

# The Null Steady-State Distribution of the CUSUM Statistic

O. A. Grigg <sup>†</sup>  
D. J. Spiegelhalter

June 4, 2007

## Abstract

We develop an empirical approximation to the null steady-state distribution of the cumulative sum (CUSUM) statistic, defined as the distribution of values obtained by running a CUSUM with no upper boundary under the null state for an indefinite period of time. The derivation is part theoretical and part empirical and the approximation is valid for CUSUMs applied to Normal data with known variance (although the theoretical result is true in general for exponential family data, which we show in the Appendix). The result leads to an easily-applied formula for steady-state p-values corresponding to CUSUM values, where the steady-state p-value is obtained as the tail area of the null steady-state distribution and represents the expected proportion of time, under repeated application of the CUSUM to null data, that the CUSUM statistic is greater than some particular value. When designing individual charts with fixed boundaries, this measure could be used alongside the average run-length (ARL) value, which we show by way of examples may be approximately related to the steady-state p-value. For multiple CUSUM schemes, use of a p-value enables application of a signalling procedure that adopts a false discovery rate (FDR) approach to multiplicity control. Under this signalling procedure, boundaries on each individual chart change at every observation according to the ordering of CUSUM values across the group of charts. We demonstrate practical application of the steady-state p-value to a single chart where the example data are number of earthquakes per year registering  $> 7$  on the Richter scale measured from 1900 to 1998. Simulation results relating to the empirical approximation of the null steady-state distribution are summarised, and those relating to the statistical properties of the proposed signalling procedure for multiple CUSUM schemes are presented.

*Keywords:* CUSUM (cumulative sum) steady-state distribution P-value FDR (false discovery rate) SPC (statistical process control) ARL (average run length) multiple charts

## 1 Introduction

The Page (1954) cumulative sum (CUSUM) chart is a well-known tool for statistical process control. The chart, in essence, sequentially tests an alternative hypothesis  $H_1$ , where the process is deemed ‘out of control’, against the null  $H_0$ , representing the ‘in control’ scenario. The cumulative log likelihood ratio (LLR) of  $H_1$  versus  $H_0$  is plotted at every observation and the test stops in favour of  $H_1$  when the LLR becomes large. To maintain sensitivity to the alternative  $H_1$ , and since accepting the null in the context of continuous monitoring makes little sense, the chart is restricted from falling below zero. In this way, credit (or inertia (Woodall & Mahmoud, 2005)) cannot be built up in favour of  $H_0$  and so any evidence in favour of  $H_1$  has a more immediate effect.

Figure 1 illustrates a CUSUM chart for some example data. The data are the number per year worldwide of earthquakes whose magnitudes are  $> 7$  on the Richter scale, measured from 1900 to 1998 (Hyndman, 2005). The null hypothesis,  $H_0$ , is here based on the behaviour of the series in the first 40 data points, from 1900 to 1939. Since the data are counts, we take the square root of the data with the aim of stabilising a variance that is proportional to the mean. The mean and variance (on the square-root scale) of the pilot data are 4.40

---

<sup>†</sup>Address for correspondence: MRC Biostatistics Unit, Institute of Public Health, Robinson Way, Cambridge UK. Email: Olivia.Grigg@mrc-bsu.cam.ac.uk

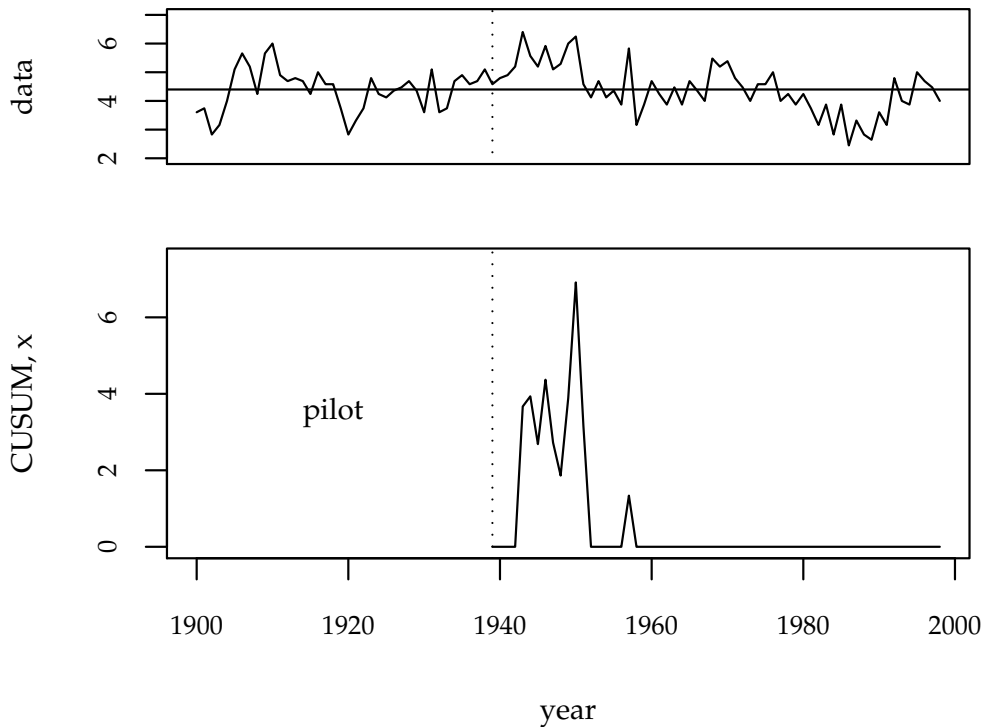


Figure 1: Lower panel: a CUSUM testing for a shift from  $N(4.40, 0.736)$  to  $N(6.61, 0.736)$  in the (square-rooted) number of earthquakes per year worldwide (with magnitude  $> 7$  on the Richter scale) from 1940 to 1998. The null  $N(4.40, 0.736)$  was generated from data in the period 1900—1939. Upper panel: the transformed data, including the pilot data.

and 0.74 respectively. The CUSUM drawn in the figure is sequentially testing the null hypothesis that the square-rooted data from 1940 onwards are  $N(4.40, 0.74)$  versus the alternative that they are  $N(6.61, 0.74)$ , corresponding to a 3 sd shift in mean.

Typically, the LLR is deemed to be large when it crosses a fixed threshold  $h > 0$  and, returning to Figure 1, one would usually see a threshold drawn at  $h$  on such a figure (see, for instance, Hawkins & Olwell (1997), p. 25). It does not make sense to consider eventual type I and II error probabilities associated with crossing such a threshold  $h$  and hence choosing  $H_1$  over  $H_0$  in an absolute sense, since the restriction of the chart means that  $H_1$  will be favoured at some point. Specifically, the type I error of the overall test (rejecting the null when it is in fact true) is necessarily 1 and the type II error (not accepting  $H_1$  in the case that it is true) is zero. Instead of type I and II errors, the distribution of the time taken for the chart to exceed  $h$  is usually considered (Hawkins & Olwell (1997), pp. 33—34). In particular, the average run length (ARL) under  $H_0$  and  $H_1$  are favourite operating characteristics to investigate and can be viewed as analogous to size and power in a standard hypothesis test. As with the size of a test, it is desirable for the rate of false alarms to be small and so for the ARL under  $H_0$  to be long. For a powerful CUSUM, we would like practically relevant changes to be detected quickly, and so for the ARL under  $H_1$  to be short. Approximate calculation of the ARL can be carried out using a Markov chain approach (Brook & Evans, 1972), by solving an integral equation (Goel & Wu, 1971) or by an empirical approximation (Siegmund, 1985). Routinely,  $h$  would be established using trial and error so that a value leading to an appropriate compromise between the ARL under  $H_0$  and that under  $H_1$  could be reached.

Here we derive the null steady-state distribution of the CUSUM statistic. This distribution does not depend

on a boundary  $h$  in any way. One important instance where consideration of such a distribution is useful is in the design of a multiple CUSUM scheme, where interpretation of operating characteristics is made more natural and straightforward: marginal, or ‘steady-state’ p-values associated with plotted CUSUM values can be calculated, and then fed into a cross-sectional multiplicity control procedure which can be implemented for an indefinite period of monitoring. The implicit thresholds on the charts would then vary across charts and time. The ‘steady-state’ p-values, which correspond to tail areas of the steady-state distribution of CUSUM values, can be interpreted as the expected proportion of time spent above a given CUSUM value over repeated experiments made under null conditions. For stand-alone charts, the p-value associated with a particular threshold  $h$  could be used to inform choice of  $h$  (in terms of expected frequency of false alarms) alongside the better known ARL-value. We show by way of examples how the ARL and p-value described here are approximately related through a simple formula.

