the power of a conditional test given $x+y$ as an increasing function of $x+y$. This is true because based on $x+y$ binomial trials, it is the power of the optimum test. The conditioning variable $X+Y$ has a Poisson distribution and its distribution forms an exponential family for varying λ. From earlier established results it is then concluded that the overall power, $E(1-\beta_1|X+Y)$, is an increasing function of λ.

2.8. Continuous Approximation as an Alternative to Randomization.

The purpose of fixing the level of the test at exactly α is to make comparisons between tests and schemes more equitable. An alternative to randomization for fixing the test level is to use a continuous random variable as an approximation to the distribution of cases. This eliminates the problem of unequal testing levels since the levels of a continuous test can be made equal to any α. Let X be the number of cases in a week and assume that

$$\tau X \sim \chi^2(\delta)$$

where τ, $\delta > 0$. Then, $E(X)=\delta/\tau$ and $V(X)=2\delta/\tau^2$. The infinite-divisibility property of χ^2 makes it flexible in considering different time periods since if Y is the number of cases in t weeks then $\tau Y \sim \chi^2(t\delta)$. Following the scheme derived above, the number of cases observed, say A, in a period of a weeks is compared to the number of cases observed, say S, in a memory period of s weeks. The alarm is sounded if $A/(A+S)$ is large or equivalently if A/S is large. Since under non-epidemic conditions

$$\tau A \sim \chi^2(a\delta)$$

and

$$\tau S \sim \chi^2(s\delta)$$

then

$$A/S \sim \frac{\chi^2(a\delta)}{\chi^2(s\delta)}$$

and

$$\frac{A}{a}\frac{s}{S} = \frac{A/a\delta}{S/s\delta} \sim \frac{\chi^2(a\delta)/a\delta}{\chi^2(s\delta)/s\delta} = F(a\delta,\ s\delta).$$

For a test that sounds an alarm with a Type I error of α there is a c such that

$$\Pr(A/S > c) = \alpha.$$

And

$$\Pr(sA/aS > sc/a) = \alpha \rightarrow \Pr(F(a\delta,\ s\delta) > sc/a) = \alpha.$$

So, c is found from the inverse $F(a\delta,\ s\delta)$ distribution and the alarm is sounded if $A/S > ac/s$.

Suppose that under epidemic conditions the scale parameter τ in the distribution of A remains the same but that the shape parameter changes to $\Delta = \gamma\delta > \delta$. If the entire period of a weeks occurs under epidemic conditions then $A/\tau \sim \chi^2(a\Delta) = \chi^2(a\gamma\delta)$. Then,

$$\Pr(A/S>c) = \Pr\left\{\frac{A/a\gamma\delta}{S/a\delta} > \frac{s}{a}c\frac{\delta}{\gamma\delta}\right\} = \Pr(F(\gamma a\delta,\ s\delta)>cs/\gamma a).$$

If the period of a weeks has a_1 non-epidemic weeks and a_2 epidemic weeks then $A/\tau \sim \chi^2(a_1\delta + a_2\Delta) = \chi^2(a_1\delta + a_2\gamma\delta)$. Then, letting $a' = a_1 + a_2\gamma$,

$$\Pr(A/S>c) = \Pr\left\{\frac{A/a'}{S/s\delta} > \frac{s}{a'}c\frac{\delta}{\delta}\right\} = \Pr(F(a'\delta,\ s\delta)>cs/a').$$

The run length distribution can be found from the tail probabilities since

$$\Pr(R>r) = \Pr\left\{\frac{X_2}{X_1}<c,\ \frac{X_3}{X_2}<c,\ \ldots,\ \frac{X_{r-1}}{X_{r-2}}<c,\ \frac{X_r}{X_{r-1}}<c\right\}$$

$$= \Pr\left\{cX_1>X_2>\frac{X_3}{c}> \ldots >\frac{X_{r-2}}{c^{r-4}}>\frac{X_{r-1}}{c^{r-3}}>\frac{X_r}{c^{r-2}}\right\}$$

$$= \Pr\left\{X_1>\frac{X_2}{c}> \ldots >\frac{X_{r-2}}{c^{r-3}}>\frac{X_{r-1}}{c^{r-2}}>\frac{X_r}{c^{r-1}}\right\}$$

and

$$\Pr(R=r) = \Pr(R>r-1) - \Pr(R>r)\ .$$

The expression for $\Pr(R>r)$, which is a joint distribution of non-independent identically distributed random variables, is rewritten to expose a joint distribution of independent non-identically distributed random variables. Since the pdf of X_t is known, the above probability can be calculated. Under non-epidemic conditions the probability distribution of τX is

$$f(\tau X) = \frac{e^{-y/2}y^{\delta/2-1}}{2^{\delta/2}\Gamma(\frac{\delta}{2})}$$

and $f(X) = \tau f(\tau X)$. So,

$$\Pr(R>r) = \int\int\cdots\int f(X_1)f\left(\frac{X_2}{c}\right)\ldots f\left(\frac{X_{r-1}}{c^{r-2}}\right)f\left(\frac{X_r}{c^{r-1}}\right)dx_1dx_2\cdots dx_{r-1}dx_r$$

where the r-fold integral is evaluated over the surface defined by $\infty > X_1 > \frac{X_2}{c} > \dots > \frac{X_{r-2}}{c^{r-3}} > \frac{X_{r-1}}{c^{r-2}} > \frac{X_r}{c^{r-1}} > 0$. Using the result that

$$\Pr(X>t) = \int_t^\infty \frac{e^{-x/\tau} x^{\frac{\delta}{2}-1}}{\tau^{\delta/2}\Gamma(\frac{\delta}{2})}\, dx = \sum_{j=0}^{\frac{\delta}{2}-1} \frac{e^{-t/\tau}\, t^j}{\tau^j j!},$$

provided that $\delta/2$ is an integer, $\Pr(R>r)$ can be written down as a product of finite sums. First, let $u_k = \sum_{i=1}^{k-1} j_i + \frac{\delta}{2} - 1$. Then,

$$\Pr(R>r) = \prod_{k=1}^{r} \sum_{j_k=0}^{u_k} \binom{j_k+\frac{\delta}{2}-1}{j_k} \left(\frac{c^k}{h(c^k)}\right)^{\frac{\delta}{2}} \left(1-\frac{c^k}{h(c^k)}\right)^{j_k}$$

where $h(c^k) = c^k + c^{k-1} + \cdots + 1$.

The change in the shape parameter that occurs under epidemic conditions manifests itself as follows. If the change in the parameter from δ to $k\delta$ occurs just after time t, then for X_i, $i>t$ the contribution to the above probability becomes

$$\sum_{j_i=0}^{u_i} \binom{j_i+\frac{k\delta}{2}-1}{j_i} \frac{c^{i\frac{k\delta}{2}}}{h(c^i)^{\frac{k\delta}{2}+j_i}},$$

where $u_i = \sum_{i=1}^{i-1} j_i + \frac{k\delta}{2} - 1$.

Because this continuous approximation to the distribution of number of cases is a multiparameter family it is very flexible. For instance, if $\tau=2$ then $E(X)=V(X)=\delta/2$ and one important characteristic of the Poisson distribution is imitated. However, the use of a continuous approximation still presents a serious problem to the user of a detection scheme, namely, having an alarm sound for a certain number of new cases that involve a fraction of a case. In practice, this is no better a solution to the problem than having

to randomize tests, since, if the number of new cases were less than one case away from the number necessary to sound an alarm, there is a good chance that a randomized test would take place *post haste*. In this event, nothing will have been gained by using a continuous approximation.

Chapter III

Properties of SM Scheme when Memory Length is One Week.

3.1. The Run Length Distribution Under Non-epidemic Conditions.

When the memory length of the SM scheme is equal to one, more properties of the scheme can be described since manipulation is possible that is extremely cumbersome when dealing with longer memory schemes. It is interesting to compare the properties of the run length distribution with those of a geometric distribution since that is the distribution that would result if the individual tests were not correlated or if the tests were independent. The probability distribution for the run length is found by setting $s=1$ and recalling the formula from Chapter II. Dropping s from the notation then,

$$\Pr(R>1)\equiv\theta_1 \text{ and } \Pr(R>2)\equiv\theta_2.$$

Then,

$$\Pr(R=1)=1-\theta_1$$

and for $r>1$

$$\Pr(R=r)= (\theta_1-\theta_2)\left(\frac{\theta_2}{\theta_1}\right)^{r-2}$$

$$= \theta_1\left(1-\frac{\theta_2}{\theta_1}\right)\left(\frac{\theta_2}{\theta_1}\right)^{r-2}$$

$$= \Pr(R>1)\ \Pr(R=r|R>1).$$

The moments of the run length can be found from the moment generating

function given below:

$$E(e^{tr}) = (1-\theta_1)e^t + \theta_1\frac{(\theta_1-\theta_2)e^{2t}}{\theta_1-\theta_2 e^t} .$$

Thus,

$$E(R|H_0) = 1 + \theta_1 + \frac{\theta_1\theta_2}{\theta_1-\theta_2}$$

$$= 1 + \frac{\theta_1^2}{\theta_1-\theta_2}$$

and

$$V(R|H_0) = (1-\theta_1)+\theta_1\left[\frac{2\theta_1+\theta_2+2\theta_2^2}{(\theta_1-\theta_2)}\right] - E^2(R|H_0)$$

$$= (1-\theta_1)+\left[\frac{2\theta_1^2+\theta_1(\theta_2+2\theta_2^2)}{(\theta_1-\theta_2)}\right] - \left[1+\frac{2\theta_1^2}{\theta_1-\theta_2}+\frac{\theta_1^4}{(\theta_1-\theta_2)^2}\right]$$

$$= \frac{\theta_1}{\theta_1-\theta_2}\left[2\theta_2(1+\theta_2) - \theta_1 - \frac{\theta_1^3}{\theta_1-\theta_2}\right] .$$

In Chapter II it was proven that in the SM scheme it is always true that

$$Pr(R=1) > Pr(R=2)$$

which says $\quad 1-\theta_1 > \theta_1-\theta_2 ,$

so $\quad \theta_2 > 2\theta_1-1 .$

If the tests which make up a scheme were uncorrelated then it would be that

$$Pr(R=2) = Pr(R\neq 1)\ Pr(R=1)$$

$$= (1-Pr(R=1))\ Pr(R=1).$$

Thus,

$$\theta_1-\theta_2 = \theta_1\ (1-\theta_1) ,$$

so $\quad \theta_2 = \theta_1^2 .$

It will next be proven that in the SM scheme, $\theta_1^2 \geq \theta_2$ for $s=1$. Thus, $\theta_2 \in (2\theta_1-1,\ \theta_1^2)$.

The proof that $\theta_1^2>\theta_2$ is as follows. Let $p_i=p_i(\lambda_0)$ be the underlying probability of seeing i cases in a non-epidemic week. This distribution depends on λ_0. Then,

$$\Pr(R>1) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} p_i\, C_{js} = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} p_i p_j (1-\alpha_{ij})$$

and

$$[\Pr(R>1)]^2 = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} p_i p_j p_k p_l (1-\alpha_{ij})(1-\alpha_{kl}).$$

Similarly,

$$\Pr(R>2) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} p_i p_j p_k (1-\alpha_{ij})(1-\alpha_{jk})$$

$$= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{l=0}^{\infty} p_i p_j p_l (1-\alpha_{ij})(1-\alpha_{jl})$$

$$= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{l=0}^{\infty} p_i p_j p_l (1-\alpha_{ij})(1-\alpha_{jl}) \sum_{k=0}^{\infty} p_k.$$

Now,

$$\theta_1^2-\theta_2 = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} p_i p_j p_k p_l (1-\alpha_{ij})\big[(1-\alpha_{kl})-(1-\alpha_{jl})\big]$$

$$= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} p_i p_j p_k p_l (1-\alpha_{ij})(\alpha_{jl}-\alpha_{kl})$$

$$= \sum_{i=0}^{\infty}\sum_{l=0}^{\infty}\Big[\sum_{j<k}\sum + \sum_{j=k}\sum + \sum_{j>k}\sum\Big] p_i p_j p_k p_l (1-\alpha_{ij})(\alpha_{jl}-\alpha_{kl}).$$

There are three cases to consider.

(i) $j=k \rightarrow \alpha_{jl}=\alpha_{kl} \rightarrow \alpha_{jl}-\alpha_{kl}=0$.

(ii) $j<k \rightarrow \alpha_{jl}\geq\alpha_{kl} \rightarrow \alpha_{jl}-\alpha_{kl}\geq 0$.

(iii) $j>k \rightarrow \alpha_{jl}\leq\alpha_{kl} \rightarrow \alpha_{jl}-\alpha_{kl}\leq 0$.

For case (iii), rename the indices as follows: $j \rightarrow k$ and $k \rightarrow j$. Then

$$\text{(iii) } k>j \rightarrow \alpha_{kt} \leq \alpha_{jt} \rightarrow \alpha_{kt}-\alpha_{jt} \leq 0 ,$$

and

$$\theta_1^2-\theta_2 = \sum_{t=0}^{\infty} \sum_{t=0}^{\infty} \sum\sum_{j<k} p_t p_j p_k p_t \Big[(1-\alpha_{tj})(\alpha_{jt}-\alpha_{kt})+(1-\alpha_{tk})(\alpha_{kt}-\alpha_{jt})\Big]$$

$$= \sum_{t=0}^{\infty} \sum_{t=0}^{\infty} \sum\sum_{j<k} p_t p_j p_k p_t (\alpha_{jt}-\alpha_{kt})\Big[(1-\alpha_{tj})-(1-\alpha_{tk})\Big]$$

$$= \sum_{t=0}^{\infty} \sum_{t=0}^{\infty} \sum\sum_{j<k} p_t p_j p_k p_t (\alpha_{jt}-\alpha_{kt})(\alpha_{tk}-\alpha_{tj}).$$

Since $j<k \rightarrow \alpha_{jt}-\alpha_{kt} \geq 0$ and $\alpha_{tk}-\alpha_{tj} \geq 0$ and there is at least one inequality, it follows that $\theta_1^2-\theta_2>0$. Q.E.D.

The bounds on θ_2 lead to lower and upper bounds on $E(R|H_0)$. First the range of θ_1 will be found. It must be that $\theta_1<1$ since otherwise the scheme will never reject. So, for $\theta_1<1$ and $2\theta_1-1<\theta_2<\theta_1^2$,

$$E(R|H_0)=1+\frac{\theta_1^2}{\theta_1-\theta_2} > 1+\frac{\theta_1^2}{\theta_1-(2\theta_1-1)}$$

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

And

$$E(R|H_0)=1+\frac{\theta_1^2}{\theta_1-\theta_2} < 1+\frac{\theta_1^2}{\theta_1-\theta_1^2}$$

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

Thus, $E(R|H_0) \in \left(\frac{1}{1-\theta_1}-\theta_1, \frac{1}{1-\theta_1}\right)$. The expected value of R under non-epidemic conditions for the SM scheme is less than what would result if the tests performed each week were uncorrelated but it is no more than θ_1 less.