Applying multiplicity control to a multiple CUSUM scheme has been considered before in the context of monitoring UK hospitals (Marshall *et al.*, 2004; Grigg & Spiegelhalter, 2005). Marshall *et al.* (2004) use a fixed boundary  $h$  on each chart, and the multiplicity control is active over a fixed period only. Grigg & Spiegelhalter (2005) apply a variant of the method described here. An alternative approach to both methods would be to choose  $h$  by trial and error so as to yield appropriate ARL values for the overall scheme. However, calculating an accurate approximation to the ARL for the overall scheme given a particular  $h$  can be computationally very intensive, particularly if the number of charts is large. Regardless of the approach taken to designing the scheme, definition of in- and out-of-control states of the entire system is less straightforward than in the single chart case: rather than all units being either in or out of control, the most likely scenario is that at any given time the majority of units are ‘in control’ while only a small proportion are not (they need not be the same ones over time). This model, where the main body of units are assumed ‘in control’, is utilised by Marshall *et al.* (2004) and Grigg & Spiegelhalter (2005) in their analyses of multiple CUSUM schemes.

Benjamini & Kling (1999) also describe p-values for the CUSUM, to facilitate visual understanding of the chart and allow for the use of multiplicity control techniques – in particular, false discovery rate (FDR) control (Benjamini & Hochberg, 1995), also used in the context of charting by Marshall *et al.* (2004) and Grigg & Spiegelhalter (2005). The p-value measure they adopt is the inverted ARL.

In Section 2, we present the proposed approximation to the steady-state distribution of the CUSUM statistic, which is valid for CUSUMs applied to approximately Normal data with known variance and illustrate how it may be approximately related to the ARL. We describe the development of the formula in detail in Section 3, and in Section 4 we return to the earthquake data of Figure 1 to demonstrate application of the formula to a single chart. In Section 5, a signalling procedure for multiple CUSUM schemes is put forward, which describes how signals can be raised within a system of CUSUM charts. The procedure involves the use of an FDR technique that powerfully picks out unusually small p-values across charts and time, where the p-values are calculated as tail areas of the approximated steady-state distribution. A simple proof is given to demonstrate that, under the proposed signalling procedure, the FDR (or expected proportion of false signals amongst all signals from charts) is indeed controlled at every time point and over time. Simulation results are presented in order to demonstrate the effective power of the approach in terms of time to true signals and to give additional properties of the sampling distribution of the observed proportion of false signals. In Section 6 some concluding remarks are made.

## 2 Approximation to the Null Steady-State Distribution $f_\delta(x)$

For Normal data  $y_1, \dots, y_t \sim N(\mu, 1)$ , a CUSUM testing the hypothesis  $H_0 : \mu = 0$  vs.  $H_1 : \mu = \delta$  can be written as

$$\begin{aligned} x_0 &= 0 \\ x_t &= \max(0, x_{t-1} + w_t), \quad t = 1, 2, \dots \end{aligned} \tag{2.1}$$

where  $w_t$  is the log-likelihood ratio weight at time  $t$ . Denoting the distribution of  $y_t$  under  $H_0$  to be  $\psi_0(y_t)$  and under  $H_1$  to be  $\psi_1(y_t)$  we can write

$$\begin{aligned} w_t &= \log \psi_1(y_t) - \log \psi_0(y_t) \\ &= \delta y_t - \frac{\delta^2}{2}. \end{aligned} \quad (2.2)$$

Testing a more general null hypothesis  $H_0 : y'_t \sim N(\mu_0, \sigma^2)$  vs.  $H_1 : y'_t \sim N(\mu_0 + \delta\sigma, \sigma^2)$  can be achieved by transforming to  $y_t = (y'_t - \mu_0)/\sigma$  and using the CUSUM weights given by (2.2). Hence,  $\delta$  can be interpreted as an amount (in standard deviations) that would constitute a ‘practically relevant’ shift in the underlying mean  $\mu$ .

Usually an absorbing barrier is placed at  $h$  so that the CUSUM signals as soon as  $x_t > h$ . Here, however, no boundary is placed and instead the stationary or steady-state distribution,  $f_\delta(x)$ , of values that could be generated by running the chart *with no boundary* for an indefinite period is considered. From (2.2) the weights  $w_t$  depend directly on  $\delta$  and hence so does the steady-state distribution  $f_\delta(x)$ . In the next section we use a combination of theory and simulation to show that, over a range of  $\delta$ ,

$$\begin{aligned} f_\delta(x) &= 1 - \gamma_0(\delta), & \text{if } x = 0 & \quad (2.3) \\ &\approx \frac{x - x_0(\delta)}{\sqrt{r^2(\delta) - \{x - x_0(\delta)\}^2}} \exp \left[ y_0(\delta) - \sqrt{r^2(\delta) - \{x - x_0(\delta)\}^2} \right], & \text{if } 0 < x \leq x'(\delta) \\ &= \gamma(\delta)e^{-x}, & \text{otherwise,} \end{aligned}$$

where  $x'$  is suitably large. Approximate functional forms for  $\gamma_0$ ,  $\gamma$  and  $x'$  are empirically derived and can be described by

$$\begin{aligned} \widehat{\gamma}_0(\delta) &= e^{-0.651(\delta-0.277)} + 0.031\delta - 0.189 & (2.4) \\ \widehat{\gamma}(\delta) &= e^{-0.578(\delta+0.024)} + 0.006\delta \\ \widehat{x}'(\delta) &= 0.170\delta^2 + 1.052\delta - 0.02. \end{aligned}$$

The other functions,  $r$ ,  $y_0$  and  $x_0$ , which are respectively the radius and origin coordinates of a circle segment (see Figure 5), can be calculated from  $\gamma_0$ ,  $\gamma$  and  $x'$  as follows. Let  $\rho_0 = \log(\gamma_0)$ ,  $\rho = \log(\gamma) - \rho_0$ ,  $u = \sqrt{x'^2 + (\rho - x')^2}$ ,  $v = \sqrt{2}x' - \rho/\sqrt{2}$  and  $\theta = \sin^{-1}(v/u)$ , then we have that

$$\begin{aligned} r &= \frac{u}{2 \sin\left(\frac{\pi}{2} - \theta\right)} & (2.5) \\ x_0 &= -r \sin\left(2\theta - \frac{3\pi}{4}\right) \\ y_0 &= \rho_0 - r \cos\left(2\theta - \frac{3\pi}{4}\right) \end{aligned}$$

A plug-in formula for steady-state p-values,  $\mathcal{P}_\delta(x) = \int_x^\infty f_\delta(x)dx$ , can be obtained from (2.3) as

$$\begin{aligned} \mathcal{P}_\delta(x) &= 1, & \text{if } x = 0 & \quad (2.6) \\ &\approx \exp \left[ y_0(\delta) - \sqrt{r^2(\delta) - \{x - x_0(\delta)\}^2} \right], & \text{if } 0 < x \leq x'(\delta) \\ &= \gamma(\delta)e^{-x}, & \text{otherwise.} \end{aligned}$$

To give a concrete example of how (2.3) through (2.6) can be used to approximate the steady-state distribution of CUSUM values or steady-state p-values for a given  $\delta$ , for  $\delta = 2$  we have that

$$\begin{aligned} \widehat{f}_2(x) &= 0.801, & \text{if } x = 0 & \quad (2.7) \\ &= \frac{x + 10.5}{\sqrt{352 - (x + 10.5)^2}} \exp \left[ -17.2 - \sqrt{352 - (x + 10.5)^2} \right], & \text{if } 0 < x \leq 2.76 \\ &= 0.322e^{-x}, & \text{otherwise.} \end{aligned}$$

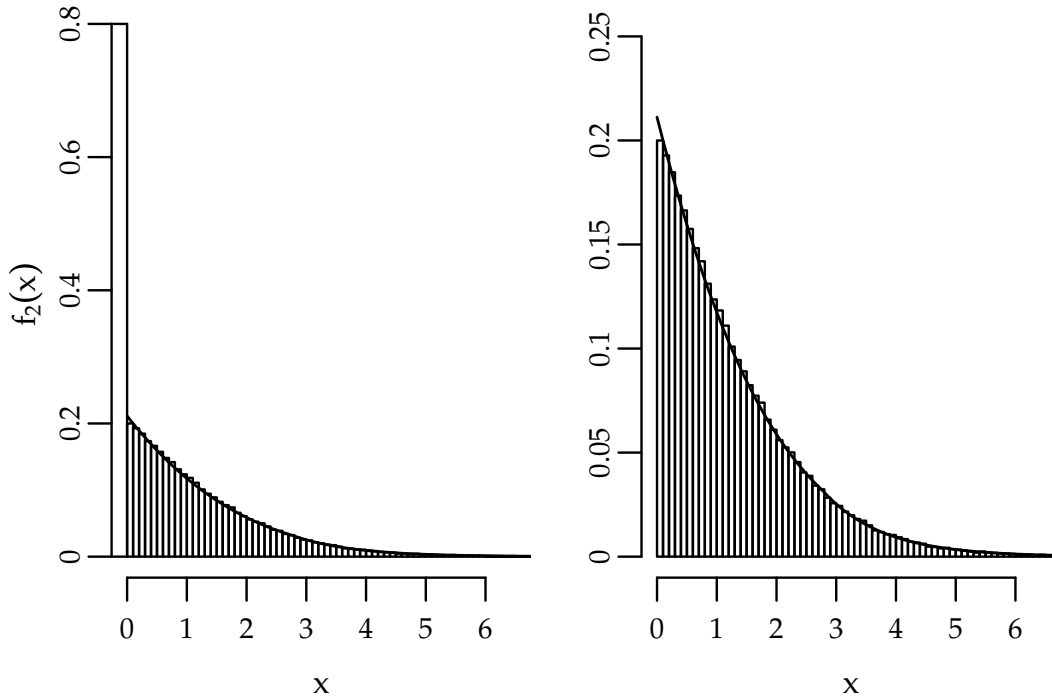


Figure 2: Simulated steady-state distribution  $f_2(x)$  for a CUSUM with  $\delta = 2$ , with the approximation given in (2.7),  $\hat{f}_2(x)$ , superimposed. The lump on zero of 0.801 can be seen in the lefthand panel; the righthand panel is ‘zoomed in’ on the main body of the distribution.