3.2. The Run Length Distribution under Epidemic Conditions.

Under epidemic conditions, as shown in Chapter II, the run length distribution does not settle down to a geometric distribution until further out in the tail (conditional on $R>2s$) than when operating under non-epidemic conditions (conditional on $R>s$). When $s=1$, the run length distribution has this conditional geometric 'tail' from the third week on. Recall from Chapter II that

$$\Pr(R>1|H_1)\equiv\beta_1,\ \Pr(R>2|H_1)\equiv\beta_2 \text{ and } \Pr(R>3|H_1)\equiv\beta_3\ .$$

Then,

$$\Pr(R=1) = 1-\beta_1$$

and

$$\Pr(R=2) = \beta_1-\beta_2\ .$$

For $r>2$,

$$\Pr(R=r) = \beta_2\left(1-\frac{\beta_3}{\beta_2}\right)\left(\frac{\beta_3}{\beta_2}\right)^{r-3}$$

$$= \beta_2\left(1-\frac{\theta_2[\lambda_1]}{\theta_1[\lambda_1]}\right)\left(\frac{\theta_2[\lambda_1]}{\theta_1[\lambda_1]}\right)^{r-3}.$$

Since $\Pr(R>1)$ is set *a priori* to be θ_1 then it is the case that $\theta_1[\lambda_1]=\theta_1[\lambda_0]=\theta_1$. Then for $r>2$,

$$\Pr(R=r) = \beta_2\left(1-\frac{\theta_2[\lambda_1]}{\theta_1}\right)\left(\frac{\theta_2[\lambda_1]}{\theta_1}\right)^{r-3}.$$

The first test compares one week of epidemic level data with one week of non-epidemic level data and the second test compares two weeks of epidemic level data for the first time. Thus, if the tests which make up a scheme were uncorrelated then it would be that

$$\Pr(R=2) = \Pr(R\neq 1|H_1)\ \Pr(R=1|H_0)$$
$$= (1-\Pr(R=1|H_1))\ \Pr(R=1|H_0).$$

The first test has probability equal to the power of a single test. In the second test it is expected that the number of cases be split evenly between the two weeks in spite of the circumstance that more cases are expected in each of the weeks. If the tests were uncorrelated then,

$$\beta_1-\beta_2 = \beta_1(1-\theta_1)\ ,$$

so

$$\beta_2 = \beta_1\theta_1\ .$$

In a manner similar to the non-epidemic case, it can be proven that $\beta_1\theta_1 > \beta_2$ in the SM scheme for $s=1$. The proof is set up as follows. Let $p_i'=p_i(\lambda_1)$ be the underlying probability of i cases in a week under epidemic conditions. Then,

$$\Pr(R>1|H_1) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} p_i\ C'_{js} = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} p_i p_j'(1-\alpha_{ij})$$

and

$$\Pr(R>1|H_1)\ \Pr(R>1|H_0) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} p_i p_j' p_k p_l(1-\alpha_{ij})(1-\alpha_{kl}),$$

and because θ_1 is independent of λ, which implies that $\theta_1[\gamma\lambda_0]=\theta_1[\lambda_0]=\theta_1$, then

$$=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} p_i p_j' p_k' p_l'(1-\alpha_{ij})(1-\alpha_{kl})\ .$$

Similarly,

$$\Pr(R>2|H_1) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} p_i p_j' p_k'(1-\alpha_{ij})(1-\alpha_{jk})$$

$$=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{l=0}^{\infty} p_i p_j' p_l'(1-\alpha_{ij})(1-\alpha_{jl})$$

$$=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{l=0}^{\infty} p_i p_j' p_l'(1-\alpha_{ij})(1-\alpha_{jl})\sum_{k=0}^{\infty} p_k'.$$

Now,

$$\beta_1\theta_1-\beta_2 = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} p_i p'_j p'_k p'_l (1-\alpha_{ij})\left[(1-\alpha_{kl})-(1-\alpha_{jl})\right]$$

$$= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} p_i p'_j p'_k p'_l (1-\alpha_{ij})(\alpha_{jl}-\alpha_{kl}) \ .$$

The proof that $\beta_1\theta_1>\beta_2$ now follows the same argument used for proving that $\theta_1^2>\theta_2$.

Also, similar to the non-epidemic situation, $\Pr(R=1|H_1)>\Pr(R=2|H_1)$. To see this, consider

$$\Pr(R=1|H_1) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} p_i p'_j \alpha_{ij}$$

and

$$\Pr(R=2|H_1) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} p_i p'_j p'_k (1-\alpha_{ij})\alpha_{jk} \ .$$

Now,

$$\Pr(R=1|H_1)-\Pr(R=2|H_1) = \sum_{i=0}^{\infty} p_i \left[\sum_{j=0}^{\infty} p'_j \alpha_{ij} - \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} p'_j p'_k (1-\alpha_{ij})\alpha_{jk}\right]$$

$$= \sum_{i=0}^{\infty} p_i \left[\sum_{j=0}^{\infty} p'_j \alpha_{ij} - \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} p'_j p'_k (\alpha_{jk}-\alpha_{ij}\alpha_{jk})\right]$$

$$= \sum_{i=0}^{\infty} p_i \sum_{j=0}^{\infty} p'_j \alpha_{ij} - \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} p'_j p'_k \alpha_{jk} \sum_{i=0}^{\infty} p_i$$

$$+ \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} p_i p'_j p'_k \alpha_{ij}\alpha_{jk} \ .$$

Now, using the fact that θ_1 is independent of λ so that $\theta_1[\lambda_1]=\theta_1=\theta_1$,

$$\sum_{i=0}^{\infty} p_i \sum_{j=0}^{\infty} p'_j \alpha_{ij} - \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} p'_j p'_k \alpha_{jk} = (1-\beta_1)-(1-\theta_1) \geq 0$$

and

$$\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} p_i p'_j p'_k \alpha_{ij} \alpha_{jk} > 0.$$

Thus, $Pr(R=1|H_1) > Pr(R=2|H_1)$. Q.E.D.

Since $1-\beta_1 \equiv Pr(R=1|H_1) > Pr(R=2|H_1) \equiv \beta_1 - \beta_2$ then, $\beta_2 \in \left(2\beta_1 - 1,\ \beta_1\theta_1\right)$.

The expected value of R under epidemic conditions is

$$E(R|H_1) = 1+\beta_1+\beta_2+\beta_2\left(\frac{\beta_3}{\beta_2-\beta_3}\right)$$

$$= 1+\beta_1+\left(\frac{\beta_2^2}{\beta_2-\beta_3}\right).$$

Since $\frac{\beta_3}{\beta_2} = \frac{\theta_2[\lambda_1]}{\theta_1}$ and $\frac{2\theta_1-1}{\theta_1} < \frac{\theta_2[\lambda_1]}{\theta_1} = \frac{\beta_3}{\beta_2} < \theta_1^2/\theta_1$, then

$$\beta_2 \frac{2\theta_1-1}{\theta_1} < \beta_3 < \beta_2\theta_1.$$

Thus,

$$\frac{\theta_1\beta_2}{1-\theta_1} < \frac{\beta_2^2}{\beta_2-\beta_3} < \frac{\beta_2}{1-\theta_1}.$$

Thus, $E(R|H_1)$ is bounded with $E(R|H_1) \in \left[1+\beta_1+\frac{\beta_2}{1-\theta_1}-\beta_2,\ 1+\beta_1+\frac{\beta_2}{1-\theta_1}\right]$.

Thus, the width of the interval that contains $E(R|H_1)$ is β_2. Since $\theta_1 > \beta_2$, this is a narrower interval than the interval that contains $E(R|H_0)$ but involves three parameters. Substituting the upper and lower bounds for β_2 yields bounds on $E(R|H_1)$ that are slightly wider but depend on only two parameters. Thus, an interval which has width $\frac{\theta_1(1-\beta_1)}{1-\theta_1}$ is

$E(R|H_1) \in \left[1+\frac{\beta_1-\theta_1(1-\beta_1)}{1-\theta_1},\ 1+\frac{\beta_1}{1-\theta_1}\right]$. The upper bound of this interval is

less than the upper bound of $\frac{1}{1-\theta_1}$ for $E(R|H_0)$. To see this, note that $\theta_1 > \beta_1$ so that $1-\theta_1+\beta_1 < 1$.

3.3. The Run Length Distribution Using Independent Tests.

One of the ways to use a given length of memory is to perform a conditional binomial test every $s+1$ weeks. There is no repeat use of the data so these tests are mutually independent. Since this is a somewhat extreme use of the memory it is interesting to compare the expected run length distribution of this scheme with what the SM scheme yields. The independent scheme will be abbreviated, when useful, as the I scheme.

For s=1, independent tests can be done every two weeks. The test weeks are either the 'odd' weeks, 1, 3, 5, ... , or the 'even' weeks, 2, 4, 6, ... , and the two situations are equally likely. In the same period of time there are, at most, only half as many independent tests done as there are tests done by the SM scheme. Define the probability of rejection of the first test as

$$\alpha_I \equiv \Pr(\text{Test Rejects}|H_0, I) = \Pr(T_1 = 1|H_0, I).$$

Then since each of the sub-schemes, where testing is done on odd weeks or is done on even weeks, has a 1/2 probability of occurring it must be that

$$\begin{aligned} &\Pr(\text{Test performed and rejects}) \\ &= \Pr(\text{Test performed})\ \Pr(\text{Test rejects}|\text{Test performed}) \\ &= \frac{1}{2}\alpha_I \end{aligned}$$

For comparison purposes, it is desirable to set $E(R|H_0,SM)$ and $E(R|H_0,I)$ equal so α_I will be chosen to accomplish this.

The run length distribution can be found from the random variable representing the number of tests until the first alarm and then using the

appropriate transformation. Let R_1 be the random variable that represents the run length when the tests are done on odd weeks. Then, let T be the random variable which represents the number of independent tests until the first alarm. Then, $R_1 = 2T - 1$ so that

$$\begin{aligned} \Pr(R_1 = r) &= \Pr(R = 2t - 1) \\ &= \alpha_I (1 - \alpha_I)^{t-1} \qquad t = 1, 2, \ldots \ . \end{aligned}$$

Let R_2 be defined similarly for when the tests are done on even weeks. Then, since $R_2 = 2T$,

$$\begin{aligned} \Pr(R_2 = r) &= \Pr(R = 2t) \\ &= \alpha_I (1 - \alpha_I)^{t-1} \qquad t = 1, 2, \ldots \end{aligned}$$

Now,

$$\begin{aligned} E(R_1|H_0) &= \sum_{t=1}^{\infty} (2t - 1)\left[\alpha_I (1 - \alpha_I)^{t-1}\right] \\ &= \sum_{t=1}^{\infty} 2t\left[\alpha_I (1 - \alpha_I)^{t-1}\right] - 1 \\ &= \frac{2}{\alpha_I} - 1 \end{aligned}$$

and

$$\begin{aligned} E(R_2|H_0) &= \sum_{t=1}^{\infty} 2t\left[\alpha_I (1 - \alpha_I)^{t-1}\right] \\ &= \frac{2}{\alpha_I} \end{aligned}$$

Thus,

$$\begin{aligned} E(R|H_0) &= \frac{E(R_1) + E(R_2)}{2} \\ &= \frac{4 - \alpha_I}{2\alpha_I} \ . \end{aligned}$$

The expected run length from the SM scheme is in the open interval

$$\left[\frac{1}{1-\theta_1}-\theta_1\ ,\ \frac{1}{1-\theta_1}\right] = \left[\frac{1}{\alpha_{SM}}-(1-\alpha_{SM})\ ,\ \frac{1}{\alpha_{SM}}\right]$$

which has length $\theta_1=(1-\alpha_{SM})$. Thus, it must be that $\alpha_I \doteq 2\alpha_{SM}$ and it is not far different from setting $\Pr(T_1=1|H_0,SM) = \Pr(T_1=1|H_0,I)$. In fact, if $E(R|H_0,SM)$ and $E(R|H_0,I)$ are set equal then the bounds given above are also bounds for the expected run length for the independent testing scheme. Thus,

$$\frac{1}{\alpha_{SM}}-(1-\alpha_{SM}) < \frac{4-\alpha_I}{2\alpha_I} < \frac{1}{\alpha_{SM}}\ .$$

Solving for α_I yields

$$\frac{4\alpha_{SM}}{2+\alpha_{SM}} < \alpha_I < \frac{4\alpha_{SM}}{2-\alpha_{SM}+2\alpha_{SM}^2}\ .$$

For example, if $\alpha_{SM}=0.05$ then

$$\frac{.20}{2.05} < \alpha_I < \frac{.20}{1.955}$$

so

$$0.0976 < \alpha_I < 0.1023\ .$$

The proper α_I for setting $E(R|H_0,I)$ equal to a given $E(R|H_0,SM)$ can be found by solving

$$E(R|H_0,SM) = \frac{4-\alpha_I}{2\alpha_I}$$

for α_I, which yields,

$$\alpha_I = \frac{4}{2\ E(R|H_0,SM) + 1}\ .$$

Under epidemic conditions, the run length distribution for tests on the odd weeks depends on one new definition. Let $\beta_I \equiv \Pr(T=0|H_1,I)$. Then

$$\Pr(R_1=1|H_1,I) = \frac{1}{2}(1-\beta_I)$$

and

$$\Pr(R_1=r|H_1,I) = \Pr(R_1=2t-1|H_1,I)$$
$$= \beta_I\ \alpha_I(1-\alpha_I)^{t-2}\ , \qquad t=2,\ 3,\ 4,\ \ldots\ .$$

The expected value is then found from

$$E(R_1|H_1) = (1-\beta_I) + \sum_{t=2}^{\infty}(2t-1)\ \beta_I\ \alpha_I(1-\alpha_I)^{t-2}$$

$$= (1-\beta_I) + 2\beta_I \sum_{t=2}^{\infty} t\alpha_I(1-\alpha_I)^{t-2} - \beta_I$$

$$= 1 + 2\beta_I \left[\frac{1-\alpha_I}{\alpha_I}+2-1\right]$$

$$= 1 + \frac{2\beta_I}{\alpha_I}\ .$$

The run length distribution for when the epidemic starts at week two is given by

$$\Pr(R_2=r|H_1,I) = \Pr(R_2=2t|H_1,I)$$
$$= \alpha_I\ (1-\alpha_I)^{t-1}\ ,\ t=1,2,3,\ \ldots$$

This is the run length distribution under non-epidemic conditions. When the epidemic starts at week 0 then by the time the test in week 2 is performed, the number of cases in both the memory and the current week are generated from the same underlying probability distribution. Although the rate has increased, the cases are expected to be equally distributed between the memory week and the comparison week. Because the expected value of the test statistic is independent of λ, this is just the distribution under non-epidemic conditions, so,

$$E(R_2|H_1) = E(R_2|H_0) = \frac{2}{\alpha_I}\ .$$

Then,

$$E(R|H_1,I) = \frac{E(R_1|H_1)+E(R_2|H_1)}{2}$$

$$= \frac{1}{2} + \frac{1+\beta_I}{\alpha_I}$$

It was shown above that $\alpha_I \doteq 2\alpha_{SM}$ so $E(R|H_0,I)$ can be rewritten as

$$E(R|H_1,I) \doteq \frac{1}{2} + \frac{(1+\beta_I)/2}{\alpha_{SM}} .$$

Under the SM scheme, the expected run length under epidemic conditions is $1+\beta_1+\beta_2+\beta_2\left[\frac{\beta_3}{\beta_2-\beta_3}\right] = 1+\beta_1+\frac{\beta_2^2}{\beta_2-\beta_3}$. The third term of this expected value was shown to be bounded as follows:

$$\frac{(1-\alpha_{SM})(2\beta_1-1)}{\alpha_{SM}} < \frac{(1-\alpha_{SM})\beta_2}{\alpha_{SM}} < \frac{\beta_2^2}{\beta_2-\beta_3} < \frac{\beta_2}{\alpha_{SM}} < \frac{(1-\alpha_{SM})\beta_1}{\alpha_{SM}}.$$

In comparing the expected run lengths of the SM and the I schemes, it is clear that $(1+\beta_1)-\frac{1}{2} \in \left[\frac{1}{2}, \frac{3}{2}-\alpha_{SM}\right]$. The dominant terms in the expected values are the terms involving $\frac{1}{\alpha_{SM}}$. Thus, if it is true that

$$(1-\alpha_{SM})\beta_2 < \frac{(1+\beta_I)}{2} < \beta_2$$

then since $\frac{\beta_2}{\alpha_{SM}} - \frac{\beta_2(1-\alpha_{SM})}{\alpha_{SM}} = \beta_2$, the SM and the Independent testing schemes will have run lengths not more than $\frac{1}{2}+\beta_1+\beta_2$ apart under epidemic conditions. The relationship between the two schemes' expected run lengths depends upon the relationships between β_I, α_{SM} and β_2. These can take values that result in different relationships between the expected run lengths.