A plug-in formula for approximate p-values when  $\delta = 2$  is given by

$$\begin{aligned}
 \hat{\mathcal{P}}_2(x) &= 1, & \text{if } x = 0 \\
 &= \exp \left[ -17.2 - \sqrt{352 - (x + 10.5)^2} \right], & \text{if } 0 < x \leq 2.76 \\
 &= 0.322e^{-x}, & \text{otherwise.}
 \end{aligned} \tag{2.8}$$

Figure 2 displays the empirical steady-state distribution of the CUSUM for  $\delta = 2$ , based on  $1 \times 10^7$  simulated values for  $x$ . The approximation to  $f_2(x)$  given in (2.7) is also given. From the lefthand panel it can be seen that, for this value of  $\delta$ , the CUSUM is expected to spend in the long run 80% of the time at zero.

Figure 3 shows that the approximation described by (2.3) and (2.6) may be simply related to the average run length (ARL), or first passage time. The relationship appears to be

$$ARL_\delta(x) \approx \left( \frac{1 - \hat{\mathcal{P}}_\delta(x)}{f_\delta(0)} \right)^2 / \hat{\mathcal{P}}_\delta(x). \tag{2.9}$$

This also gives an approximate relationship between  $\hat{\mathcal{P}}_\delta(x)$  and the p-value measure proposed by Benjamini & Kling (1999),  $ARL_\delta^{-1}(x)$ . Since  $f_\delta(0)/\{1 - \hat{\mathcal{P}}_\delta(x)\} < 1$  for  $x > 0$ ,  $\mathcal{P}_\delta(x)$  presents as a less extreme p-value measure than  $ARL_\delta^{-1}(x)$ .

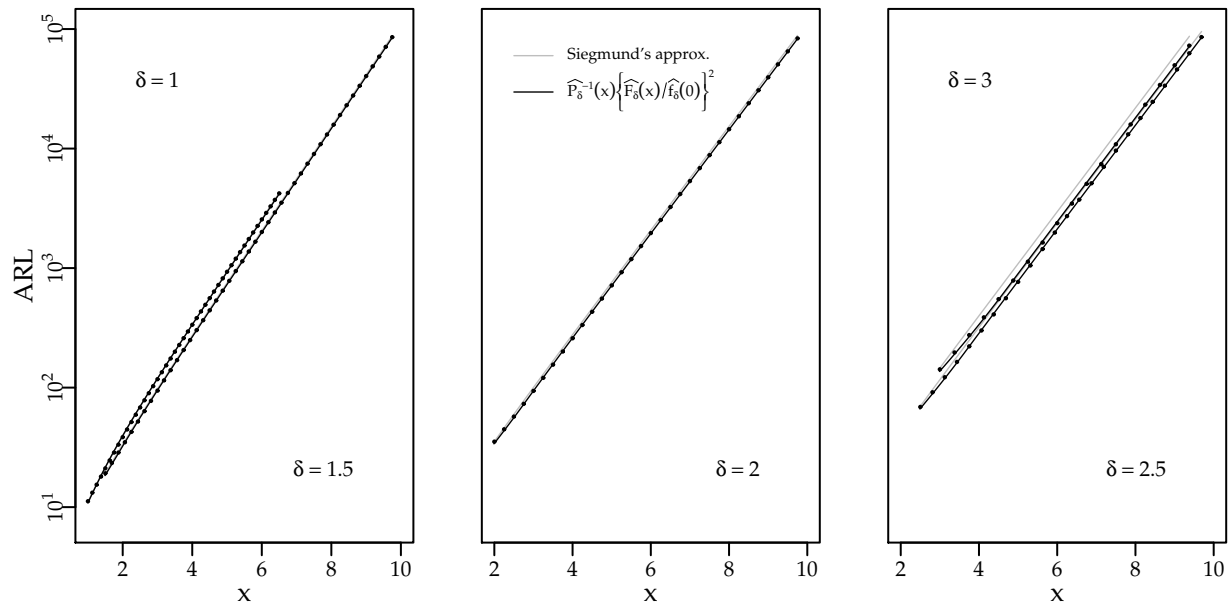


Figure 3: Approximate average run length (ARL) to a CUSUM value of  $x$  given the shift parameter  $\delta$ . Siegmund's (1985) approximation is also given. The ARL points are from Hawkins & Olwell (1997, p. 48).

### 3 Derivation of Approximation to $f_\delta(x)$

In this section we firstly show that the tail of  $f_\delta(x)$  is proportional to an exponential distribution with mean 1, denoted  $\text{Exp}(1)$ , and hence that its shape does not depend on  $\delta$ . The constant of proportionality,  $\gamma$ , does however depend on  $\delta$  and so we accordingly describe it as  $\gamma(\delta)$ . In the Appendix we show that a CUSUM on exponential family data in general will result in a steady-state distribution with an  $\text{Exp}(1)$  tail, regardless of the size of the shift  $\delta$  from the null natural parameter value.

To complete the approximation given the theoretical results for large CUSUM values, we consider the parameters  $\gamma_0(\delta)$ ,  $\gamma(\delta)$  and  $x'(\delta)$ . The value of  $\gamma_0(\delta)$  is equal to one minus the steady-state lump of probability at zero,  $1 - f_\delta(0)$ , or one minus the proportion of time the CUSUM spends at zero under long-run null conditions and is the value of the log p-value as it tends downwards towards zero. The value of  $x'(\delta)$  is the point beyond which CUSUM values are assumed large enough for the  $\text{Exp}(1)$  model to be reasonable.

Empirical forms for the three parameters  $\gamma_0(\delta)$ ,  $\gamma(\delta)$  and  $x'(\delta)$  were fitted through joint minimisation of the squared distance on the log scale of modelled p-values given by (2.6) away from simulated p-values and of the squared distance on the log scale of  $\gamma_0(\delta)$  away from one minus the simulated lump of probability on zero. This joint minimisation was carried out over a range of values for  $\delta$  and at a number of CUSUM values at certain realisations of  $\delta$ .

#### 3.1 Theoretical Results for Large CUSUM Values

The CUSUM, as it is described in (2.1), can be formally viewed as a random walk with negative drift under the null equal to  $\mathbb{E}[w_t|\mu = 0] = -\delta^2/2$ , and with a holding barrier at zero (Page (1954), Cox & Miller (1965), pp. 22–75, Sornette & Cont (1997)). Under the specified alternative,  $\mu = \delta$ , the chart has positive drift equal to  $\mathbb{E}[w_t|\mu = \delta] = \delta^2/2$  and so in that case does not reach a steady-state, or equilibrium distribution. For large  $x$ , say greater than some value  $x'(\delta)$ , the holding barrier at zero is unimportant in the sense that the CUSUM has practically no memory of it. Under these circumstances  $f_\delta(x)$  can be viewed as the solution to the Wiener-Hopf type integral equation (Cox & Miller (1965), p. 63, Feller (1966), p. 385; Sornette & Cont (1997))

$$f_\delta(x) = \int_{-\infty}^x f_\delta(x-w)\xi(w)dw \quad (3.1)$$

where, if the test data  $y$  are  $N(0, 1)$  (that is, are null data), the distribution  $\xi(w)$  of the weights is  $N(-\delta^2/2, \delta^2)$ .

It can easily be checked that the general solution to (3.1) can be given in the form

$$f_\delta(x) \propto \beta e^{-\beta x}, \quad \beta > 0 \quad (3.2)$$

where (Feller (1966), p. 387, Sornette & Cont (1997))

$$\int_{-\infty}^{\infty} \xi(w)e^{\beta w}dw = 1. \quad (3.3)$$

That is, under the null hypothesis and in the tail of the distribution,  $f_\delta(x)$  is proportional to a standard exponential distribution with mean  $1/\beta$ , where  $\beta$  is given by (3.3). This is true whatever the distribution of the weights  $w$ .

Recognising that the left-hand side of (3.3) is just the moment generating function of the weight  $w$ , and given that the distribution of  $w$  is  $N(-\delta^2/2, \delta^2)$ , we can show that

$$\frac{\delta^2\beta}{2}(\beta - 1) = 0. \quad (3.4)$$

Since  $\beta > 0$ , this implies that  $\beta = 1$ . In the Appendix, we show that  $\beta = 1$  for exponential family data in general, that is, where  $w$  is the log-likelihood ratio in favour of a shift in the natural parameter of an exponential family distribution.

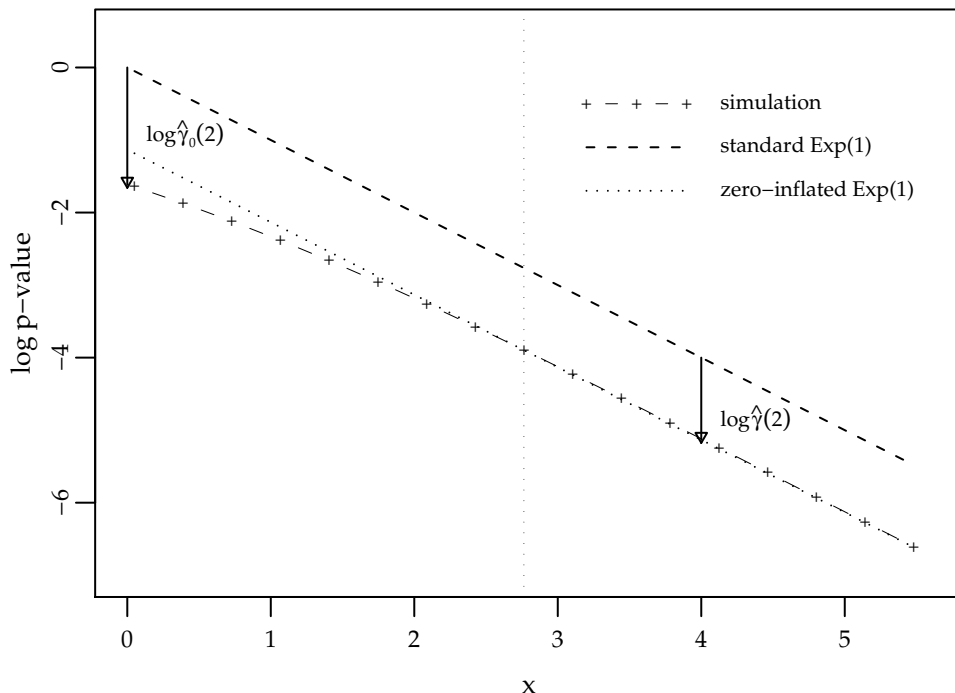


Figure 4: Simulated steady-state log p-values,  $\log \tilde{\mathcal{P}}_2(x)$ , for a CUSUM with  $\delta = 2$ . The approximate cut-off for when the zero-inflated Exp(1) model for  $f_2(x)$  becomes valid,  $\hat{x}'(2)$ , is indicated by a vertical dotted line.

Therefore, when the data  $y$  are  $N(0,1)$ ,  $f_\delta(x)$  is distributed as Exp(1) in the tail, up to a constant  $\gamma(\delta)$ . In other words, for large CUSUM values the steady-state distribution is a zero-inflated exponential distribution with rate parameter 1 and inflation factor  $1 - \gamma(\delta)$ . The inflation factor  $1 - \gamma(\delta)$  depends on  $\delta$ , since it depends on the behaviour of the chart close to its lower boundary of zero. More precisely it depends on the expected time spent at zero and the shape of the distribution for small  $x$  where  $0 < x < x'(\delta)$ .

### 3.2 Empirical Approach

Figure 4 illustrates simulated and fitted log p-values for the case  $\delta = 2$ . In particular, the roles of the parameters  $\gamma_0(\delta)$ ,  $\gamma(\delta)$  and  $x'(\delta)$  can be seen. That beyond  $x'(\delta)$  the log p-value decreases close to linearly with gradient -1 can also be seen, as can the resemblance between the log p-value over the range  $0 < x \leq x'(\delta)$  and a segment of a circle. Denoting the radius of the circle  $r(\delta)$  and its origin  $(x_0(\delta), y_0(\delta))$ , the relationship between these three parameters and  $\gamma_0(\delta)$ ,  $\gamma(\delta)$  and  $x'(\delta)$  is described in Figure 5 and Equation (2.5).

Smooth functional forms were fitted to describe  $\gamma_0(\delta)$ ,  $\gamma(\delta)$  and  $x'(\delta)$  over  $\delta$ . The functions are assumed to be of the form

$$\begin{aligned}
 \gamma_0(\delta) &= e^{-a_1(\delta+b_1)} + c_1\delta + d_1 \\
 \gamma(\delta) &= e^{-a_2(\delta+b_2)} + c_2\delta + d_2 \\
 x'(\delta) &= a_3\delta^2 + b_3\delta + c_3.
 \end{aligned}
 \tag{3.5}$$

The first two functions are of ‘non-linear decay + linear’ form and the third of quadratic form. Approximate values for the parameters  $a_i, b_i, c_i, i \in (1, 2, 3)$  were obtained through fitting the formula given in (2.6) to simulated p-values, on the log scale, with the constraint that  $1 - \gamma_0(\delta)$  should approximately equal the simulated lump on zero for a particular value of  $\delta$ . The p-values and lump on zero were simulated by



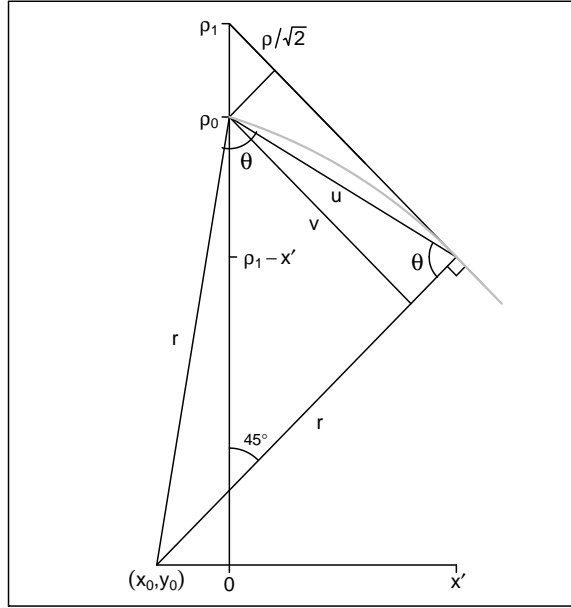


Figure 5: Circle segment (in grey) assumed to describe log p-values,  $\log \mathcal{P}_\delta(x)$ , corresponding to CUSUM values  $x$  in the range for  $0 < x \leq x'(\delta)$ . All parameters depend on  $\delta$ .

generating between  $2 \times 10^7$  and  $1.4 \times 10^9$  null CUSUM values under each of a range of values for  $\delta$ , equal to 0.5 to 4 by steps of 0.25. For  $\delta$  in this range, the simulated p-value at each value of a vector of 20–30 CUSUM values was considered, so that the formula in (2.6) was fitted to a grid of 340 p-values using the R `optim()` function (R project, version 2.4.0). The error sum of squares for small  $x$ , large  $x$  and of  $\log(\gamma_0)$  from the logarithm of one minus the simulated lump on zero were minimised conjointly.

The resulting optimised values for  $a_i, b_i, c_i$ ,  $i \in (1, 2, 3)$  are given in Equation (2.4). Figure 6 shows the simulated lump on zero and the fitted lump,  $1 - \gamma_0(\delta)$ , from Equation (2.4). Extra simulated values relating to more extreme values of  $\delta$  are added to show that the fitted function extrapolates well.

### 3.3 Assessing the Quality of the Approximation

We consider two sources of error: the error due to simulation, and the error due to the empirical approximation itself. The error measures calculated were averaged over a vector of  $M_\delta$  CUSUM points for the part of the approximation relating to large CUSUM values only and  $N_\delta$  for the complete approximation (2.3) valid over the entire range of CUSUM values. The range of  $x$ 's covered in each case is reported in the second and third columns of Table 1 and  $N_\delta$  and  $M_\delta$  in columns 4 and 5 respectively. The number of simulations upon which the error measures are based for each  $\delta$  value,  $l_\delta$ , is given in column 6.