However, two important orderings can be established. If $\beta_2 \leq 1/2$ then $(1+\beta_I)/2 > \beta_2$. That is, if the term in the expected run length for independent testing exceeds the term in the upper bound for the expected run length of the SM scheme then it must be that $(1+\beta_I)/2 > \beta_2$. Consider that

$(1+\beta_I)/2 > \beta_2$ means that $\beta_I > 2\beta_2-1$. This has to be true if $2\beta_2-1 \leq 0$. Thus, as long as $\beta_2 \leq 1/2$ then the condition is satisfied since $\beta_I \geq 0$. This restriction on β_2, that $2\beta_2-1 < 0$ or $\beta_2 \leq 1/2$, defines a desirable situation, namely, when the SM scheme has a decent amount of power.

Similarly, if $\alpha_{SM} \leq 1/2$ then $(1-\alpha_{SM})\beta_2 < (1+\beta_I)/2$. That is, the term for the expected run length for independent testing is bounded from below by $(1-\alpha_{SM})\beta_2$ as long as $\alpha_{SM} < 1/2$. Suppose that the term in the expected run length for the independent testing is smaller than the term in the lower bound for the expected run length of the SM scheme. Then $(1+\beta_I)/2 < (1-\alpha_{SM})\beta_2$ which says that $\beta_I < 2(1-\alpha_{SM})\beta_2-1$. This is impossible when $\beta_2 < \frac{1}{2(1-\alpha_{SM})}$ since then $2(1-\alpha_{SM})\beta_2-1 < 0$ so that $\beta_I < 0$. It is the case that $\beta_2 < \frac{1}{2(1-\alpha_{SM})}$ whenever $\alpha_{SM} \leq 1/2$ since this says that $\beta_2 < 1$. So, for $\alpha_{SM} \leq 1/2$ then $(1-\alpha_{SM})\beta_2 < (1+\beta_I)/2$. Thus, there are situations which are desirable and reasonably unrestrictive under which the independent testing scheme will have a larger expected run length than does the SM scheme.

The relationship given above for exceeding the term in the upper bound can be exploited. To find the conditions necessary for the term for the independent testing scheme to exceed the term for the SM scheme by a factor of, say, f then use the fact that this is true when $\frac{1+\beta_I}{2} > f\beta_2$. That is, when $\beta_I > 2f\beta_2-1$ which is always true for $\beta_2 \leq \frac{1}{2f}$.

3.4. Effects of Ignoring the Randomization Weights.

The tests that make up the SM scheme and the independent testing scheme are randomized and can therefore be set to give any specified error rate. This can not otherwise be achieved when the underlying probability

distribution is discrete. The effect of ignoring some or all of the weights associated with the randomization can be studied in specific circumstances. Three rules will be used to show the effect of using different randomization requirements that range from fully randomized tests to fully unrandomized tests. They are as follows:

1) Tests are randomized to specified α.

2) Tests will not reject by randomization when there are zero cases in the new week.

3) Tests are not randomized at all.

The first rule has been discussed above and is full randomization. The second rule only affects the probability of an alarm when each of the two weeks has zero cases and thus the n for the binomial test is zero. This is the only situation where zero cases can sound an alarm under randomization. This rule dictates dropping only one weight but it avoids a situation that would disturb many users of a scheme, namely, sounding an alarm when no cases are observed. The third rule corresponds to when the tests are not randomized at all and by choice the nominal level of an independent test is always greater than or equal to the actual level. That is, the unconditional probability of rejection of a test is conservative.

For the sake of a specific example, the underlying probability distribution of cases is assumed to be Poisson with $\lambda_0 = 1.0$ with a two-, a four-, and a five-fold increase under epidemic conditions. The nominal error rate is .05 so that

$$P_1(R=1|H_0, SM) = .05 \text{ and } Pr(T=1|H_0, I) = \frac{4}{2\ E(R|H_0, SM)+1}.$$

Choosing such a small number for the Poisson parameter ensures that the event of seeing zero number of cases will occur with high probability and the

effect of rule 2 will be evident.

For the SM scheme the effects on the error rates and the expected run lengths under non-epidemic conditions and an epidemic condition represented by a two-fold increase in λ_0 are shown in Table 3.1 below:

Table 3.1

Error Rates and Expected Run Lengths for SM Scheme

$\lambda_0=1$, $\gamma=2$ and $\alpha_{SM}=0.05$

Rule	θ_1	θ_2	β_1	β_2	β_3	$E(R\|H_0)$	$E(R\|H_1)$
1	0.950	0.901	0.880	0.832	0.788	19.4	17.7
2	0.957	0.914	0.883	0.834	0.791	22.2	18.0
3	0.999	0.997	0.979	0.970	0.962	725.1	116.2

By not allowing the scheme to sound an alarm based on zero cases both weeks, the expected run length is increased by 2.8 weeks under non-epidemic conditions but only by 0.3 weeks under epidemic conditions. This is because under epidemic conditions the probability of seeing zero cases is reduced. The effect of easing all of the randomization requirement is seen to be quite severe. The effect of the third rule is to force the scheme to operate with $1-\theta_1$ about 36 times smaller than under Rule 1 and the expected run length is increased by about a factor of 37 under non-epidemic conditions and by about a factor of 6.5 under epidemic conditions.

For the independent testing scheme, the effects of easing the randomization requirements for the error rates are also quite evident. When the tests are fully randomized as in Rule 1 or partially randomized as in Rule 2 then $2\alpha_{SM} \doteq \alpha_I$. This relationship does not hold under Rule 3 and will not

hold in general for partial randomization. The non-zero probability of sounding an alarm when making a transition from zero cases to zero cases has a symmetry under rules 1 and 2 that will not hold for all transitions in general. For the zero cases to zero cases transition, the probability of sounding an alarm is α_{SM} for the SM scheme and α_I for the independent testing scheme. Thus, the inclusion or exclusion of these randomization weights does not change the relationship between α_{SM} and α_I. The expected run lengths are presented below in Table 3.2 below for the case where $\lambda_0=1$ and $\lambda_1=2$.

Table 3.2

Error Rates and Expected Run Lengths for Independent Scheme

$\lambda_0=1$ and $\gamma=2$

Rule	α_I	$1-\alpha_I$	β_I	$E(R\|H_0)$	$E(R\|H_1)$
1	0.101	0.899	0.782	19.4	18.2
2	0.087	0.913	0.787	22.5	21.1
3	0.007	0.997	0.940	276.5	269.3

For rule 3, under non-epidemic conditions, the independent testing scheme does not have as large an expected run length as the SM scheme does. For this example, Plan 3 is more impervious than Plan 1 to the lack of randomization. However, under epidemic conditions, Plan 3 has a much larger expected run length than Plan 1. Thus, it seems that the lack of randomization can lead to a situation where one test which has a smaller Type I error than another test can also have a smaller Type II error than

that test. Thus, the SM scheme is preferred over the Independent scheme for this example.

It is interesting to note that under epidemic conditions, the SM scheme has a shorter expected run length than the independent testing scheme even though $\beta_2 >> 1/2$. Recall that $\beta_2 < 1/2$ is a condition under which the upper bound for $E(R|H_1,SM)$ would likely to be exceeded by $E(R|H_1,I)$ under Rule 1. The upper and lower bounds for $E(R|H_1,SM)$ are 18.6 and 16.3, respectively, under Rule 1.

The results for a four-fold and for a five-fold increase in the rate are shown below in Tables 3.3 and 3.5 for the SM scheme and in Tables 3.4 and 3.6 for the independent testing scheme. The description of these results parallels the discussion given above.

Table 3.3

Error Rates and Expected Run Lengths for SM Scheme

$\lambda_0 = 1$, $\gamma = 4$ and $\alpha_{SM} = 0.05$

Rule	θ_1	θ_2	β_1	β_2	β_3	$E(R\|H_0)$	$E(R\|H_1)$
1	0.950	0.901	0.649	0.599	0.567	19.4	12.8
2	0.957	0.914	0.646	0.599	0.567	22.2	12.8
3	0.999	0.997	0.818	0.798	0.782	725.1	41.1

Table 3.4

Error Rates and Expected Run Lengths for Independent Scheme

$\lambda_0=1$ and $\gamma=4$

Rule	α_I	$1-\alpha_I$	β_I	$E(R\|H_0)$	$E(R\|H_1)$
1	0.101	0.899	0.489	19.4	15.3
2	0.087	0.913	0.489	22.5	17.6
3	0.007	0.993	0.691	276.5	234.7

Table 3.5

Error Rates and Expected Run Lengths for SM Scheme

$\lambda_0=1$, $\gamma=5$ and $\alpha_{SM}=0.05$

Rule	θ_1	θ_2	β_1	β_2	β_3	$E(R\|H_0)$	$E(R\|H_1)$
1	0.950	0.901	0.514	0.469	0.444	19.4	10.1
2	0.957	0.914	0.514	0.469	0.444	22.2	10.1
3	0.999	0.997	0.692	0.669	0.653	725.1	29.4

Table 3.6

Error Rates and Expected Run Lengths for Independent Scheme

$\lambda_0=1$ and $\gamma=5$

Rule	α_I	$1-\alpha_I$	β_I	$E(R\|H_0)$	$E(R\|H_1)$
1	0.101	0.899	0.358	19.4	14.0
2	0.087	0.913	0.358	22.5	16.1
3	0.007	0.993	0.540	276.5	213.8

Chapter IV

Properties of SM Scheme when Memory Length is Greater Than 1 Week

4.1. Limiting Scheme as Memory Becomes Infinite.

The SM scheme can be investigated in the limit, as the memory length increases, to see what kinds of tests it would involve. By definition, the SM scheme will not be used in the limit since that would be obviating its short memory characteristics. Because of its nature, this section will use the notation of Section 2.3.

The value of lambda can be estimated from the memory using the maximum likelihood estimate. Let $X=\sum_{i=1}^{s} X_i$ be the total number of cases from the s weeks of memory and let Y be the number of cases in the current week. It is assumed that the X_i are independent and identically distributed Poisson random variables with mean equal to λ so that X is a Poisson random variable with mean $s\lambda$. Then the likelihood

$$\Pr(X=x) = \frac{e^{-s\lambda}(s\lambda)^{x}}{x!}$$

can be maximized for λ yielding

$$\hat{\lambda} = \frac{X}{s} = \frac{\sum_{i=1}^{s} x_i}{s}.$$

Since $\hat{\lambda}$ is the m.l.e. for λ, it is consistent. Also $E(\hat{\lambda})=\lambda$, so the estimate is unbiased. Thus, as $s \rightarrow \infty$, $\hat{\lambda} \rightarrow \lambda$ with probability one.

The likelihood ratio principle can be used to find a test for testing whether or not the ratio of two Poisson means is equal to a specified value. The value of the ratio of the two Poisson means described above for X and Y is $\frac{\lambda_y}{\lambda_x}=\frac{\gamma\lambda}{s\lambda}=\frac{\gamma}{s}$. This single test amounts to an unconditional test in the SM scheme. Hoel (1945) observed that the critical regions for this likelihood ratio test statistic and the test statistic for a χ^2-test used for testing the equality of two Poisson means are essentially the same. The test of interest is

$$H_0\colon \gamma=1 \text{ vs. } H_1\colon \gamma>1.$$

The hypotheses H_1 and H_0 correspond to the parameter space

$$\Omega\colon \lambda>0, \quad \gamma>1,$$

and the subspace

$$\omega\colon \lambda>0, \ \gamma=1,$$

respectively. The likelihood

$$L = \Pr(X=x, Y=y) = \Pr(X=x)\Pr(Y=y) = \frac{e^{-\lambda s}(\lambda s)^x}{x!}\ \frac{e^{-\gamma\lambda}(\gamma\lambda)^y}{y!}$$

is maximized over the parameter space Ω and over the subspace ω. The likelihood ratio for the conditional binomial test, defined as $\mathcal{LR}$(CBT), is

$$\mathcal{LR}(\text{CBT}) = \frac{\max L_\omega}{\max L_\Omega} = \frac{(x+y)^{x+y}}{x^x\, y^y}\ \frac{s^x}{(s+1)^{x+y}}$$

and small $\mathcal{LR}$(CBT) leads to rejection of H_0. Since the testing remains invariant, this is the same likelihood ratio that results from using a Binomial likelihood and maximizing over the parameter spaces corresponding to the hypotheses

$$H_0\colon p=\frac{1}{s+1} \quad \text{and} \quad H_1\colon p=\frac{\gamma}{s+\gamma}.$$

The $\mathcal{LR}$(CBT) uses all of the available data so there is no apparent benefit, for this test, in ignoring part of the data or in treating a subset of the memory differently.

If the value of λ was known then the likelihood ratio principle can be used to derive a test statistic used for testing that the value of Y, which is the number of cases in the current week and a maximum likelihood estimate of λ_y, is equal to this known value. The test of interest is

$$H_0\colon \lambda_y = \lambda \text{ against } H_1\colon \lambda_y > \lambda \ .$$

The hypotheses H_1 and H_0 correspond to the parameter spaces

$$\Omega\colon \lambda_y > 0 \text{ and } \omega\colon \lambda_y = \lambda \ ,$$

respectively.