The error due to simulation can be summarised by the monte carlo error associated with each simulated p-value  $\tilde{\mathcal{P}}_\delta(x_j)$ , calculated as

$$me(\delta, j) = \sqrt{\frac{\sigma_{x_j}^2 \{1 + \rho_{1j}(\delta)\}}{l_\delta \{1 - \rho_{1j}(\delta)\}}}, \quad (3.6)$$

where  $\sigma_{x_j}^2(\delta) = \tilde{\mathcal{P}}_\delta(x_j)[1 - \tilde{\mathcal{P}}_\delta(x_j)]$  is the estimated sample variance of the indicator function  $I(x > x_j)$  calculated for a particular value  $x_j$  of the CUSUM,  $\rho_{1j}(\delta)$  is the sample autocorrelation of  $I(x > x_j)$  at lag

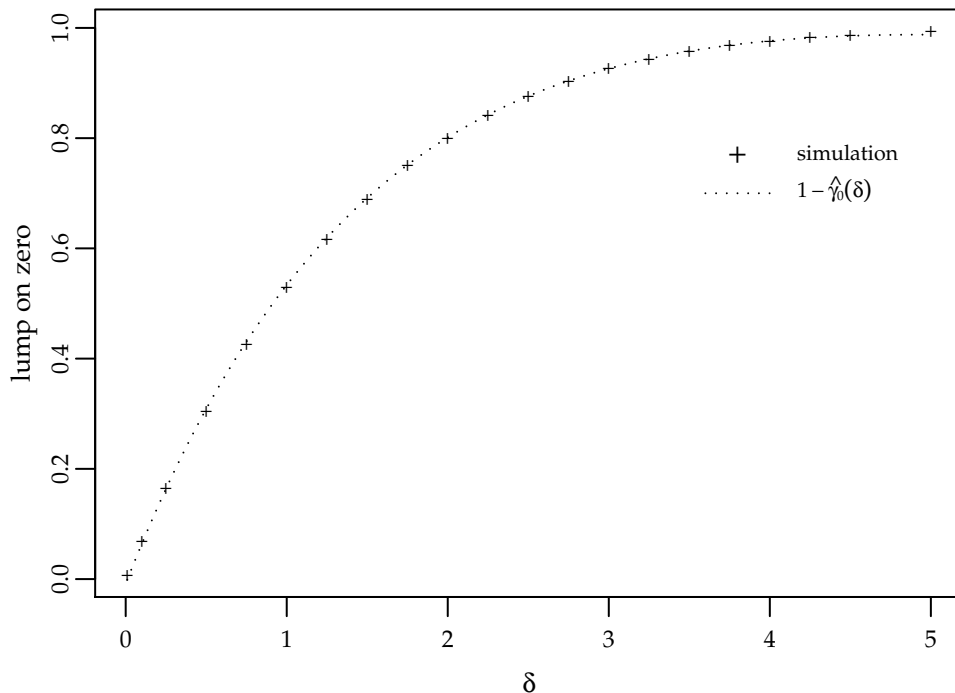


Figure 6: Simulated steady-state lump on zero, or proportion of time CUSUM spends at zero under long-run null conditions.

1 and  $l_\delta$  is the number of simulated values at each  $\delta$ . The error due to the empirical approximation can be measured on the log p-value scale as the squared error of the fitted log p-values from the simulated log p-values for  $x_j$ :

$$se(\delta, j) = \left\{ \log \widehat{\mathcal{P}}_\delta(x_j) - \log \widetilde{\mathcal{P}}_\delta(x_j) \right\}^2. \quad (3.7)$$

For the part of the approximation relating to large CUSUM values, the monte carlo error averaged over the vector of  $M_\delta$   $x_j$ 's,  $\sum_{j=1}^{M_\delta} se(\delta, j)/M_\delta$ , did not exceed 0.0002 and the mean squared error,  $\sum_{j=1}^{M_\delta} se(\delta, j)/M_\delta$ , did not exceed 0.004. For the complete approximation, the mean monte carlo error calculated in a similar way was no greater than 0.0002 and the mean squared error, 0.0025.

Although it is reasonable to minimise the error of the fitted log p-value, since the log p-value is roughly linear over  $x$ , it may be more appropriate to consider the size of the error in terms of actual p-values. Values of mean percentage relative error from simulated p-values are given in columns 7 and 8 of Table 1, where mean percentage relative error is measured as

$$re(\delta) = \sum_{j=1}^{M_\delta} \frac{100}{M_\delta} \left| \exp \sqrt{se(\delta, j)} - 1 \right| \% = \sum_{j=1}^{M_\delta} \frac{100}{M_\delta} \left| \frac{\widehat{\mathcal{P}}_\delta(x_j)}{\widetilde{\mathcal{P}}_\delta(x_j)} - 1 \right| \% \quad (3.8)$$

for the part of the approximation relating to large CUSUM values only, and similarly for the complete approximation. The error defined by (3.8) is seen to not exceed 5%, and the mean percentage relative error relating to the complete approximation, 4%.

The monte carlo error from simulations can also be considered in terms of percentage relative error, but from the true p-value  $\mathcal{P}_\delta(x)$ . The true p-value can be approximated using the derived formula (2.6), and the

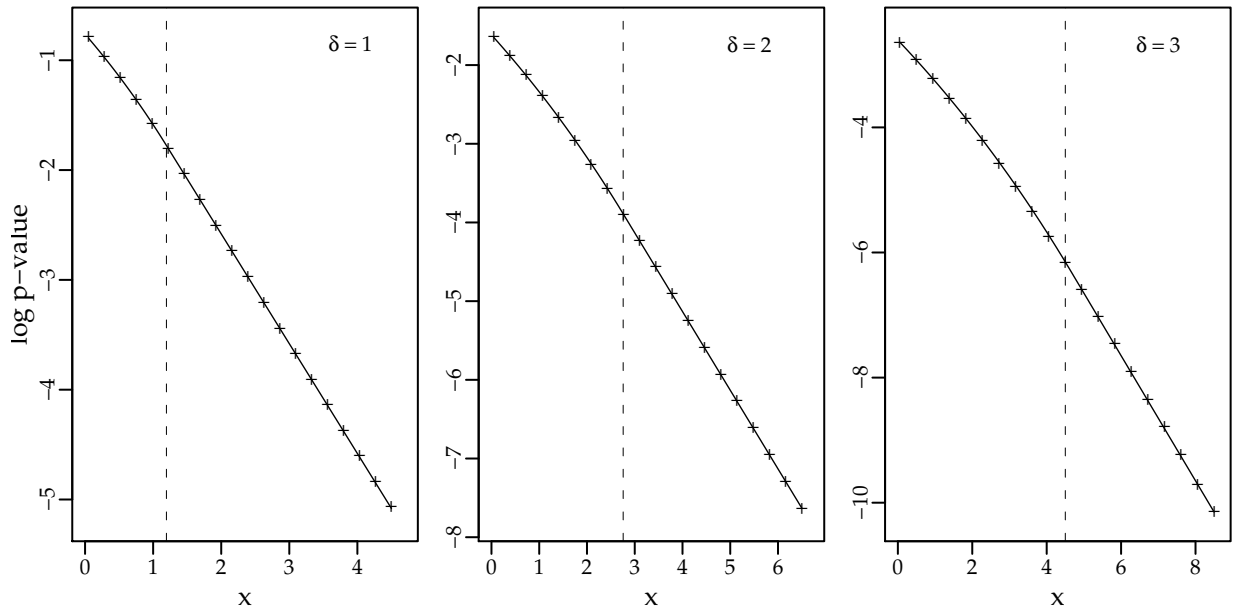


Figure 7: Quality of fit of approximate steady-state log p-value  $\log \widehat{\mathcal{P}}_\delta(x)$  (solid line) to simulated p-values  $\log \widetilde{\mathcal{P}}_\delta(x)$  ('+') for shift size  $\delta = 1, 2$  and  $3$ . The dotted lines indicate the approximated value of  $x'(\delta)$ ,  $\widehat{x}'(\delta)$ .

mean percentage relative error due to simulation calculated as

$$rme(\delta) = \sum_{j=1}^{M_\delta} \frac{100}{M_\delta} \left[ \frac{me(\delta, j)}{\widetilde{\mathcal{P}}_\delta(x_j)} \right] \% \quad (3.9)$$

for the part of the approximation relating to large CUSUM values only, and similarly for the complete approximation. From Table 1, the error defined by (3.9) increases with  $\delta$  up to a value of roughly 5.5%. Those relating to the entire range of CUSUM increase with  $\delta$  up to approximately 2.5%. As we would expect, the relative error due to the approximation is in most instances larger than that purely attributable to sampling error. In places where it is not, there may be slight overfitting to simulated values.

For the complete approximation (2.3), the mean percentage relative error  $re(\delta)$  can be seen to be reasonably small, especially for smaller  $\delta$ . The increase in  $re(\delta)$  with  $\delta$  is in tune with the increase in sampling error with the same. Figure 7 shows the quality of the fit for  $\delta = 1, 2$  and  $3$  respectively, comparing approximated p-values  $\widehat{\mathcal{P}}_\delta(x)$  to simulated values  $\widetilde{\mathcal{P}}_\delta(x)$ . Figure 8 gives a graphic comparison of  $re(\delta)$  for the approximation relating to large  $x$  and the complete approximation, from which it can be seen that the error is roughly balanced over the two.

## 4 Example: Earthquakes

We return to the earthquake data illustrated in Figure 1. Recall that this CUSUM is sequentially testing the null hypothesis that the square-rooted data from 1940 onwards are  $N(4.40, 0.74)$  versus the alternative that they are  $N(6.61, 0.74)$ , thus it is testing for a 3 standard deviation shift upwards in the mean, so that  $\delta = 3$ . Suppose we are interested in the marginal significance of the peak in the CUSUM around 1950 – if

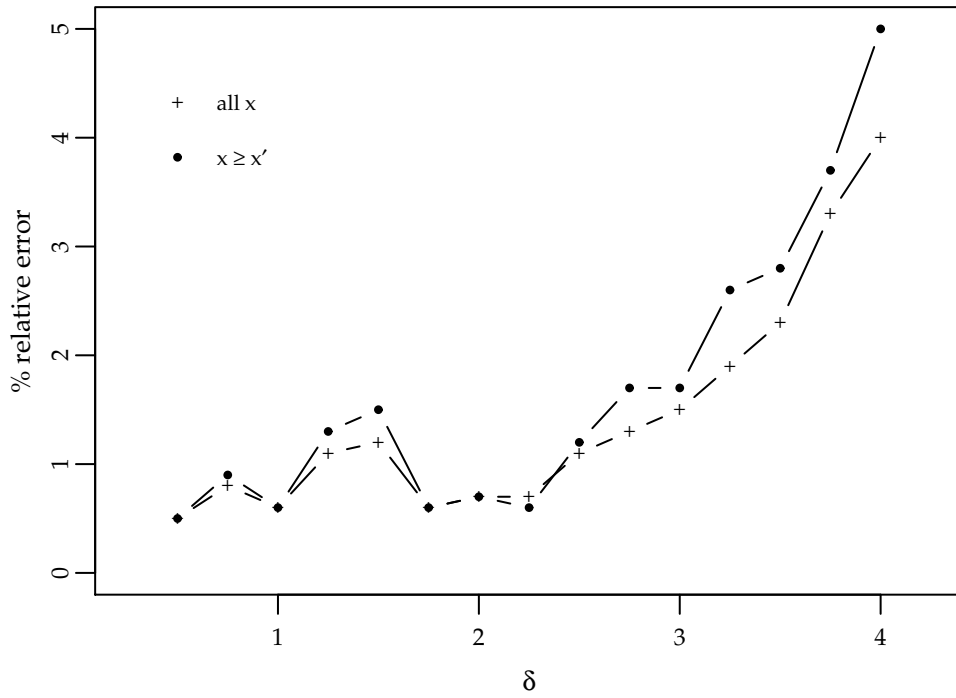


Figure 8: Mean percentage relative error  $re(\delta)$  of fitted steady-state p-values (calculated from (2.6)) from simulated p-values. The error value associated with each  $\delta$  in  $[0.5, 4]$  is plotted. The mean percentage relative error relating to the part of (2.6) valid for large CUSUM values only is given for comparison.

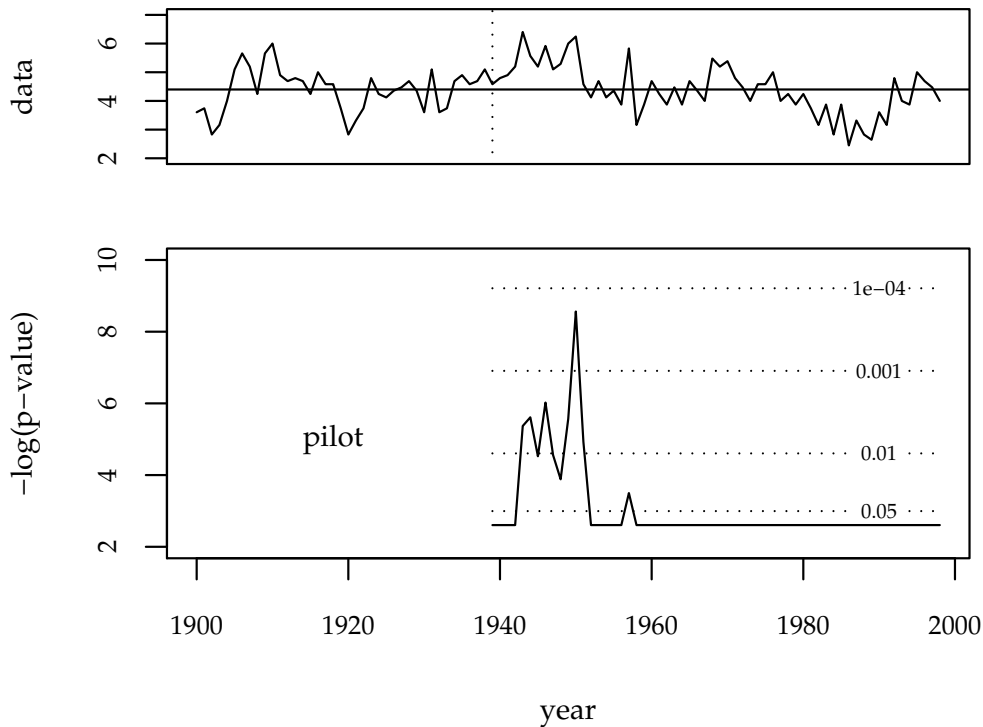


Figure 9: Lower panel: minus log steady-state p-value, calculated from (2.6), of a CUSUM testing for a 3 sd shift in mean square-rooted number of earthquakes per year worldwide (with magnitude  $> 7$  on the Richter scale) where the null is taken to be  $N(4.40, 0.736)$ . P-value thresholds have been added to the plot. Upper panel: the transformed data, including the pilot data.

we ran such a CUSUM on the standardised data  $(y' - 4.40)/\sqrt{0.74}$  under null conditions, how often would we expect to see values this large? The steady-state p-values associated with the CUSUM values drawn in Figure 1 can be approximated using (2.6) with  $\gamma_0(3) = 0.074$ ,  $\gamma(3) = 0.192$  and  $x'(3) = 4.67$ , all from (2.4). The lump of steady-state probability on zero is approximated as 0.926 and so approximate p-values  $\hat{P}_3(x)$  associated with non-zero CUSUM values cannot be greater than 0.074.

Figure 9 is a plot of minus the log steady-state p-value against year for the earthquake data. From the figure, it can be seen that plotting the p-values on this scale results in their roughly following the shape of the CUSUM (compare with Figure 1). This is because the relationship between the p-value on this scale and the CUSUM value  $x_t$  is exactly linear for large  $x_t$  (see (2.6)). Consulting Figure 9, in 1950 the steady-state p-value falls below 0.001 and  $x_t$  above 5.26. Under the null hypothesis that the square-root of the number of earthquakes of this magnitude occurring per year is Normal with mean 4.40 and variance 0.736, we would only expect the CUSUM value to be this extreme once every 1000 years. The null ARL to a CUSUM value of at least 5.26 (using Equation (2.9) and the spc package in R (R project, version 2.4.0)) is approximately 1150, so with resetting at  $x_t > 5.26$  we would only expect to see a CUSUM value this extreme once every 1150 years.

## 5 Signalling Procedure for Multiple CUSUM Schemes

In this section we propose a signalling procedure for a multiple CUSUM scheme. The procedure utilises the false discovery rate (FDR) approach of Benjamini & Hochberg (1995), which is a powerful approach to identifying outlying observations in the presence of multiplicity.

In a system with a large number of sub-processes or units, where each is to be monitored by a scheme such as the CUSUM, it is practical to consider that a small percentage of those (which need not always be the same ones) are likely to be extreme in comparison to the main body and hence we can expect a certain number of signals at each inspection point. Where true signals from the group of charts may be mixed in with false signals, controlling the rate of false signals is less useful than controlling the rate of false signals amongst all signals. This is primarily because inference is simpler and more direct if the latter is controlled: if we have a number of signals, we can immediately infer a soft upper bound (because the control is in expectation only) for the number that are likely to be false without needing to consider the expected proportions of false and true signals.

The idea of using an FDR approach in the context of multiple CUSUM charting procedures has been propounded by Benjamini & Kling (1999); Marshall *et al.* (2004) and Grigg & Spiegelhalter (2005). Benjamini & Kling (1999) suggest applying the Benjamini & Hochberg (1995) FDR approach across charts and over time to p-values obtained as inverted ARL-values corresponding to each CUSUM value. Contrastingly, Marshall *et al.* (2004) advocate implementing Storey's (2002) positive FDR procedure to estimate, given signals across charts with fixed boundaries and in a fixed window of time, how many of those signals are false. The fixed window aspect of this method seems restrictive in a continuous monitoring setting, as it turns what is a sequential test in origin into a fixed sample size test.

Benjamini & Hochberg (1995) give a definition of their Bonferroni-type multiple-testing procedure as follows: suppose  $P_{(1)} \leq P_{(2)} \leq \dots \leq P_{(I)}$  are the ordered p-values to be analysed, where the null hypothesis corresponding to the  $i^{\text{th}}$  ordered p-value is denoted  $H_{(i)}$ . Defining  $k$  to be the largest  $i$  for which

$$P_{(i)} \leq \frac{i}{I}q^* \quad (5.1)$$

the procedure rejects all  $H_{(i)}$  for  $i = 1, 2, \dots, k$ . Benjamini & Hochberg (1995) prove that, for independent test statistics and for any configuration of false null hypotheses, under the null their procedure controls the FDR at the specified level  $q^*$ . They define FDR to be the expected rate of false discoveries amongst all discoveries.