Maximizing the likelihood, $\Pr(Y=y) = \frac{e^{-\lambda}\lambda^y}{y!}$, over Ω yields $L_\Omega = \frac{e^{-y}y^y}{y!}$ and over ω yields $L_\omega = \frac{e^{-\lambda}\lambda^y}{y!}$. The likelihood ratio, defined as $\mathcal{LR}$(P), is then

$$\mathcal{LR}(P) = e^{-\lambda+y}\left(\frac{\lambda}{y}\right)^y .$$

The point now is to see how these two tests, $\mathcal{LR}$(CBT) and $\mathcal{LR}$(P) are related to each other. Rewritting $\mathcal{LR}$(CBT) yields

$$\mathcal{LR}(CBT) = \frac{(x+y)^{x+y}}{x^x\, y^y}\,\frac{s^x}{(s+1)^{x+y}}$$

$$= \left(\frac{x+y}{y}\,\frac{1}{s+1}\right)^y \left(\frac{x+y}{x}\,\frac{s}{s+1}\right)^x$$

$$= \left(\frac{1}{y}\right)^y \left(\frac{x+y}{s+1}\right)^y \left(1+\frac{y}{x}\right)^x \left(\frac{s}{s+1}\right)^x \ .$$

When passing to the limit as $s \to \infty$ the value of Y is unaffected since it is

not a function of the memory in any way. Thus, $\left(\frac{1}{y}\right)^y$ is a constant with respect to s and x. As $s \to \infty$, $\frac{x}{s+1} \to \lambda$ with probability 1 and $\frac{y}{s+1} \to 0$. By writing

$$\left(\frac{x+y}{s+1}\right)^y = \left(\frac{x}{s+1}+\frac{y}{s+1}\right)^y$$

and noting that since $\hat{\lambda} \to \lambda$ almost surely, then $g(\hat{\lambda}) \to g(\lambda)$ almost surely when g is continuous at λ, it must be that for every $\epsilon > 0$,

$$\Pr\left(\left| \left(\frac{x+y}{s+1}\right)^y - \lambda^y \right| < \epsilon\right) \to 1 \text{ as } s \to \infty .$$

It is necessary to bound the random variables to pass to the limits for the other terms in $\mathcal{LR}$(CBT). Since $\lambda_y < \infty$, for any $\delta > 0$ there exists a y_0 such that $\Pr(Y > y_0) < \delta$. That is, Y can be bounded away from y_0 with probability δ. For any $y_0 > 0$ there exists an $x_0 > 0$ such that $\Pr(X < x_0) < \delta$. That is, X can be bounded away from x_0 with probability δ. Thus, if $0 < y < y_0$ and $x > x_0 > 0$, then, for any $\epsilon > 0$

$$\Pr\left(\left|\left(1+\frac{y}{x}\right)^x - e^y\right| > \epsilon\right) < 2\delta .$$

Since $\Pr(X < x_0) \to 0$ as $s \to \infty$, then,

$$\Pr\left(\left|\left(1+\frac{y}{x}\right)^x - e^y\right| > \epsilon\right) \to 0 \text{ as } x \to \infty .$$

Also,

$$\left(\frac{s}{s+1}\right)^x = \left(\frac{s+1}{s}\right)^{-x} = \left(1+\frac{1}{s}\right)^{-x} = \left(1+\frac{\hat{\lambda}}{x}\right)^{-x}.$$

From the Cauchy-Schwarz inequality it is known that

$$\Pr\left(\left|\left(1+\frac{\hat{\lambda}}{x}\right)^{-x} - e^{-\lambda}\right| < \epsilon\right)$$

$$\leq \Pr\left(\left|\left(1+\frac{\hat{\lambda}}{x}\right)^{-x} - \left(1+\frac{\lambda}{x}\right)^{-x}\right| < \epsilon\right) + \Pr\left(\left|\left(1+\frac{\lambda}{x}\right)^{-x} - e^{-\lambda}\right| < \epsilon\right) .$$

Since $\hat{\lambda} \rightarrow \lambda$ almost surely, then $g(\hat{\lambda}) \rightarrow g(\lambda)$ almost surely when g is continuous at λ, so

$$\Pr\left(\left|\left(1+\frac{\hat{\lambda}}{x}\right)^{-x} - \left(1+\frac{\lambda}{x}\right)^{-x}\right| < \epsilon\right) \rightarrow 0 \quad \text{as } x \rightarrow \infty \quad .$$

Then, similar to what was argued above, for any $\delta > 0$ there exists an $x_0 > 0$ such that $\Pr(X < x_0) < \delta$. Thus, if $x > x_0 > 0$, then, for any $\epsilon > 0$

$$\Pr\left(\left|\left(1+\frac{\lambda}{x}\right)^{-x} - e^{-\lambda}\right| > \epsilon\right) < \delta \ .$$

Thus,

$$\Pr\left(\left|\left(1+\frac{\hat{\lambda}}{x}\right)^{-x} - e^{-\lambda}\right| < \epsilon\right) \rightarrow 0 \quad \text{as } x \rightarrow \infty \quad .$$

Since each of the terms of the $\mathcal{LR}$(CBT) shown above converge in probability, the product of these terms converges to the product of the limits. Thus,

$$\lim_{x \rightarrow \infty} \mathcal{LR}(\text{CBT}) = e^{-\lambda+y}\left(\frac{\lambda}{y}\right)^{y} = \mathcal{LR}(\text{P}) \ .$$

That is, as the memory increases, the test of the equality of two Poisson parameters becomes the test of a Poisson mean against a population parameter. The addition of the current data to the memory has no effect in

the limit for the second test since it is only adding one random variable to an infinite sum of random variables. The dependencies between tests break down and the second test has the same form as the first test which leads to a geometric distribution for the run length.

4.2. Alternative Ways of Using the Available Memory.

When the SM scheme has a memory length of one week there are two ways to use the memory. As the memory length increases, so does the number of ways that the memory can be used. A rule for testing that yields the most power for a test at a given point in time may put the scheme in a less than optimal position in terms of what the probabilities of rejection for subsequent tests might be. It may be that by sacrificing some power in the present test it is possible to make a subsequent test be more powerful than it would otherwise be if the present test is as powerful as possible. Take a scheme with a memory length of s and suppose that only the earliest $s-1$ of the s weeks' worth of data are used to compare with the current week. Then the first test will not have as much power as when all s of the available weeks are used but the next test may have more power than the equivalent next test since the memory will still contain only non-epidemic generated data, albeit $s-1$ weeks worth of it. If s is large, giving up the use of one week of memory may not reduce the power of the first test very much. Of course, as s grows, the impact of any single week on the data in memory is diminished anyway. But if the increase is large enough or perhaps if s is moderate then these two factors may not simply be offsetting. This lag in incorporating data into the memory does not have to be limited to one week. There could be a waiting time of many weeks before the data are put in

memory. As the lag time grows, the time when the data in the memory were collected is further back with respect to the current week. Of course, the practice of 'lagging' the data does not have to reduce the 'in use' length of the memory. However, any scheme which has a memory length of s weeks and a lag time of l weeks is a special case of an SM scheme with a memory length of $s+l$ weeks.

Similarly, an SM scheme with memory length equal to one week can be viewed as a special case of a larger memory SM scheme where only the most recent week of the memory is used. If only the most recent part of the memory is available then the SM scheme reduces to the smaller memory SM scheme for both the non-epidemic and epidemic situations. It does not seem useful to view the SM schemes in this way. It should be that under epidemic conditions SM schemes with longer memory lengths have higher probabilities of rejection at the tests involving combined epidemic and non-epidemic data than SM schemes with shorter memory lengths do, all other things being the same.

Consider a scheme with a memory length of 2 weeks. There are many different ways of using the data that involve a comparison of the current week's data to some or all of the data in memory. For example:

Plan 1) Compare current week to previous two weeks,

Plan 2) Compare current week to last week,

Plan 3) Compare current week to second previous week ,

Plan 4) Compare current week to previous two weeks but test once every 3rd week.

Plan 1 is the standard SM scheme for $s=2$, while Plan 2 reduces to the SM

scheme for s=1. Plan 3 is the lagged SM scheme and Plan 4 involves independent tests.

The parameters of the schemes defined by the four plans above will be chosen so as to make the expected run lengths under non-epidemic conditions equal. To avoid confusion in comparing the run lengths for schemes with different memory lengths define the parameters below. Let

$$\theta_r^{[2]} \equiv P_2(R>r \mid H_0) = Pr(R>r \mid H_0, s=2)$$

where $\alpha_{P1} \equiv 1-\theta_1^{[2]}$. This parameter was defined in general in Chapter II only now the memory parameter is explicitly shown as an argument.

Then

$$E(R|H_0, \text{Plan } 1) = 1 + \theta_1^{[2]} + \frac{\left(\theta_2^{[2]}\right)^2}{\theta_2^{[2]} - \theta_3^{[2]}}$$

$$= 1 + \theta_1 + \frac{\left(\theta_2^{[2]}\right)^2}{\theta_2^{[2]} - \theta_3^{[2]}}$$

and

$$E(R|H_1, \text{Plan } 1) = 1 + \sum_{r=1}^{3} \beta_r^{[2]} + \frac{\left(\beta_4^{[2]}\right)^2}{\beta_4^{[2]} - \beta_5^{[2]}} .$$

where $\beta_r^{[2]} = Pr(R>r \mid H_1, s=2)$ shows the memory paramter explicitly.

For Plan 2, it was shown in Chapter III that

$$E(R|H_0, \text{Plan } 2) = 1 + \theta_1 + \frac{\theta_1\theta_2}{\theta_1 - \theta_2}$$

where $\alpha_{P2} \equiv 1-\theta_1$ and it is chosen so that $E(R|H_0,\text{Plan } 2)=E(R|H_0,\text{Plan } 1)$.

Also

$$E(R|H_1, \text{Plan } 2) = 1 + \beta_1 + \frac{\beta_2^2}{\beta_2 - \beta_3} .$$

The run length distribution under Plan 3 can be derived using only the parameters defined under Plan 2. For Plan 3, a test does not depend on the test in the week immediately preceding it since they do not share any data. Thus under H_0,

$$\Pr(R>1) = \Pr(T_1=0) = \theta_1$$

where $\alpha_{P3} \equiv 1-\theta_1$ and it is chosen so that $E(R|H_0,\text{Plan } 3)=E(R|H_0,\text{Plan } 1)$. Then

$$\Pr(R>2) = \Pr(T_1=0,\ T_2=0) = \Pr(T_2=0)\ \Pr(T_1=0) = \theta_1^2 .$$

Similarly,

$$\begin{aligned}\Pr(R>3) &= \Pr(T_3=0|T_2=T_1=0)\ \Pr(T_2=0|T_1=0)\ \Pr(T_1=0)\\ &= \Pr(T_3=0|T_1=0)\ \Pr(T_2=0)\ \Pr(T_1=0)\\ &= \frac{\theta_2}{\theta_1}\,\theta_1\,\theta_1 = \theta_2\,\theta_1 .\end{aligned}$$

For larger R, the Markov property is invoked so that the probability is conditionalized on at most two weeks. Thus,

$$\Pr(R>4) = \Pr(T_4=0|T_3=T_2=T_1=0)\ \Pr(T_3=0|T_2=T_1=0)\ \Pr(T_2=0|T_1=0)\ \Pr(T_1=0)$$

and by the Markov property

$$= \Pr(T_4=0|T_3=T_2=0)\ \Pr(T_3=0|T_2=T_1=0)\ \Pr(T_2=0|T_1=0)\ \Pr(T_1=0)$$

and by the pairwise independence of successive tests

$$\begin{aligned}&= \Pr(T_4=0|T_2=0)\ \Pr(T_3=0|T_1=0)\ \Pr(T_2=0)\ \Pr(T_1=0)\\ &= \frac{\theta_2}{\theta_1}\,\frac{\theta_2}{\theta_1}\,\theta_1\,\theta_1 = \theta_2^2.\end{aligned}$$

Similarly,

$$\begin{aligned}\Pr(R>5) &= \Pr(T_5=0|T_3=0)\ \Pr(T_4=0|T_2=0)\ \Pr(T_3=0|T_1=0)\ \Pr(T_2=0)\ \Pr(T_1=0)\\ &= \left(\frac{\theta_2}{\theta_1}\right)^3 \theta_1^2 = \theta_2^2\,\frac{\theta_2}{\theta_1} .\end{aligned}$$

In general,

$$\Pr(R>4+t) = \theta_2^2 \left(\frac{\theta_2}{\theta_1}\right)^{t-1}.$$

Then,

$$E(R|H_0, \text{Plan } 3) = 1 + \theta_1 + \theta_1^2 + \theta_1\theta_2 + \theta_2^2 \sum_{r=4}^{\infty} \left(\frac{\theta_2}{\theta_1}\right)^{r-4}$$
$$= 1 + \theta_1 + \frac{\theta_1^3}{\theta_1 - \theta_2} .$$

Now,

$$E(R|H_0, \text{Plan } 3) - E(R|H_0, \text{Plan } 2) = \theta_1 \frac{\theta_1^2 - \theta_2}{\theta_1 - \theta_2} .$$

In similar fashion the run length distribution can be derived for Plan 3 under epidemic conditions, where,

$$\Pr(R>1)=\beta_1, \qquad \Pr(R>2)=\beta_1^2, \qquad \Pr(R>3)=\beta_1\beta_2,$$
$$\Pr(R>4)=\beta_2^2, \qquad \Pr(R>5)=\beta_2\beta_3$$

and for $r>6$,

$$\Pr(R>6+t) = \beta_3^2 \left(\frac{\beta_3}{\beta_2}\right)^{t-1}.$$

Thus,

$$E(R|H_1, \text{Plan } 3) = 1 + \beta_1 + \beta_1^2 + \beta_1\beta_2 + \beta_2^2 + \beta_2\beta_3 + \frac{\beta_2\beta_3^2}{\beta_2 - \beta_3}$$

$$= 1 + \beta_1 + \beta_1^2 + \beta_1\beta_2 + \frac{\beta_2^3}{\beta_2 - \beta_3} .$$

Under Plan 4, there are independent tests every three weeks. Similar to what was described in Section 3.3, there is a probability of 1/3 that the scheme will be testing during week 1, when the epidemic has started. If the probability of rejection for a test which is performed is set at $\alpha_{P4} \equiv 1 - \theta_1$, then,

Pr(Test performed and rejects)

$$= \Pr(\text{Test performed}) \Pr(\text{Test rejects}|\text{Test performed})$$
$$= \left(\frac{1}{3}\right) \alpha_{P4} .$$

where α_{P4} is determined so that $E(R|H_0, \text{Plan } 4) = E(R|H_0, \text{Plan } 1)$.

Let R_1, R_2 and R_3 be the random variables which represent the run length when testing starts at week 1, week 2 and week 3, respectively and let T be the random variable which represents the number of tests until the first alarm. The run length distribution is found by transforming to R_1, R_2 and R_3 from the probability distribution of T. Then

$$\left.\begin{aligned} \Pr(R_1=r) &= \Pr(R_1=3t-2) \\ \Pr(R_2=r) &= \Pr(R_2=3t-1) \\ \Pr(R_3=r) &= \Pr(R_3=3t) \end{aligned}\right\} = 3(1-\theta_1)(3\theta_1-2)^{t-1} \quad t=1,2,3,\ldots$$

And,

$$E(R_1|H_0) = \sum_{t=1}^{\infty}(3t-2)\ 3(1-\theta_1)(3\theta_1-2)^{t-1} = \frac{3}{3(1-\theta_1)} - 2 = \frac{2\theta_1-1}{1-\theta_1},$$

$$E(R_2|H_0) = \frac{\theta_1}{1-\theta_1},$$

$$E(R_3|H_0) = \frac{1}{1-\theta_1}.$$

Then,

$$E(R|H_0, \text{Plan 4}) = \frac{1}{3}\left[E(R_1|H_0) + E(R_2|H_0) + E(R_3|H_0)\right]$$

$$= \frac{1}{3}\left[\frac{2\theta_1-1+\theta_1+1}{1-\theta_1}\right]$$

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

That $E(R|H_0, \text{Plan 4}) = E(R_2|H_0)$ is not surprising since $R_1=R_2-1$ and $R_3=R_2+1$. Thus, $E(R|H_0) = \frac{1}{3}\left[3E(R_2|H_0)-1+1\right] = E(R_2|H_0)$.

Under epidemic conditions, additional definitions are needed to specify the run length distributions of R_1 and R_2. For R_1, it is necessary to define the power of the test which operates at an error rate of $3(1-\theta_1)$ and is

comparing one epidemic week to two non-epidemic weeks. For R_2, it is necessary to define the power of a test which also operates at an error rate of $3(1-\theta_1)$ but which compares one epidemic week to a mixture of one epidemic week and one non-epidemic week.