The signalling procedure described by equation (5.1) is conservative: if we are prepared to estimate, prior to monitoring, the proportion  $\pi_0$  of 'in control' units, we can increase the bound to  $iq^*/(I\pi_0)$  and retain FDR control at the level  $q^*$  (Storey *et al.*, 2004; Storey, 2002).

For our multiple CUSUM scheme, we propose use of rule (5.1) at every time point  $t$  across the  $I$  charts drawn. Only the current p-values  $p_{it}$  are sorted and examined at each time point  $t$ . Since we have multiple charts, 'false discovery rate' is here synonymous with 'proportion of false signals amongst all signals' across  $I$  charts at time  $t$ , where a chart is taken to signal if its (ordered) p-value satisfies (5.1). The parameter  $q^*$  defines a signalling threshold that does apply uniformly to all charts in that the same rule, (5.1), is used. However, at any one time, the ordering of the p-values across charts (a facet of the data) determines the individual chartwise critical thresholds, and so these can vary over time and charts.

Since the FDR rule (5.1) controls in expectation the FDR at time  $t$ , it follows by the linearity of expectation that the rate over any window of time will be controlled in expectation at the same level. Formally, if  $q_t$  is the proportion of the observed signals at time  $t$  which are false, or the current false discovery rate, then

$$\mathbb{E}\left(\frac{\sum_{t=1}^T q_t}{T}\right) = \frac{\sum_{t=1}^T \mathbb{E}(q_t)}{T} \leq q^* \quad (5.2)$$

if  $\mathbb{E}(q_t) \leq q^*$  for all  $t \in 1, 2, \dots, T$ .

A simulation study was carried out in order to investigate the properties of the proposed signalling procedure. Data generated from  $I = 500$  units were charted over a monitoring period of length  $T = 100$ , with 90%

of units assumed to be ‘in-control’ such that each observation  $y \sim N(0, 1)$  and 10% assumed to be ‘out-of-control’ so that  $y \sim N(2, 1)$ . Charts were taken to signal according to the FDR controlling procedure described above. Following a signal charts were restarted at the subsequent time point. This is not taken into account in the above simple proof (5.2), but the results in Table 2 show that the FDR is still controlled in expectation. The monitoring scenario was simulated 500 times at each chosen value of the controlling level parameter  $q^*$ . For each simulation run, the FDR over a window  $T$ ,  $\sum_1^T q_t/T$ , was recorded as well as the mean rate of signals,  $\sum_1^J a_j/J$ , across charts corresponding to ‘in-control’ units (indexed by  $j$  and totalling  $J = 450$ ). The statistic  $\sum_1^J a_j/J$  is an estimator of the proportion  $\alpha$  of null charts signalling at an arbitrary time point  $t$  (the incidental, or marginal type I error rate). Because the charts are reset just after signalling,  $1/\alpha$  is the average time to signal under this stopping rule for a single ‘in control’ unit and from a chart position of zero.

Also measured were the successful detection rate up to time  $t$  ( $\text{SDR}_t$ ) and the false signalling rate up to time  $t$  ( $\text{FSR}_t$ ). These were based on 1000 simulation runs. The SDR is effectively a measure of power of the signalling procedure, and is defined to be the cumulative rate of true signals across charts corresponding to ‘out-of-control’ units.  $\text{FSR}_t$  is the cumulative rate of false signals across charts corresponding to ‘in-control’ units and tends to  $1 - (1 - \alpha)^t$  as time elapses. Over short runs of length  $t$  from the zero state,  $\text{FSR}_t$  is less than  $1 - (1 - \alpha)^t$ . This follows since small CUSUM values are positively correlated on null charts (due to the effect of the reflecting boundary at zero and the negative drift under the null), making quick signalling from nearby the zero state less likely than under a geometrically distributed waiting time distribution where the sequence of logicals regarding the signalling status would be independent.

Table 2 displays the summary statistics collected for each value of the design parameter  $q^*$ . It can be seen that the FDR is well controlled in expectation at the specified level, such that the expected value  $\pi_T = \mathbb{E}(\sum_{t=1}^T q_t/T)$  is estimated to be closer to half  $q^*$  than  $q^*$  in all cases. However, given that the control is in expectation only, at single instances of time the proportion of discoveries that are false could be larger than  $q^*$ .

The SDR is seen to increase with  $q^*$  and rapidly with  $t$ . For  $t = 4$ , the SDR is seen to not approach 90% until around  $q^* = 0.15$ ; whereas, for  $t = 6$ , the SDR is approaching 90% around  $q^* = 0.02$ . The FSR up to  $t = 8$  is seen to be small, and not exceeding 10% for moderate values of  $q^*$ . For  $q^* = 0.2$  it can be observed that, on average, 49% of null charts signalled at least once over the 100 time points. It can be checked that  $\text{FSR}_t$  is less than  $1 - (1 - \hat{\alpha})^t$  for the smaller values of  $t$ , where  $\hat{\alpha} = \sum_1^J a_j/J$ , and that  $\text{FSR}_{100} \approx 1 - (1 - \hat{\alpha})^{100}$ . As might be expected, the distribution of number of signals within the  $T = 100$  periods on charts relating to ‘in control’ units is close to  $\text{Poisson}(\hat{\alpha}T)$ .

## 6 Conclusion

We have presented here a new and useful characterisation of the standard Normal CUSUM. An approximation to the steady-state distribution of the CUSUM statistic  $f_\delta(x)$  is provided in Equation (2.3), and this approximation leads directly to a plug-in formula (2.6) for a steady-state p-value associated with a CUSUM value  $x$ . Use of the formula is demonstrated via an example where the data are earthquakes worldwide registering  $> 7$  on the Richter scale over the period 1900 to 1998. The overall quality of the approximation in terms of percentage relative error is seen to be good (Table 1), especially for small  $\delta$ . Also, from Figure 8, the fit is balanced in terms of relative error measured at smaller and larger CUSUM values.

The steady-state p-value measure presented appears to be relatable to the popular null ARL measure via Equation (2.9) and hence also the inverted ARL, another possible p-value measure. That measure was proposed by Benjamini & Kling (1999) and yields more extreme (smaller) p-values than the measure described here. Use of an empirical formula for the ARL under the null, such as Siegmund’s 1985 formula or Equation (2.9) combined with (2.6) might facilitate calculation of Benjamini & Kling’s p-value measure. More generally, Equation (2.9) could be viewed as an alternative to Siegmund’s approximation and other numerical methods used to calculate the ARL.

The theory relating to large CUSUM values  $x$  given in Section 3.1 is applicable when the data  $y$  being

monitored are non-Normal (details are given in the Appendix). This allows, for non-Normal data, steady-state p-values to be calculated correctly up to a constant of proportionality. Approximating the absolute steady-state p-values for such data could perhaps be approached in a similar way to here, but empirical forms would in general need to vary over parameters associated with the data (such as the rate parameter for Poisson counts) as well as over the CUSUM shift parameter  $\delta$ .

Monitoring multiple series of outcomes is becoming of increasing importance in the healthcare setting. Objective prospective monitoring necessarily involves the use of some kind of control chart. Where there are many units to be monitored, for example hospital trusts, multiple charts would inevitably be needed. Implementing a signalling procedure like that described here where the FDR is controlled across charts and time seems a sensible approach to controlling the multiplicity of such a charting scheme.

## Acknowledgements

We are grateful to Vern Farewell for many helpful discussions.

## Appendix

Suppose we would like to test for a shift in the natural parameter  $\theta$  of an exponential family distribution  $\psi(y|\theta)$ , so that

$$\begin{aligned} H_0 : \psi_0(y) &= e^{A(y)+B(\theta)+y\theta} \\ H_1 : \psi_1(y) &= e^{A(y)+B(\theta+\delta)+y(\theta+\delta)}, \quad |\delta| > 0, \end{aligned}$$

then the log-likelihood ratio statistic  $w$  takes the form

$$w = \delta y + B(\theta + \delta) - B(\theta). \quad (6.1)$$

From (3.2), we have that the general solution to the Wiener-Hopf integral equation (3.1) is proportional to an  $\text{Exp}(\beta)$  distribution when the CUSUM value  $x$  is large, and that the moment-generating function of the weights  $w$  equates to 1 (see Equation (3.3)).

Strictly, if  $y$  is discrete, the integral in the Wiener-Hopf equation becomes a summand (over the domain of  $w$ , with  $w < x$ ), as does the integral in equation (3.2) (over the full domain of  $w$ ). However, this does not prevent a solution (for large  $x$ ) of the form (3.2). If  $x$  can only take a countable number of values, then the steady-state distribution, provided it exists, will be discrete. Specifically, it will be proportional in the tail to a geometric distribution, and hence be expressible as (up to a constant)

$$\Pr_\delta[X = x | \zeta] = (1 - \zeta)\zeta^x, \quad x \in \mathcal{X}, \quad x > x'(\delta)$$

where  $\mathcal{X}$  is the set of values that  $x$  can take under steady-state conditions, and  $\zeta = e^{-\beta}$  leads to equivalence with (3.2).

Now if  $\mathbb{E}[e^{\beta w}] = 1$  then from (6.1)

$$\begin{aligned} \mathbb{E}[e^{\beta\{\delta y + B(\theta+\delta) - B(\theta)\}}] &= 1, \quad \text{which implies that} \\ e^{\beta[B(\theta+\delta) - B(\theta)]} \mathbb{E}[e^{\beta\delta y}] &= 1. \end{aligned} \quad (6.2)$$

The term  $\mathbb{E}[e^{\beta\delta y}]$  is equivalent to the moment-generating function of  $y$  and can be shown to be equal to  $e^{B(\theta) - B(\theta + \beta\delta)}$ . Using (6.2), we have that

$$\beta B(\theta + \delta) + (1 - \beta)B(\theta) - B(\theta + \beta\delta) = 0.$$

Now  $B(\theta + \beta\delta)$  is concave (since  $B(\theta)$  is concave) so the line  $\beta B(\theta + \delta) + (1 - \beta)B(\theta)$  must cross it twice, at  $\beta = 0$  and  $\beta = 1$ . Thus, if the steady-state distribution of  $x$ ,  $f_\delta(x)$ , exists it is  $\text{Exp}(1)$  in the tail.



## References

- Benjamini, Y., & Hochberg, Y. 1995. Controlling the False Discovery Rate: a practical and powerful approach to multiple testing. *J. R. Statist. Soc. B*, **57**(1), 289–300.
- Benjamini, Y., & Kling, Y. 1999 (August). *A look at statistical process control through the p-values*. Tech. rept. RP-SOR-99-08. Tel Aviv University, Israel. <http://www.math.tau.ac.il/~kling/SPCpvalue.html> Accessed on 10/04/05.
- Brook, D., & Evans, D.A. 1972. An approach to the probability distribution of cusum run length. *Biometrika*, **59**(3), 539–549.
- Cox, D.R., & Miller, H.D. 1965. *The Theory of Stochastic Processes*. New York: John Wiley.
- Feller, W. 1966. *An Introduction to Probability Theory and its Applications*. Vol. II. New York: Wiley.
- Goel, A.L., & Wu, S.M. 1971. Determination of ARL and a contour nomogram for CUSUM charts to control Normal mean. *Technometrics*, **13**, 221–230.
- Grigg, O.A., & Spiegelhalter, D.J. 2005. Random-effects CUSUMs to monitor hospital mortality. *Pages 239–246 of: Francis, A.R., Matawie, K.M., Oshlack, A., & Smyth, G.K. (eds), Statistical Solutions to Modern Problems: Proceedings of the 20th International Workshop on Statistical Modelling*. Sydney: University of Western Sydney Press.
- Hawkins, D.M., & Olwell, D.H. 1997. *Cumulative Sum Charts and Charting for Quality Improvement*. New York: Springer.
- Marshall, C., Best, N., Bottle, A., & Aylin, P. 2004. Statistical issues in the prospective monitoring of health outcomes across multiple units. *J. R. Statist. Soc. A*, **167**(3), 541–559.
- Page, E.S. 1954. Continuous inspection schemes. *Biometrika*, **41**, 100–115.
- Siegmund, D. 1985. *Sequential Analysis Tests and Confidence Intervals*. New York: Springer-Verlag.
- Sornette, D., & Cont, R. 1997. Convergent multiplicative processes repelled from zero: power laws and truncated power laws. *Journal de Physique I*, **7**, 431–444.
- Storey, J. D. 2002. A direct approach to false discovery rates. *J. R. Statist. Soc. B*, **64**, 479–498.
- Storey, J. D., Taylor, J. E., & Siegmund, D. 2004. Strong control, conservative point estimation and simultaneous conservative consistency of false discovery rates: a unified approach. *J. R. Statist. Soc. B*, **66**, 187–205.
- Woodall, W.H., & Mahmoud, M.A. 2005. The inertial properties of quality control charts. *Technometrics*, **47**, 425–436.

Table 1: Mean percentage relative error  $re(\delta)$  of approximate steady-state p-values for CUSUM values  $x$  compared to the mean percentage pure monte carlo error  $rme(\delta)$ . The range and number ( $N_\delta$  or  $M_\delta$ ) of  $x$  values upon which the errors are based are given as well as the number of simulated values  $l_\delta$ .

$\delta$	range of $x$		no. of $x$ values		no. of sims $\times 10^{-6}$ $l_\delta \times 10^{-6}$	mean % relative error (monte carlo)		
	$x \geq 0$	$x > x'(\delta)$	$N_\delta$	$M_\delta$		$x \geq 0$	$x > x'(\delta)$	
0.50	0.01—3.5	0.55—3.5	20	17	20	0.5 (0.2)	0.5 (0.2)	
0.75	0.05—4.0	0.86—4.0	20	16	20	0.8 (0.2)	0.9 (0.2)	
1.00	4.5	1.20—4.5		15		0.6 (0.2)	0.6 (0.3)	
1.25	5.0	1.56—5.0		14		1.1 (0.3)	1.3 (0.3)	
1.50	5.5	1.94—5.5		13		1.2 (0.3)	1.5 (0.4)	
1.75	6.0	2.34—6.0		13		0.6 (0.3)	0.6 (0.5)	
2.00	6.5	2.76—6.5		12		30	0.7 (0.4)	0.7 (0.5)
2.25	7.0	3.21—7.0		11		30	0.7 (0.4)	0.6 (0.7)
2.50	7.5	3.67—7.5		11		30	1.1 (0.5)	1.2 (0.9)
2.75	8.0	4.16—8.0		10		40	1.3 (0.6)	1.7 (1.0)
3.00	8.5	4.67—8.5		9		40	1.5 (0.7)	1.7 (1.4)
3.25	9.0	5.19—9.0	30	13	50	1.9 (0.8)	2.6 (1.5)	
3.50	9.5	5.74—9.5	13	13	70	2.3 (0.8)	2.8 (1.6)	
3.75	10.0	6.32—10.0	12	12	120	3.3 (0.9)	3.7 (2.0)	
4.00	0.05—12.0	6.91—12.0	13	13	140	4.0 (2.4)	5.0 (5.3)	

Table 2: Properties of FDR signalling procedure, given  $J = 450$  ‘in control’ units where  $y \sim N(0, 1)$ , 50 ‘out of control’ units where  $y \sim N(2, 1)$ , and duration of monitoring  $T = 100$ . The mean proportion of false signals amongst signals,  $\sum_1^T q_t/T$ , the mean rate of signals across ‘in-control’ charts,  $\sum_1^J a_j/J$ , and associated standard errors are based on 500 simulations. The successful detection rate amongst ‘out of control’ charts and false signalling rate amongst ‘in control’ charts up to and including time  $t$ ,  $SDR_t$  and  $FSR_t$ , are based on 1000 simulations.

$q^*$	$\sum_1^T q_t/T$ (se $\times 10^3$ )	$\sum_1^J a_j/J$ (se $\times 10^4$ )	SDR <sub><i>t</i></sub> (%) FSR <sub><i>t</i></sub>			
			$t = 4$	$t = 6$	$t = 8$	$t = 100$
0.01	0.006 (2.2)	0.00017 (0.62)	61.3 0.03	87.9 0.07	96.4 0.11	1.69
0.05	0.028 (4.1)	0.0011 (1.5)	79.4 0.28	94.9 0.51	98.8 0.74	10.5
0.075	0.042 (5.1)	0.0018 (2.1)	83.9 0.50	96.2 0.85	99.2 1.20	16.9
0.1	0.057 (5.7)	0.0026 (2.7)	86.7 0.78	97.0 1.30	99.3 1.82	23.3
0.15	0.086 (7.1)	0.0045 (3.8)	90.2 1.44	98.0 2.29	99.6 3.18	36.7
0.2	0.12 (7.8)	0.0067 (4.8)	92.3 2.18	98.5 3.44	99.8 4.73	49.2
0.25	0.15 (8.1)	0.0092 (5.9)	93.9 3.15	98.9 4.97	99.8 6.74	61.0
0.3	0.18 (7.8)	0.012 (6.6)	94.9 4.16	99.1 6.48	99.9 8.70	71.1
0.35	0.21 (8.3)	0.016 (8.0)	95.9 5.36	99.4 8.35	99.9 11.2	79.6
0.4	0.24 (8.3)	0.020 (9.5)	96.5 6.84	99.6 10.6	99.9 14.1	86.6
0.45	0.27 (8.9)	0.025 (12)	97.1 8.62	99.6 13.1	99.9 17.4	92.0