For R_1, define

$$\Pr(R_1=1|H_1) \equiv 1-\beta_1'$$

and for $r>2$

$$\Pr(R_1=r|H_1) = \beta_1' \; 3(1-\theta_1)(3\theta_1-2)^{t-2} \; .$$

Then,

$$E(R_1|H_1) = 1-\beta_1' + \beta_1' \sum_{t=2}^{\infty}(3t-2) \; 3(1-\theta_1)(3\theta_1-2)^{t-2}$$

$$= 1-\beta_1' + \beta_1'\left[\frac{3}{3(1-\theta_1)} + 3 - 2\right]$$

$$= 1 + \frac{\beta_1'}{1-\theta_1}.$$

For R_2, define

$$\Pr(R_2=1|H_1) \equiv 1-\beta_1''$$

and for $r>2$

$$\Pr(R_2=r|H_1) = \beta_1'' \; 3(1-\theta_1)(3\theta_1-2)^{t-2} \; .$$

Then,

$$E(R_2|H_1) = 2(1-\beta_1'') + \beta_1'' \sum_{t=2}^{\infty}(3t-1)3(1-\theta_1)(3\theta_1-2)^{t-2}$$

$$= 2-2\beta_1'' + \beta_1''\left[\frac{3}{3(1-\theta_1)} + 3 - 1\right]$$

$$= 2 + \frac{\beta_1''}{1-\theta_1}.$$

Since R_3 is only testing epidemic data under epidemic conditions, then the run

length distribution is the same as under H_0. Thus,

$$E(R_3|H_1) = \frac{1}{1-\theta_1} \quad .$$

And,

$$E(R|H_1, \text{Plan } 4) = 1 + \frac{\frac{1}{3}\left(\beta_1' + \beta_1'' + 1\right)}{(1-\theta_1)} \quad .$$

The expected run lengths for the four plans under epidemic and nonepidemic conditions are presented in Table 4.1 below.

Table 4.1

Expected Run Lengths Under 4 Plans, s=2

Plan	$E(R\|H_0)$	$E(R\|H_1)$
1	$1 + \theta_1 + \frac{\left(\theta_2^{[2]}\right)^2}{\theta_2^{[2]} - \theta_3^{[2]}}$	$1 + \sum_{r=1}^{3} \beta_r^{[2]} + \frac{\left(\beta_4^{[2]}\right)^2}{\beta_4^{[2]} - \beta_5^{[2]}}$
2	$1 + \theta_1 + \frac{\theta_1\theta_2}{\theta_1 - \theta_2}$	$1 + \beta_1 + \frac{\beta_2^2}{\beta_2 - \beta_3}$
3	$1 + \theta_1 + \frac{\theta_1^3}{\theta_1 - \theta_2}$	$1 + \beta_1 + \beta_1^2 + \beta_1\beta_2 + \frac{\beta_2^3}{\beta_2 - \beta_3}$
4	$\frac{\theta_1}{1-\theta_1}$	$1 + \frac{\frac{1}{3}\left(\beta_1' + \beta_1'' + 1\right)}{(1-\theta_1)}$

4.3. Numerical Example of the Four Plans.

The formulas in Table 4.1 are difficult to compare. Certainly an SM scheme with s=2 is expected to outperform an SM scheme with s=1 for the first test under epidemic conditions when the θ_1 is equal for both schemes

This is because a test that a binomial parameter has changed from 1/2 to 2/3 is less powerful than the test that it has changed from 1/3 to 1/2. Numerical examples will go a long way towards deciphering some of the relationships of the parameters and the expected run lengths of the four different plans. By setting the expected run lengths of the four plans to be equal under non-epidemic conditions the expected run lengths under epidemic conditions can be compared fairly. The results for an example with $\lambda_0=1$ and $\gamma=4$ and $\gamma=10$ with $E(R|H_0)=18.8$ are shown below in Table 4.2.

Table 4.2

Comparison of Four Plans with $E(R|H_0)=18.8$, $\lambda=1$

	Plan			
	1	2	3	4
α_{p_q}	.05	.0517	.0517	.0506
ARL, $\gamma=4$	8.2	12.2	8.7	13.7
ARL, $\gamma=10$	1.06	1.28	1.12	8.89

When $\gamma=4$, Table 4.2 shows that Plan 1 outperforms Plan 3 by one-half of a week. The other two plans are distinctly inferior but it is interesting to note that Plan 4 which uses two weeks of memory but tests every third week does not do as well as Plan 2 which only uses one week of memory but tests every week.

When $\gamma=10$, Plan 1 has the smallest expected run length but Plan 3 is not inferior by much. Plan 2 is now almost as powerful as Plans 1 and 3 with

the larger fold increase. The performance of Plan 4 is not very good with an expected run length under epidemic conditions that is 7 to 8 times as large as Plans 1, 2 or 3.

The results for an example with $\lambda_0=1$ and $\gamma=4$ and $\gamma=10$ when $E(R|H_1)=98.6$ are shown below in Table 4.3.

Table 4.3

Comparison of Four Plans with $E(R|H_0)=98.6$, $\lambda_0=1$

	Plan			
	1	2	3	4
α_{pq}	.01	.01008	.01008	.01004
ARL, $\gamma=4$	69.7	82.3	73.0	83.8
ARL, $\gamma=10$	1.35	1.91	1.53	49.38

In Table 4.3, when $\gamma=4$, Plan 4 again has the smallest expected run length followed by Plan 3 while Plans 2 and 3 are very similar to each other. When $\gamma=10$, Plans 1 and 3 look similar. Plan 2 has a dramatic drop in its expected run length from when $\gamma=4$ and is almost equal to Plans 1 and 3. It is clear from these results that Plan 4 is not a viable competitor. This is especially true when $\gamma=10$ in Table 4.3 and the scheme under Plan 4 only has a 2/3 probability of testing with data that is non-identically distributed since R_3 only involves epidemic data. In both of tables, Plan 3 is between Plan 2 and Plan 1. Judging by the size of the expected run length under epidemic conditions, the winner is Plan 1.

Chapter V

Comparison of Short Memory Scheme with CUSUM Scheme

5.1. Comparison of SM Scheme with Tabled Values of the CUSUM Scheme.

Ewan and Kemp (1960) have calculated the values of the parameters, h and k, necessary to operate a CUSUM scheme for Poisson variates. Their tables cover a selected set of acceptable means (.15 - 9.0) and rejectable means (.76 - 12.81) representing a range of fold-increases (5.11 - 1.42). The CUSUMs have an average run length of 500 under H_0 and an ARL of 7 under H_1 for the values of h and k given. That the CUSUM scheme can perform so well with expected values of this magnitude shows that it is a powerful technique, especially for detecting a 5.07-fold increase in a mean as small as 0.15 or for detecting an increase of 1.65-fold from a mean of 3.84.

The SETS scheme is not compared to the SM scheme here. The SM scheme and the CUSUM scheme both do tests at regularly scheduled times on grouped data. The SETS scheme only does a test when the data dictate that one be done and exact times of the event under study are available. Thus, the occurence of an event will cause a test that judges the sizes of the sets to be done. The expected run length of the SETS scheme could be compared to the expected run length of either of the other two schemes. However, in the surveillance situation of monitoring from previously grouped data the applicability of the SETS scheme is doubtful.

Table 5.1 below, reproduced from Ewan and Kemp (1960), shows the means and parameters of CUSUM schemes with $E(R|H_0)=500$ and $E(R|H_1)=7$. The expected run lengths of the SM scheme will be calculated using the means in this table.

Table 5.1*

Values of $\gamma=\frac{\lambda_1}{\lambda_0}$, h and k from CUSUM schemes with ARL=500 at AQL and ARL=7 at RQL

λ_0	λ_1	γ	h	k
.19	.97	5.11	4.00	.30
.15	.76	5.07	3.00	.30
.16	.81	5.06	3.25	.30
.17	.86	5.06	3.50	.30
.18	.90	5.00	3.75	.30
.22	1.02	4.64	2.00	1.00
1.01	2.45	2.53	4.00	2.00
2.37	4.40	1.85	9.00	3.00
3.84	6.33	1.65	9.00	5.00

*Reprinted from Ewan and Kemp, 1960.

To compare the SM scheme with the CUSUM scheme it is necessary to set $E(R|H_0)$ for the SM scheme equal to the ARL for the CUSUM scheme. The expected run lengths and the β_1, β_2 and β_3 parameters for the SM scheme with memory length equal to one week are presented in Table 5.2 below and for memory length equal to two weeks they are presented below in Table 5.3.

Table 5.2

Expected Run Lengths for SM Scheme with $E(R|H_0)=500$, $s=1$

λ_0/λ_1	β_1	β_2	β_3	$E(R\|H_1)$
.19/.97	.9956	.9936	.9917	498
.15/.76	.9963	.9943	.9923	499
.16/.81	.9962	.9942	.9922	499
.17/.86	.9960	.9940	.9920	498
.18/.90	.9959	.9939	.9919	498
.22/1.02	.9955	.9936	.9916	498
1.01/2.45	.9917	.9897	.9877	496
2.37/4.40	.9871	.9851	.9831	494
3.84/6.33	.9849	.9829	.9810	492

Table 5.3

Expected Run Lengths for SM Scheme with $E(R|H_0)=500$, $s=2$

λ_0/λ_1	β_1	β_2	β_3	β_4	β_5	$E(R\|H_1)$
.19/.97	.9908	.9865	.9845	.9825	.9805	493
.15/.76	.9933	.9896	.9877	.9857	.9837	495
.16/.81	.9928	.9890	.9870	.9850	.9830	495
.17/.86	.9922	.9883	.9863	.9843	.9823	494
.18/.90	.9918	.9877	.9857	.9837	.9817	494
.22/1.02	.9905	.9861	.9841	.9821	.9801	492
1.01/2.45	.9793	.9729	.9709	.9689	.9670	487
2.37/4.40	.9754	.9683	.9663	.9643	.9624	485
3.84/6.33	.9746	.9673	.9653	.9634	.9614	484

Despite the fact that there is a an increase as large as 5.07-fold, the SM scheme does not do as well as the CUSUM in detecting it since it is an increase from a mean that is small, 0.15. For the larger means, the increase that is used is not very large considering the Type I error. So in either case the SM scheme does not do very well when compared to the CUSUM.

When the mean of the underlying Poisson distribution is small, it likely will not yield sufficient information about $\frac{\lambda_y}{\lambda_x}$. This is a problem since for a fixed value of the ratio $\frac{\lambda_y}{\lambda_x}$, the power to detect an increase in λ_0 depends on λ_0, as well as on γ, and the SM scheme will not know the value of λ_0 *apriori*. Birnbaum (1954) and Lehmann (1986) show that this can be overcome in a conditional binomial test by taking samples sequentially until enough data is collected to provide the information necessary to detect a difference in the parameters. That is, find an N′ such that if $X+Y=N\geq N'$, then the power to detect the increase in the conditional binomial parameter from p_0 to p_1 can be bounded from below. As $N' \rightarrow \infty$, the power of the test goes to one. This can not be achieved by simply setting an upper limit on the number of Poisson random variables to use in a conditional binomial test. In the case where λ_0 is unknown, it may very well be small. So taking a sample of the prescribed size is no guarantee that enough information will be provided to make a decision that is powerful. The conditional binomial test used in the SM scheme can not produce a power that is arbitrarily close to one as s increases to infinity. Since $N=X+Y$ and $X \rightarrow \infty$ as $s \rightarrow \infty$, it must be that $N\geq N'$ with a high probability, even for small λ_0. However, the values of $p_0=\frac{1}{s+1}$ and $p_1=\frac{\gamma}{s+\gamma}$ are both going to zero as s goes to infinity. The effect of this is to increase the value of N′ necessary for the

conditional binomial test to achieve a certain power. Thus, as $s \rightarrow \infty$ then $N' \rightarrow \infty$.

It is clear that the CUSUM scheme is more powerful than the SM scheme when the memory length is small. However, the length of the memory of an SM scheme is a parameter of the scheme and can be selected according to a set of criteria. If the time that elapses between the previous alarm and the first test of a newly started scheme is not of interest then the memory parameter can be increased indefinitely in pursuit of a desired run length or a desired probability of detection under epidemic conditions. The resulting value of s might only be of theoretical interest since a large s may be too unwieldy in practice. For the conditions in Table 5.1 the SM scheme may be hard pressed to match the CUSUM scheme with any value for s.

As the memory length becomes larger the ability to accurately estimate λ_0 increases. Using the same notation as in Chapter 4, under non-epidemic conditions, the m.l.e. of λ_0 is $\hat{\lambda}_0=(x+y)/(s+1)$. It is consistent so the standard error of the estimate goes to zero as s goes to infinity. That is, $\sqrt{\hat{\lambda}_0/(s+1)} \rightarrow 0$ as $s \rightarrow \infty$. Also, $\hat{\lambda}_0$ is unbiased for λ_0 since $E\left[\frac{(x+y)}{(s+1)}|H_0\right] = E\left[\frac{1}{(s+1)}\left(\sum_{i=1}^{s} x_i + y\right)|H_0\right] = \frac{(s+1)\lambda_0}{s+1} = \lambda_0$. Thus, for any $\delta>0$ and $\epsilon>0$, there exists an s_0 such that, for any $s>s_0$, $Pr(|\hat{\lambda}_0-\lambda_0|>\epsilon) < \delta$.

For large values of s, the test of the Binomial parameter can be carried out using a normal approximation. Under H_0, the Type I error is

$$Pr\left[\frac{y-(x+y)p}{\sqrt{(x+y)p(1-p)}} \geq z_{1-\alpha}\right] = Pr\left[\frac{y-\left(\frac{x+y}{s+1}\right)}{\sqrt{\left(\frac{x+y}{s+1}\right)\left(\frac{s}{s+1}\right)}} \geq z_{1-\alpha}\right] = Pr\left[\frac{\hat{\lambda}_y-\hat{\lambda}_0}{\sqrt{\hat{\lambda}_0\left(\frac{s}{s+1}\right)}} \geq z_{1-\alpha}\right].$$

For large s, this is equal to $\Pr\left(\frac{\hat{\lambda}_y-\lambda_0}{\sqrt{\lambda_0}} \geq z_{1-\alpha}\right) = \Pr\left(\hat{\lambda}_y \geq z_{1-\alpha}\sqrt{\lambda_0}+\lambda_0\right)$. Then,

under H_1,

$$1-\beta = \Pr\left[\frac{y-\gamma\left(\frac{x+y}{s+\gamma}\right)}{\sqrt{\gamma\left(\frac{x+y}{s+\gamma}\right)\left(\frac{s}{s+\gamma}\right)}} \geq \frac{(z_{1-\alpha}\sqrt{\lambda_0}+\lambda_0)-\gamma\left(\frac{x+y}{s+\gamma}\right)}{\sqrt{\gamma\left(\frac{x+y}{s+\gamma}\right)\left(\frac{s}{s+\gamma}\right)}}\right].$$

In any case, in the limit as $s \rightarrow \infty$, the conditional binomial test becomes the test of a Poisson parameter, λ_y, against a known value, λ_0. The UMP test for testing H_0: $\lambda_y=\lambda_0$ versus H_1: $\lambda_y>\lambda_0$ is based on $\Pr(Y>c)$ where c is chosen so that $\Pr(Y>c|\lambda_0)\leq\alpha$ and the distribution of Y is, of course, Poisson. With randomized testing, c and w_α are found such that $\alpha=w_\alpha\Pr(Y=c|\lambda_0) + \Pr(Y>c|\lambda_0)$. Then, the power of this test, $1-\beta$, is $w_\alpha\Pr(Y=c|\lambda_y) + \Pr(Y>c|\lambda_y)$. This test can not produce power that is arbitrarily close to one as s increases because the test is dependent on the variation in Y which is the m.l.e. of λ_y. Thus, even though λ_x becomes essentially known, and is used as the estimate for λ_0, the estimate of λ_y comes from only one week of data and has variation that is not diminished as $s \rightarrow \infty$ since the estimate of the standard error of $\hat{\lambda}_y$ is simply $\sqrt{y}$.

As s increases and the dependencies between tests disappear, the run length distribution becomes geometric with parameter $p_i=\Pr(T=1|H_i)$, $i=0,1$ and $E(R|H_i)=\frac{1}{\Pr(T=1|H_i)}$, $i=0,1$. The parameter, p_i, equals α under non-epidemic conditions and equals $1-\beta$ under epidemic conditions. The limiting form of the SM scheme is compared to the CUSUM in Table 5.4 for the same means as in Table 5.1. $E(R|H_0)=1/\alpha$ is set at 500 by randomizing the test so that $\alpha=0.002$.

Table 5.4 shows that the limiting form of the SM scheme does much better than the SM scheme does when s is small. The cost is that it depends on λ_0 being known. More significant is the fact that even though this test is the UMP test of H_0: $\lambda_y=\lambda_0$ versus H_1: $\lambda_y>\lambda_0$, it does not yield an expected run length under epidemic conditions that matches the expected run length of the CUSUM scheme for these means. Thus, the CUSUM scheme operates at a level which is better than what the independent UMP tests can yield.

Table 5.4

$E(R|H_1)$ under limiting form of SM Scheme with $E(R|H_0)=500$

λ_0	λ_1	γ	Power	$E(R\|H_1)$
0.19	0.97	5.11	0.087	11.5
0.15	0.76	5.07	0.063	15.9
0.16	0.81	5.06	0.067	14.9
0.17	0.86	5.06	0.073	13.7
0.18	0.90	5.00	0.077	13.0
0.22	1.02	4.64	0.089	11.2
1.01	2.45	2.43	0.067	14.9
2.37	4.40	1.85	0.057	17.5
3.84	6.33	1.65	0.056	17.9

The CUSUM scheme operates on new data by building partial sums of deviations of the observations from their target value. This target value is made up from the known parameter and a tolerable amount of variation from it. The CUSUM resets to zero whenever the partial sum would actually fall

below zero and a new run is started. Otherwise, as long as the CUSUM is positive, a new deviation is added to the sum of the other deviations in this run of the CUSUM. Thus, under non-epidemic conditions the partial sum does not accumulate, except for large deviations due to random variation, since it will keep resetting at zero. Only under the epidemic conditions will the CUSUM steadily build towards the boundary value with positive-valued partial sums. Thus, it will not be accumulating lots of non-epidemic data before the epidemic period begins and thereby is not 'mixing' data from non-epidemic and epidemic periods.

For a short memory scheme, a similar situation would be to accumulate the new epidemic cases together with the old cases and not pass the old data into the memory, at least not immediately. In the limiting case of the SM scheme, where λ_0 is known, adding data to the memory does not have an impact on it anyway so the epidemic data is accumulated and never passed into the memory. If the starting time of the epidemic is known then the test will be done on epidemic data only. Otherwise, the accumulated data would be a mixture of cases from the epidemic and non-epidemic periods. The run length distribution is not geometric since tests are sharing data and creating dependencies. Suppose that the epidemic starts at the conclusion of week 0 as has been assumed for all of the schemes throughout. The m.l.e. of λ_y is consistent so the variance of the estimate goes to zero as more data is accumulated and used to estimate λ_y. Thus, as the scheme continues and λ_y is more accurately estimated, the power of the test converges to one since λ_y can be more easily distinguished from λ_0.

An example of how the power of the UMP test of a Poisson parameter, H_0: $\lambda = \lambda_y$ against H_1: $\lambda > \lambda_y$, increases as the number of weeks of

accumulated data increases is shown below in Table 5.5 where $\gamma=\frac{\lambda_1}{\lambda_0}=4.64$. Each test is randomized so that α, the Type I error, equals 0.002. This table shows that as the data is accumulated into the 10th week, the power of the test becomes 0.798 whereas at the first test the power was only equal to 0.089.

Table 5.5

Power of Detecting $\gamma=4.64$-fold Increase with Known λ_0

$Pr(T=1|H_0)=0.002$

	Number of Weeks in Test									
	1	2	3	4	5	6	7	8	9	10
λ_0	0.22	0.44	0.66	0.88	1.10	1.32	1.54	1.76	1.98	2.20
λ_1	1.02	2.04	3.06	4.08	5.10	6.12	7.15	8.17	9.19	10.21
$Pr(T=1\|H_1)$	.089	.169	.254	.376	.443	.541	.606	.686	.731	.798

The SM scheme can operate by accumulating cases from more than one week and testing this combined number of cases against the number of cases in memory and then passing the oldest of the 'comparison data' into the memory for the next test. Suppose that $s=2$, then the first test under epidemic conditions will compare a mixture of epidemic and non-epidemic data to non-epidemic data. The second test will compare 2 weeks of epidemic data to one week of non-epidemic data. The third test, as in the regular SM scheme, will operate on all epidemic data. The results are shown below in Table 5.6 where the tests are randomized so that the expected run length under non-epidemic conditions is 500. The results show that the first test is

not as powerful as the first test for the regular SM scheme and the subsequent tests do not make up for this initial deficiency.

Table 5.6

Expected Run Lengths for Scheme with 2 Comparison Weeks

$E(R|H_0)=500$

λ_0/λ_1	β_1	β_2	β_3	β_4	β_5	$E(R\|H_1)$
.19/.97	.9970	.9927	.9907	.9887	.9867	496
.15/.76	.9973	.9936	.9916	.9896	.9877	498
.16/.81	.9972	.9934	.9914	.9894	.9874	498
.17/.86	.9971	.9932	.9912	.9892	.9872	497
.18/.90	.9971	.9930	.9910	.9890	.9871	497
.22/1.02	.9970	.9926	.9906	.9886	.9866	495

5.2. Effect on Expected Value when Mean is Misstated.

One of the benefits to be derived from using the SM scheme is that even though the power of the scheme does depend on λ_x and λ_y, and not just on $\frac{\lambda_y}{\lambda_x}$, the effect when the underlying mean is mis-stated or poorly estimated may not be as devastating as under the CUSUM scheme . This event could be likely when good baseline records are not available or after an alarm is sounded and the subsequent investigation has changed the baseline rate from what it had been before the increase occurred. Pollak and Siegmund (1985) state that the most interesting variation from the traditional situation in which a CUSUM scheme is used is when the initial, in control or non-epidemic, parameter value is unknown. The expected value of the run length

distribution with the CUSUM scheme is unknown to these authors under these circumstances although an alternative is suggested by them.

The CUSUM is investigated under the circumstances where the baseline rate, λ_0, is underestimated by a factor, f, and then the increased rate, λ_1, is calculated as $\lambda_1 = f\gamma\lambda_0$. The parameters for the CUSUM, h and k, are determined for λ_0 and $\gamma\lambda_0$ but it is actually operating on data generated from a Poisson distribution with a mean of $f\lambda_0$ not λ_0 or with a mean of $f\gamma\lambda_0$ not $\gamma\lambda_0$. The 'true' epidemic rate, $f\gamma\lambda_0$, will not even represent an increase over the misstated non-epidemic rate, λ_0, if $f \leq 1/\gamma$. If $f = 1/\gamma$ then $E(R|H_1, f\gamma\lambda_0) = E(R|H_0, \lambda_0)$. The question under study is how robust the CUSUM is to this problem. That is, how powerful does the CUSUM technique remain when the baseline rate is misstated.

Under this scenario, the SM scheme is also operating on Poisson data generated with the mean equal to $f\lambda_0$ and $f\gamma\lambda_0$ instead of λ_0 and $\gamma\lambda_0$. The SM scheme is set up so that $E(R|H_0) = 500$ to match Ewan and Kemp's value for ARL at AQL and this is independent of λ_0 as shown in Chapter II. Then, since $E(R|H_0) \geq E(R|H_1)$ there is an upper bound on the expected run length under epidemic conditions when the true mean is misstated. Of course, for the range of values given in Ewan and Kemp (1960), the SM scheme is not as powerful as the CUSUM scheme when the mean is stated correctly, as shown in Section 5.1. However, the SM scheme affords protection against operating at a much lower power than anticipated.

The CUSUM parameters, h and k, and the fold increase were taken from Table 3 in the 1960 paper by Ewan and Kemp. These were used to simulate CUSUM schemes where λ_0 was underestimated by $(1-f)100\%$. The CUSUM scheme was run 1000 times for each underestimated value of λ_0 and

the detection times were then averaged. This was also done for the tabled values of λ_0 and $\gamma\lambda_0$ as a check and it resulted in ARLs equal to 7 and 500, respectively; the simulated values are not included in the results. The results for $\lambda_0=9.0$, $\gamma\lambda_0=12.81$ with $\gamma=\lambda_1/\lambda_0=1.42$ are in Table 5.7 and the results for $\lambda_0=3.84$ and $\gamma\lambda_0=6.33$ with $\gamma=\lambda_1/\lambda_0=1.65$ are in Table 5.8. The value of γ where $E(R|H_1)=E(R|H_0)$ is $1/1.42=.703$ in Table 5.7 and $1/1.65=.61$ in Table 5.8.

The CUSUM still seems powerful when λ_0 is being slightly to moderately underestimated in Tables 5.7 and 5.8 since the expected run length is still much smaller than what is seen for the SM scheme when the memory is very short. The CUSUM also appears robust since the expected run length does not rise rapidly to 500 as f decreases but rather increases steadily as f decreases. This is certainly a good property and affords some protection when the estimate of the baseline mean is only slightly off. However, even though the ARL does not immediately rise to 500 from seven, the increases in the expected run length under epidemic conditions are substantial. At $f=.80$, the expected run length is increased by a factor of 7.8 in Table 5.7 and by a factor of 3 in Table 5.8. For $f=.75$, these factors increase to 22.9 and 5.7, in the two tables respectively. Thus, the CUSUM scheme is operating considerably less efficiently when $f=.75$ than when $f=1$. The smallest factors of underestimation are $f=.60$ for Table 5.7 and $f=.50$ for Table 5.8. At these levels the expected run length under epidemic conditions is 1987.5 and 1215.6 times larger, respectively, than they were intended to be. Clearly, this represents an unacceptable predicament since the time to detection under epidemic conditions is far greater than expected at levels of inaccuracy that are realistic to expect.

An upper bound for the expected run length of the SM scheme which needs no calculations was seen to be $E(R|H_0)=500$. This seems like an unapproachably large value for the CUSUM scheme when it is operating under the best of circumstances. However, this value is surpassed in situations that are not unreasonable and once the ARL does surpass 500, it increases rapidly thereafter.

For the smaller means shown in Ewan and Kemp (1960), the γ-fold increase is around 4 and would need f to be around .25 for the expected run length to inflate to 500. The cases for the smaller means were not investigated. Since the CUSUM is more powerful than the SM scheme with small s and $E(R|H_1)$ does increase slowly towards $E(R|H_0)$ until $f=1/\gamma$ the true rate would have to be very small for the SM scheme to be more useful than the CUSUM scheme. Rates as small as $0.22/4=0.055$ may be realistic to expect. However, the SM scheme would not be affording very much power at that rate for small s. Thus, while the SM scheme does afford protection against misspecification of the baseline rate, if the power of the scheme to detect an increase is small it may not be a viable option, anyway.

Table 5.7

ARL of CUSUM under H_1 with Underestimated Mean

$\lambda_0=9.0 \quad \gamma\lambda_0=12.81 \quad h=12 \quad k=11$

$f\gamma\lambda_0$	$\bar{R}$	$\hat{\sigma}_{\bar{R}}$	f	$\bar{R}/E(R\|H_1)$
12.81	7.	—	1.00	1.0
11.50	14.2	.33	.90	2.0
10.22	54.7	1.65	.80	7.8
9.56	160.6	4.87	.75	22.9
9.20	334.9	10.43	.72	47.8
9.0	500.	—	.703	71.4
8.95	569.2	18.25	.699	81.3
7.67	13912.3	466.57	.60	1987.5

Table 5.8

ARL of CUSUM under H_1 with Underestimated Mean

$\lambda_0=3.84 \quad \gamma\lambda_0=6.33 \quad h=9 \quad k=5$

$f\gamma\lambda_0$	$\bar{R}$	$\hat{\sigma}_{\bar{R}}$	f	$\bar{R}/E(R\|H_1)$
6.33	7.	—	1.00	1.0
5.70	11.0	0.24	.90	1.5
5.07	21.6	0.53	.80	3.1
4.75	39.7	1.08	.75	5.7
4.44	76.3	2.13	.70	10.9
4.25	120.7	3.65	.67	17.2
4.12	195.2	5.74	.65	27.9
3.84	500.	—	.61	71.4
3.80	561.1	16.93	.60	80.2
3.17	8509.4	271.63	.50	1215.6

Chapter VI

A Numerical Example of the SM Scheme

6.1. Description of Data.

In January of 1970 the Centers of Disease Control initiated a program, the National Nosocomial Infection Survey (NNIS), to coordinate the surveillance, collection and analysis of nosocomial infections data from hospitals across the United States. The hospitals participated voluntarily and agreed to follow guidelines set down by CDC in conducting prospective clinical surveillance of their patients. In all, 84 hospitals in 31 states joined NNIS. The data that was collected for CDC included dates of admission and onset of infection, the site of the infection and the responsible pathogens. The accumulated totals were sent to CDC monthly along with the number of patients discharged from the clinical services in which patients were monitored. Each report of a nosocomial infection was reviewed at CDC by a physician-epidemiologist.

In June 1970 a manufacturer ('Source A') began distributing contaminated dextrose-containing intravenous (IV) fluids. Soon there was an increase of Erwinia and Enterobacter bacteremia, but this was not detected by NNIS using the techniques available to it at the time. It was not until December 1970 that CDC received its first report of an outbreak and began investigating. The source was located and the contaminated fluids recalled as of March 21, 1971.

The data were later reanalyzed (Goldmann, Dixon, Fulkerson, Maki, Martin and Bennett, 1978) to see if the NNIS data did contain evidence of this increase. It was decided that the NNIS data did reflect the increase of nosocomial infections due to the guilty pathogen but that the techniques NNIS used were not sufficient to detect the epidemic.

Since this outbreak occurred near the beginning of NNIS, the roster of participating hospitals changed somewhat during this time. However, there was a subgroup of regularly reporting hospitals defined by two criteria; having reported in at least 13 of the 19 months under study and having reported in either February or March 1971 and either April or May 1971. The second criterion insures bracketing the recall date, within a month, with a report from the hospital. Of the 84 hospitals in NNIS, 49 qualified as regularly reporting hospitals. The 49 regularly reporting hospitals were divided into three groups based on their answers to a questionnaire survey about their IV fluid use during the study period. The three groups are hospitals which only used IV fluid from Source A, called Group A ($n=18$); hospitals which only used IV fluid from sources other than Source A, called Group O ($n=19$); and hospitals which used IV fluids from multiple sources including Source A, called Group M ($n=12$). In the reanalysis of this data in 1978 it was determined that 10 of the Group A hospitals reported one or more cases of Enterobacter or Erwinia bacteremia while only two of the Group O hospitals did this during the epidemic period. Using a two-tailed Fisher's Exact Test, this difference in counts is significant at the $p=0.005$ level.

The NNIS data which is still available from the period covered by the IV fluid epidemic is the number of cases per month of Enterobacter and Erwinia in the period from January 1970 to July 1971 and the estimated

incidence rate of cases per 10,000 discharges. In general, the monthly hospital discharges are fairly constant so the numbers of cases per month are proportional to the monthly incidence rates per 10,000 discharges. The numbers of cases per month for Group A and Group O are presented below in Table 6.1.

Table 6.1*

Number of Cases of Enterobacter and Erwinia in Group A and Group O Hospitals

January 1970 to July 1971

Year	1970												1971						
Month	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	A	M	J	J
Group A	0	1	3	1	0	3	5	6	10	4	6	10	6	21	28	1	1	0	1
Group O	2	1	2	1	1	0	2	1	0	3	5	2	1	0	0	3	3	5	1

*Source: Goldmann, Dixon, Fulkerson, Maki, Martin and Bennett (1978).

6.2. Application and Results for SM Scheme.

The data from this epidemic will be used here towards two purposes. The first purpose is to see how the SM scheme performs with data from an actual database built from surveillance data. The 'problems' of real data are immediately evident. The rate of infection does not take a single jump upward under epidemic conditions but rather appears to keep growing until a single jump downward does occur. However, it is an interesting example and has been described in the literature. A second purpose will be to see the effects that different memory lengths have on an SM scheme. SM schemes with memory lengths from one month to five months can be used on these

data. The beginning of the epidemic is considered to be June 1970, since the contaminated IV fluid was first distributed then. Thus, there are five months of non-epidemic data available. As s increases from one to five, the cases in memory initially are the cases from May 1970, back through January 1970.

The run lengths for the SM schemes will be presented with and without randomized testing. When the testing is not randomized then the test that will signal an increase in a certain month is done at an error rate that is equal to or less than the nominal α. Thus, the schemes with different memory lengths are likely to be operating with different expected run lengths under non-epidemic conditions and it will be difficult to equate them all. The expected run length under regular testing is just the month that detection occurs in since rejection of non-epidemic conditions occurs with a probability of one in that month and a probability of zero in all other months under study. When randomized testing is used, each of the schemes may have a non-degenerate probability distribution for its run length. This is because the decision of whether to reject may have a probability (other than zero or one) associated with it and the scheme needs to continue with the complement probability of not rejecting at that time. Thus, the SM scheme under these testing conditions will have an average run length. The run lengths and average run lengths for five memory lengths, $s=1, 2, 3, 4, 5$, and four error rates, $\alpha=.005, .01, .05, .1$, for both testing methods are presented in Table 6.2 when the only data that the scheme uses are from the Group A hospitals.

Table 6.2

Average Run Length With Randomized Tests and

Month of Detection (after May 1970) without Randomized Tests

for Memory Lengths=1,2,3,4,5

Using Contaminated IV Fluid Epidemic Data from Group A Hospitals

Testing Method	Nominal α	Memory 1	2	3	4	5
Random	0.005	8.7	9.0	9.0	8.3	4.0
	0.01	8.4	9.0	9.0	5.5	4.0
	0.05	5.8	3.9	2.0	2.5	1.8
	0.10	2.6	1.1	1.8	1.7	1.3
Not Random	0.005	9	9	9	9	4
	0.01	9	9	9	9	4
	0.05	9	9	2	4	4
	0.10	9	2	2	2	2

Even the SM scheme with memory of only one week at $\alpha=0.005$ does not let the epidemic pass beyond the ninth month without detection. This is the time when the epidemic greatly accelerates and the infection rate increases rapidly. Thus, if the epidemic infection rate had stayed constant at the rate in the period from June 1970 to January 1971 that SM scheme might not have detected this epidemic. The SM scheme with memory length of 5 months does very well even at the strict error rate of $\alpha=0.005$ by detecting the epidemic at the fourth test. Thus, the epidemic is detected at the fourth month (September 1970) with a false alarm signalled only once every 16.67 years.

Because of the fluctuation in the number of cases in the memory, the average run length does not monotonically decrease as the memory length increases for a given error rate. This is true under both randomized and regular testing. Thus, schemes with longer memory lengths can retain in memory the cases from months in which there are large numbers of cases whereas the shorter memory schemes would have already dropped those data before a larger increase occurs. This can result in the shorter memory scheme having a higher detection probability than the longer memory scheme at that point in time.

Without randomized testing the schemes are operating at different expected run lengths under non-epidemic conditions and are therefore harder to compare. The actual error rates at each test for the SM schemes operating on Group A data under regular testing are presented below in Table 6.3. Each entry in the table is the largest error rate less than or equal to the nominal alpha for a binomial test conditional on the total number of cases in memory plus the current week. It is clear that for some of the tests the actual α is far less than the nominal α. In some cases the actual alpha is zero which says that the scheme can not ever reject with the total number of cases from memory and the current week.

The run length distributions for the five SM schemes and four error rates when randomized testing is used are presented in Table 6.4 for the data from Group A hospitals. Each entry is the probability of rejection in that month, starting at June 1970, for that SM scheme. When randomized testing is used, the average run length does not give the entire picture since, for example, when $s=1$ and $\alpha=0.005$ there is a small but non-negligible probability of detection at the first test and the next nonzero probability of

Table 6.3

Actual Error Rates for Each Non-randomized Test

for Memory Lengths=1,2,3,4,5

Using Contaminated IV Fluid Epidemic Data from Group A Hospitals

Nominal α	s	Test 1	2	3	4	5	6	7	8	9
.005	1	.0000	.0039	.0005	.0021	.0009	.0010	.0021	.0021	.0030
	2	.0000	.0026	.0040	.0018	.0037	.0037	.0037	.0035	.0029
	3	.0013	.0013	.0022	.0021	.0034	.0015	.0027	.0015	.0047
	4	.0012	.0039	.0042	.0038	.0050	.0044	.0025	.0025	.0029
	5	.0046	.0024	.0011	.0047	.0037	.0022	.0050	.0023	.0041
.01	1	.0000	.0039	.0059	.0021	.0065	.0010	.0021	.0021	.0096
	2	.0000	.0026	.0040	.0068	.0037	.0037	.0037	.0035	.0075
	3	.0013	.0100	.0022	.0072	.0034	.0052	.0082	.0052	.0047
	4	.0012	.0039	.0042	.0038	.0050	.0044	.0070	.0070	.0074
	5	.0046	.0024	.0053	.0047	.0037	.0067	.0050	.0064	.0097
.05	1	.0000	.0352	.0327	.0384	.0287	.0107	.0384	.0384	.0261
	2	.0123	.0197	.0174	.0212	.0376	.0376	.0376	.0327	.0384
	3	.0129	.0489	.0383	.0213	.0297	.0401	.0216	.0401	.0252
	4	.0104	.0194	.0181	.0362	.0391	.0327	.0424	.0424	.0366
	5	.0307	.0127	.0206	.0447	.0333	.0459	.0315	.0376	.0435
.10	1	.0000	.0352	.0327	.0384	.0898	.0547	.0384	.0384	.0610
	2	.0123	.0879	.0576	.0557	.0919	.0919	.0919	.0787	.0754
	3	.0706	.0489	.0383	.0547	.0713	.0909	.0507	.0909	.0962
	4	.0563	.0726	.0611	.0892	.0900	.0746	.0889	.0889	.0721
	5	.0307	.0512	.0653	.0447	.0815	.0459	.0682	.0790	.0822

Table 6.4

Run Length Distribution of SM Schemes with Five Memory Lengths Under Randomized Testing using Group A Hospitals

α	s	R 1	2	3	4	5	6	7	8	9
.005	1	.040	0	0	0	0	0	0	0	.960
	2	0	0	0	0	0	0	0	0	1
	3	0	0	0	0	0	0	0	0	1
	4	0	0	0	.137	0	0	0	0	.863
	5	0	0	0	1	0	0	0	0	0
.01	1	.080	0	0	0	0	0	0	0	.920
	2	0	0	0	0	0	0	0	0	1
	3	0	.0001	0	0	0	0	0	0	.9999
	4	0	0	0	.703	0	0	0	0	.297
	5	0	0	0	1	0	0	0	0	0
.05	1	.400	0	0	0	0	0	0	0	.600
	2	.381	.275	0	0	0	0	.078	0	.266
	3	0	1	0	0	0	0	0	0	0
	4	0	.576	.315	.109	0	0	0	0	0
	5	.186	.790	.016	.008	0	0	0	0	0
.10	1	.800	0	0	0	0	0	0	0	.200
	2	.887	.113	0	0	0	0	0	0	0
	3	.170	.830	0	0	0	0	0	0	0
	4	.298	.702	0	0	0	0	0	0	0
	5	.666	.334	0	0	0	0	0	0	0

detection does not occur until the ninth test. Even when $\alpha=0.10$ there is a significant probability of detection at the ninth test when $s=1$. When the memory length is longer and the error rate is larger then the range of values of the run length distributions that have a nonzero probability is much tighter.

An equivalent set of analyses are done on the data combined from Group A and Group O hospitals. The baseline rate in the Group O hospitals did not change due to contaminated IV fluid. Also, the baseline rate in Group O is larger than the baseline rate in Group A thus partially masking the initial increase that occurs only in the Group A hospitals. Thus, the effect of combining the data from the two groups of hospitals will be to make the detection of the epidemic by the SM scheme more difficult.

The run lengths and average run lengths for five memory lengths, $s=1, 2, 3, 4, 5$, and four error rates, $\alpha=.005, .01, .05, .1$, for both testing methods are presented in Table 6.5 when Group A and Group O hospitals are combined.

When $\alpha=0.005$ and $s=1$ or $s=2$ the SM scheme does not detect the increase before test number 15. Thus, the entries for $\alpha=0.005$ and $s=1, 2$ as shown in Table 6.5 are the minimum possible values for the average run length and the month of detection. In general, the SM scheme has more difficulty detecting the epidemic than when only data from Group A hospitals are used. In the two cases mentioned above, the epidemic has a non-zero probability of not being detected by this scheme before it has decreased to pre-epidemic levels.

Table 6.5

Average Run Length With Randomized Tests and

Month of Detection (after May 1970) without Randomized Tests

for Memory Lengths=1,2,3,4,5 Using Contaminated IV Fluid Epidemic Data

from Group A and Group O Hospitals Combined

Testing Method	Nominal α	Memory 1	2	3	4	5
Random	0.005	10.7*	12.1*	9.7	9.0	9.0
	0.01	9.0	8.7	7.4	9.0	7.7
	0.05	9.0	2.0	2.0	2.1	2.0
	0.10	5.5	2.0	2.0	2.0	2.0
Not Random	0.005	15*	15*	10	9	9
	0.01	9	9	9	9	9
	0.05	9	2	2	4	2
	0.10	9	2	2	2	2

* Minimum possible value for this entry.

Table 6.6

Run Length Distribution of SM Schemes with Five Memory Lengths Under Randomized Testing using Group A and Group O Hospitals

α	s	R 1	2	3	4	5	6	7	8	9	10	R ≥ 15
.005	1	0	0	0	0	0	0	0	0	.712	0	.288
	2	0	0	0	0	0	0	0	0	.196	.345	.459
	3	0	0	0	0	0	0	0	0	.316	.684	0
	4	0	0	0	0	0	0	0	0	1	0	0
	5	0	0	0	0	0	0	0	0	1	0	0
.01	1	0	0	0	0	0	0	0	0	1	0	0
	2	0	.039	0	0	0	0	0	0	.961	0	0
	3	0	.233	0	0	0	0	0	0	.767	0	0
	4	0	0	0	0	0	0	0	0	1	0	0
	5	0	0	0	.251	0	0	0	0	.749	0	0
.05	1	0	0	0	0	0	0	0	0	1	0	0
	2	0	1	0	0	0	0	0	0	0	0	0
	3	0	1	0	0	0	0	0	0	0	0	0
	4	0	.964	.012	.024	0	0	0	0	0	0	0
	5	0	1	0	0	0	0	0	0	0	0	0
.10	1	.150	.329	0	0	0	0	0	0	.521	0	0
	2	0	1	0	0	0	0	0	0	0	0	0
	3	0	1	0	0	0	0	0	0	0	0	0
	4	0	1	0	0	0	0	0	0	0	0	0
	5	0	1	0	0	0	0	0	0	0	0	0

The run length distributions for the five SM schemes and four error rates when randomized testing is used are presented in Table 6.6 for the data from Group A and Group O hospitals. Each entry is the probability of rejection in the month, starting at June 1970, for that SM scheme. For all but the two combinations of parameters mentioned above, when α is small the SM scheme detects the epidemic at least by test number 10. Except for $s=1$, for $\alpha=.05$ and $\alpha=.10$, under randomized testing, the expected run length is two. This is also true for the month of detection under regular testing except for $\alpha=.05$ when $s=2$. The SM scheme operating on data from the Group A and Group O hospitals combined is never more powerful and is usually less powerful than when it is operating on data from only the Group A hospitals. This is true despite the increase in the sample size for the conditional binomial test that results from incorporating Group O data. The effect is to increase the rate under non-epidemic conditions but not to contribute to an increase in the rate under epidemic conditions. However, for $\alpha=.05$ and $s=2$, 3, 4 and 5, the expected run length is 2, which is small, with probability equal to one but with the cost of a false alarm every 1.67 years, on average.

Chapter VII

Conclusions and Suggestions for Further Research

7.1. Introduction.

This dissertation develops a model which is useful for the surveillance of epidemics by detecting an increase in the mean of a Poisson probability distribution. Beginning with a test that compares two Poisson means by conditioning on the total number of cases, a scheme is built that compares the number of 'new' cases to the number of 'old' cases and then incorporates the new data into memory. The length of the memory, s, is one of the parameters of the scheme and determines the order of the Markovian property of the scheme. This s^{th} order Markov property results in dependencies between tests. Since it was expected that the chosen s would be small the scheme was dubbed the Short Memory (SM) scheme. Tests are done at fixed intervals that compare the most current data to the data stored in memory. A scheme continues until an alarm is sounded which signals an increase in the underlying rate. The type of increase under investigation was characterized by a single jump in the baseline rate. It is always possible to determine the unconditional error probability of a test independent of the baseline rate.

7.2. Summary.

The sequential testing gives rise to a random variable, R, which is the

run length of the scheme. R represents the time to the first detection of an increase in the baseline rate of the underlying probability distribution. The probability distribution of R was shown to be decreasing with a mode at $R=1$ and with a geometric tail. Under non-epidemic conditions, conditioning on $R>s$, R has a geometric distribution with parameter $\frac{Pr(R>s+1)}{Pr(r>s)}$. Under epidemic conditions, conditioning on $R>2s$, R has a geometric distribution with parameter $\frac{Pr(R>2s+1)}{Pr(r>2s)}$.

The conditional binomial tests are randomized to achieve an exact error probability for the sake of making comparisons more meaningful. A continuous approximation to the distribution of cases was examined using the Gamma probability distribution. A test was proposed that used the χ^2 probability distribution and a sequence of F tests. The probabilities of the run length distribution were then calculated from the joint distribution of independent non-identically distributed Gamma random variables. Since this could lead to decisions being made on the basis of less than a whole case it was concluded that the approximation was not justified. The effect of ignoring the randomization weights was studied for the case of $s=1$, $\lambda_0=1$ and $\gamma=2$, 4 and 5. Ignoring the weights for the case when $n=0$ has a moderate effect on E(R) for small λ. This effect diminishes as λ increases since the probability that $n=0$ becomes smaller. The effect of ignoring all of the randomization weights is quite severe in the sense that the expected run lengths are much larger than desired for $\alpha=0.05$. However, the SM scheme shows some resilience to this. For example, while $E(R|H_0)$ increases from 19.4 to 725.1, $E(R|H_1)$ only increases from 10.1 to 29.4. This non-randomized scheme may be viewed as useful since the difference between the

expected run length under non-epidemic and epidemic conditions is so much greater than it is when the tests are randomized.

For the case of $s=1$, it was shown that for the SM scheme the expected run length under nonepidemic conditions is smaller than if the run length distribution is geometric with parameter $p=1-\theta=Pr(R=1)$. This is because of a negative correlation between the first two tests resulting in $\theta_1^2>\theta_2$, e.g. $[Pr(R=1)]^2 > Pr(R>2)$. However, the expected run length is not smaller by more than θ_1. The SM scheme is compared to a testing scheme where the tests are done every two weeks. This is actually two sub-schemes each with probability equal to one-half of occurring. The tests in each sub-scheme are independent since they do not share data as in the SM scheme. It is plausibly argued that this independent testing scheme should not perform as well as the SM scheme. This is born out in the numerical examples presented. The effect of ignoring all of the randomization weights in the independent scheme is also quite large but without the redeeming quality mentioned above for the SM scheme. That is, E(R) increases dramatically under both non-epidemic and epidemic conditions if the weights are ignored.

The SM scheme with longer memory is examined. The likelihood ratio test of the ratio of two Poisson means is taken to its limit as $s \rightarrow \infty$. The limit is found to be the likelihood ratio test of a Poisson mean against a known population parameter. When the second test is performed it has the same form as the first test since adding one random variable to an infinite sum of random variables will not change the test. This is true for any finite number of random variables so the dependencies between tests break down as s increases. Thus, in the limit, R has a geometric distribution with parameter $p=Pr(R=1)$. Additional memory affords the opportunity of using the

available information in different ways. When $s=2$, there are four testing schemes which were investigated. The expected run lengths were derived for the four plans listed; 1) the standard SM scheme which compares the current week's data to the two previous weeks' data; 2) an abbreviated SM scheme which compares the current week's data to the single previous week's data; 3) a lagged SM scheme which compares the current week's data to one previous week's data which is the second week back in time; and 4) an independent scheme which has three sub-schemes where each sub-scheme tests once every three weeks and each week has exactly one of the sub-schemes testing. The expected run lengths are difficult to compare theoretically because they contain so many parameters. Two examples are shown with $\lambda=1$ and $\gamma=4$ and 10. When $\gamma=4$ the standard SM has the smallest $E(R|H_1)$ and the lagged SM scheme is very close. The other two testing schemes do not do as well although the abbreviated SM scheme does better than the independent scheme. For $\gamma=10$, $E(R|H_1)$ is esentially equal for the three SM schemes (standard, abbreviated and lagged). The independent testing scheme is clearly less useful than the other testing schemes when $\gamma=10$. The SM scheme has been studied for small s and in the limit. These are both extreme situations. The SM scheme needs to be studied for moderate and large values of s.

The usefulness of the CUSUM has been called into question when the underlying rate is not known or can not be accurately estimated. This is the situation where the SM scheme affords protection since $E(R|H_0)$ can be set independently of λ_0 and will not be exceeded by $E(R|H_1)$. Direct comparisons of the CUSUM scheme and the SM scheme showed that the CUSUM was much more powerful for detecting large increases in small means and for detecting small increases in moderate means when s is small. Assuming that the true

baseline mean is not the mean that the CUSUM is set up for leads to a situation where the CUSUM parameters are misspecified for the true means. Using a mean in the CUSUM that is not the 'assumed' mean causes an increase in $E(R|H_1)$ that will surpass $E(R|H_0)$ when the misspecification of the true mean is more than $1/\gamma$-fold. For the examples shown, the worst misspecifications of λ_0 were only 50—70% which seem like realistic values to consider. The CUSUM is robust as $E(R|H_1)$ approaches $E(R|H_0)$ steadily from below but once it surpasses it, $E(R|H_1)$ increases quite rapidly. In spite of the SM scheme faring poorly when s is small in the direct comparisons with the CUSUM, the SM scheme can do much better if the memory is allowed to increase indefinitely.

The example from the epidemic of intravenous data in 1970 that was investigated by NNIS shows that the SM scheme performs well. The baseline rate does not experience a single jump that constitutes the increase but rather it increases many times over the nine months of the epidemic period. However, when the scheme performs at its greatest efficiency, it does pick up the first large increase in the epidemic period. The randomization weights play a role since, for some combinations of s and α, ignoring them does lengthen the average time to detection. The average time to detection does not decrease monotonically as the length of the memory becomes longer, presumably because of the fluctuation in the number of cases in each of the time periods.

7.3. Characterizing an Epidemic.

It has been assumed throughout that an epidemic was characterized by a one time jump in the baseline rate. This situation is a very difficult one

for the SM scheme to deal with. This is especially true for small s where the single jump epidemic data quickly becomes all of the data in memory. The epidemic situation could be characterized by a weekly γ-fold increase up to some maximum rate or by a situation where the increase was an exponential function that eventually leveled off. Both of these are realistic situations and would improve the prospects of the SM scheme when s is small. Suppose, for example, that $s=1$ and there was a γ-fold increase each week until some point in time where the increases ceased. Then all of the tests up to that time point will compare a new mean, which represents a γ-fold increase over the old mean, to the old mean. While $E(R|H_0)$ is unaffected, $E(R|H_1)$ is smaller under this type of epidemic since the run length distribution consists of larger detection probabilities for the smaller values of R. The power of the first CBT, given n_1, is the same as before, namely,

$\sum_{i_1=0}^{n_1}\binom{n_1}{i_1}\left(\frac{\gamma}{s+\gamma}\right)^{i_1}\left(\frac{s}{s+\gamma}\right)^{n_1-i_1}\alpha_{n_1-i_1,i_1}$. The power of the second CBT, given n_2,

is $\sum_{i_2=0}^{n_2}\binom{n_2}{i_2}\left(\frac{\gamma^2}{s+\gamma+\gamma^2}\right)^{i_2}\left(\frac{s+\gamma}{s+\gamma+\gamma^2}\right)^{n_2-i_2}\alpha_{n_2-i_2,i_2}$. This pattern continues for

as long as the γ-fold increases continue.

7.4. The Underlying Distribution.

The underlying distribution of cases was assumed to be Poisson which led to the conditional binomial test for comparing two Poisson means. If the underlying distribution is not Poisson then the test performed each week will be different. For example, suppose that the number of cases in a week follows a geometric probability distribution with parameter p. Then,

$Pr(Y=y)=p(1-p)^y$. Let

$$X=\sum_{j=1}^{s} X_i \text{ where } Pr(X_i=x_i) = p(1-p)^{x_i}.$$

If the X_i's are independent, X has a Negative Binomial pdf where

$$Pr(X=x) = \binom{x+s-1}{s-1} p^s (1-p)^x.$$

Similarly, if X and Y are independent,

$$Pr(X+Y=x+y) = \binom{x+y+s}{s} p^{s+1} (1-p)^{x+y}.$$

Thus,

$$Pr(Y=y|X+Y=x+y) = \frac{Pr(Y=y)\ Pr(X=x)}{P(X+Y=x+y)} = \frac{\binom{x+s-1}{s-1}}{\binom{x+y+s}{s}}.$$

A UMP test exists for testing the hypothesis that the ratio of the two distribution parameters equals one against the hypothesis that the ratio is less than one. For $s=1$, under null conditions this conditional probability distribution is seen to be Uniform on the integers in the set $\{0, x+y\}$ where $Pr(Y=y|X+Y=x+y)=\frac{1}{x+y+1}$. In this setting, the forms of the run length distributions and the expected run lengths will not change since their derivations depend only on the assumptions made about the SM scheme and not on the assumption about the underlying distribution of cases. Of course, the probabilities, θ_i and β_i, will be calculated differently since this conditional distribution dictates that a test different from the conditional binomial test be performed each week.

An examination of different underlying distributions of cases should include the situation when the observations are not independent. Cases are expected to occur in clusters under epidemic conditions. In the IV fluid epidemic, the cases were not passing the illness to uninfected patients but the contaminated fluid was sent to certain hospitals and presumably was administered to patients randomly. But, if one patient becomes ill there is a

higher chance that other patients will become ill since this indicates that a source of infection is present. If the disease under surveillance is contagious then cases are not expected to occur independently of one another. Likewise, if there are no cases then the proabaility of not seeing any cases should increase. If there is nonindependence in the observations then the expected run length should decrease under non-epidemic conditions. Under epidemic conditions, the expected run length should decrease if seeing a few cases in one week means that there will be many cases in the following week. This is a similar situation to what was described above for increases other than a single jump.

7.5. Conclusions.

The SM scheme shows certain desirable properties that make it useful as a surveillance system. When s is small and the baseline rate is small, the SM scheme is not as powerful as the CUSUM scheme. However, it does afford a protection against the baseline rate being estimated inaccurately or when it is simply unknown. The standard SM scheme does better than the modified SM schemes and all of the SM schemes do better than the independent testing schemes that they were compared to. Thus, the data is being used more efficiently by the SM scheme than some obvious competitors.

The three requirements that were stated as a description of a good surveillance model will be addressed. It is seen that the SM scheme can be used to detect an increase in a rare event even when the increased number of cases is rare. As stated above, it is not as powerful as the CUSUM if the disease is very rare and s is small. A single small increase in consecutive

time periods does not have the effect of one large increase unless there are consecutive small increases in the time periods. However, as s increases, the effect of not detecting the increase in any single test is diminished. The framework is completely adaptable to breaking the data into subsets or combining data from several sources as was shown in Chapter VI. The SM scheme seems to address a different concern, namely, to provide protection when the baseline rate is unknown. Thus, the SM provides for a situation where there are not many alternatives now being used in the surveillance of diseases.

BIBLIOGRAPHY

Armitage, P. (1950). Sequential analysis with more than two alternative hypothesis and its relation to discriminant analysis. Journal of the Royal Statistical Society B , 137-144.

Bagshaw, M., and Johnson, R.A. (1975). The effect of serial correlation on the performance of CUSUM tests II. Technometrics 17 , 73-80.

Bagshaw, M., and Johnson, R.A. (1975). The influence of reference values and estimated variance on the ARL of CUSUM tests. Journal of the Royal Statistical Society B , 413-420.

Barbujani, G., and Calzolari, E. (1984). Comparison of two statistical techniques for the surveillance of birth defects through a Monte Carlo simulation. Statistica in Medicine 3 , 239-247.

Barnard, G.A. (1959). Control charts and stochastic processes. Journal of the Royal Statistical Society B , 239-270.

Bartlett, M.S. (1953). Proceedings of the Cambridge Philosophical Society 49.

Birnbaum, A. (1954). Statistical Methods for Poisson Processes and Exponential Populations. Journal of the American Statistical Association 49 , 254-266.

Bjerkedal, T., and Bakketeig, L.S. (1975). Surveillance of congenital malformations and other conditions of the newborn. International Journal of Epidemiology 4 , 31-36.

Brook, D., and Evans, D.A. (1972). An approach to the probability distribution of CUSUM run length. Biometrika 59 , 539-549.

Chen, R. (1978). A surveillance system for congenital malformations. Journal of the American Statistical Association 73 , 323-327.

Chen, R. (1978). Statistical techniques in birth defects surveillance systems. Contributions to Epidemiology and Biostatistics 1 , 184-189.

Chen, R. (1987). The relative efficiency of the SETS and the CUSUM techniques in monitoring the occurrence of a rare event. Statistics in Medicine 6 , 517-525.

Chen, R. (1985). Letter to the Editor: Comparison of two statistical techniques for the surveillance of birth defects through a Monte Carlo simulation, Barbujani, G. and Calzolari, E., Statistics in Medicine 4 , 389-390.

Chen, R., Mantel, N., Connely, R.R., and Isacson, P. (1982). A monitoring system for chronic diseases. Methods of Information in Medicine 21 , 86-90.

Cox, D.R., and Lewis, P.A.W. (1966). The Analysis of Series of Events. London: Methuen.

Ewan, W.D., and Kemp, K.W. (1960). Sampling inspection of continuous processes with no autocorrelation between successive results. Biometrika 47, 363-380.

Feller, W. (1970). An Introduction to Probability Theory and Its Application Volume I , 3rd edition. New York: John Wiley & Sons.

Gallus, G., Mandelli, C., Marchi, M., and Radaelli, G. (1986). On surveillance methods for congenital malformations. Statistics in Medicine 5 , 565-571.

Goldmann, D.A., Dixon, R.E., Fulkerson, C.C., Maki, D.G., Martin, S.M., and Bennett, J.V. (1978). The role of nationwide nosocomial infection surveillance in detecting epidemic bacteremia due to contaminated intravenous fluids. American Journal of Epidemiology 108 , 207-213.

Goldsmith, P.L., and Whitfield, H. (1961). Average run lengths in cumulative sum quality control schemes. Technometrics 3 , 11-20.

Haley, R.W. and Garner, J.S. Infection surveillance and control programs. Chapter 3, Hospital Infections. Bennett, J.V. and Brachman, P.S.; Boston: Little, Brown and Co., 2nd Edition.
Hill, G.B., Spicer, C.C., and Weatherall, J.A.C. (1968). The computer surveillance of congenital malformations. British Medical Bulletin 24 , 215-218.

Johnson, N.L. (1961). A simple theoretical approach to cumulative sum control charts. Journal of the American Statistical Association 56 , 836-840.

Johnson, R.A., and Bagshaw, M. (1974). The effect of serial correlation on the performance of CUSUM tests I. Technometrics 16 , 103-112.

Kemp, K.W. (1958). Formulae for calculating the operating characteristic and the average sample number of some sequential tests. Journal of the Royal Statistical Society B 20 , 379-396.

Kemp, K.W. (1961). The average run length of the cumulative sum chart when a v-mask is used. Journal of the Royal Statistical Society B 23 , 149-153.

Kemp, K.W. (1967). A simple procedure for determining upper and lower limits for the average sample run length of a cumulative sum scheme. Journal of the Royal Statistical Society B 29 , 263-265.

Kemp, K.W. (1971). Formal expressions which can be applied to CUSUM charts. Journal of the Royal Statistical Society B 3 , 321-360.

Kenett, R., and Pollak, M. (1983). On sequential detection of a shift in the probability of a rare event. Journal of the American Statistical Society 78 , 389-395.

Lehmann, E.L. (1986). Testing Statistical Hypothesis , 2nd edition. New York: John Wiley & Sons.

Nadler, S., and Robbins, N.B. (1971). Some characteristics of Page's two-sided procedure for detecting a change in a location parameter. Annals of Mathematical Statistics 42 , 538-551.

Page, E.S. (1954). Continuous Inspection Schemes. Biometrika 41 , 100-115.

Page, E.S. (1955). A test for a change in a parameter occuring at an unknown point. Biometrika 42 , 523-527.

Pollak, M. and Siegmund, D. (1985). A diffusion process and its applications to detecting a change in the drift of Brownian motion. Biometrika 72 , 267-280.

Reynolds, M.R. (1972). A sequential nonparametric test for symmetry with applications to process control. Technical Report No. 148, Department of Operations Research and Department of Statistics, Stanford University.

Reynolds, M.R. (1975). Approximations to the average run length in cumulative sum control charts. Technometrics 17 , 65-71.

Shewart, W.A. (1931). Economic control of quality of manufactured product. New York: Van Nostrand.

van Dobben de Bruyn, C.S. (1968). Cumulative Sum Tests: Theory and Practice. London: Griffin.

Wetherall, J.A.C., and Haskey, J.C. (1976). Surveillance of malformations. British Medical Bulletin 32 , 39-44.

Wetherill, G.B. (1977). Sampling Inspection and Quality Control. London: Metheun.

Woodward, R.H., and Goldsmith, P.L. (1964). Cumulative Sum Techniques, ICI Monograph No. 3, Oliver & Boyd.