J. Opl Res. Soc. Vol. 35, No. 9, pp. 811–820, 1984
Printed in Great Britain. All rights reserved

0160-5682/84 $3.00 + 0.00

"Trial by Computer"—A Case Study of the Use of Simple Statistical Techniques in the Detection of a Fraud

DAVID G. JOHNSON
Department of Management Studies, Loughborough University of Technology

This case study describes an investigation which was undertaken for a local shopkeeper after one of his shop assistants had been convicted of theft. The person concerned had admitted stealing a few pounds on just one occasion, but a series of worsening year-end results over a period of three or four years caused the shopkeeper to suspect that stealing had been taking place on a very large scale over a long period of time. Could this be 'proved' by examining in detail the records of the shop's business for the years in question?

The analysis undertaken involved principally the number of customers each week and the revenue per customer. The most usual form of theft involves 'understriking' (ringing into the till less than the value of the sale and subsequently taking out the difference). This practice will be expected to reduce the revenue per customer and increase the number of customers (because the incorrect entry is usually disguised by immediately ringing in a no-sale or a zero amount). Clear evidence was found of stealing on a much larger scale than had been admitted.

Key words: cusum, change-point analysis, intervention analysis, autoregressive-moving average processes

INTRODUCTION

THIS CASE STUDY describes an investigation which was undertaken for a local shopkeeper after one of his shop assistants had been convicted of theft. The story begins in March 1979, when a disastrous set of year-end results caused the auditors to suspect that someone connected with the business was systematically stealing either cash or stock from the premises, and as a consequence, the owner was on the look-out for any indications of dishonesty by his staff. Some time afterwards, the owner happened to notice one of his shop assistants, whom we shall call Mrs F, ringing into the till amounts of 1p and 2p, and as the shop did not sell any items costing as little as this, the owner was naturally suspicious. As a result, the police were informed and, by means of spot purchases made by plain-clothed officers, they ultimately caught Mrs F understriking on two occasions. (Understriking refers to the procedure of ringing into the till an amount less than the actual value of the purchase and stealing from the till the difference between these two amounts.) At the ensuing court case, Mrs F admitted stealing £5 and £8 on the two occasions which had been observed and in total a sum of £250 during the period February 1979 to June 1979.

The owner, however, was highly dissatisfied with the outcome of the court case. He was convinced that Mrs F (or perhaps even someone else) was stealing on a much larger scale for the effect on business profits to have been so noticeable. His own detailed analysis of the previous year's performance had led him to suspect that the sum stolen amounted not to £250 but to many thousands of pounds.

THE PROCEDURE EMPLOYED

Understriking is always difficult to detect if it is carefully done, but it was thought that there may be four different ways in which it could be highlighted.

(i) If the amount stolen is sufficiently large, it should cause a significant fall in the average revenue per customer. The practice of understriking is often disguised by the thief ringing in a zero amount or a 'no sale' immediately after the incorrect amount has been rung in. In this way the customer will probably not have time to notice the incorrect amount. If this procedure is used, the average revenue per customer will be doubly depressed for not only is the total revenue being reduced, but also the total number of 'customers' is being artificially increased.

(ii) For the reason just described, one might expect the number of 'customers' to be inflated by understriking.

(iii) Again, if a no-sale entry is used to disguise understriking, then the number of 'no sales' or 'zero' totals should be inflated.

(iv) If the thief is stealing on a large scale, he or she will often not remove money from the till at the time of the purchase, as this may well be noticed by the customer. The thief, therefore, has to remember how much has been understruck on each occasion and the total amount removed later when no-one is watching. In this situation, it is easy for the thief to lose track of the amount stolen and they will often err on the side of caution and remove less from the till than the amount by which the till was actually understruck. This causes there to be more money in the till than is actually recorded on the till roll, creating a positive till 'difference'. With normal working, one might expect the differences due to routine error to average zero, and so a consistently positive difference may provide further evidence of understriking.

There are, of course, a number of factors other than stealing which might lead to the four indicators listed above, such as, for example, the number of customers increasing owing to seasonal changes in buying patterns. It is therefore important that none of these factors be taken in isolation, and only if they all point to the same conclusion will it be seen as evidence for the occurrence of stealing by means of persistent understriking.

THE DATA

The owner of the shop had for many years kept detailed records of daily takings in the form of

(i) the number of customers,
(ii) the till readings, i.e. the amount rung into each till since it was last cashed up,
(iii) the difference between the till reading and the amount of money actually in each till.

To reduce the data to more manageable proportions, each of the above three pieces of information was aggregated to give a weekly reading from November 1972 onwards (this date marking a significant change in the nature of the shop premises). Additionally, the days when Mrs F was absent for any reason were extracted from shop records and noted down.

In analyzing the data with a view to detecting the occurrence of stealing, it is clearly necessary to take account of the time of the year. For example, most shops would be significantly busier at Christmas time than any other time of year, with the number of customers and the average revenue per customer increasing many fold in the weeks prior to Christmas. If no account were taken of these seasonal influences, it would be difficult to detect which changes in business activity might be due to theft, and so some deseasonalizaton of the data seems essential. Accordingly, before the data were further analyzed, the average revenue per customer was seasonally adjusted, having estimated from the data a seasonal factor for each week of the year. As may be anticipated, the only really significant factors were for the few weeks around Christmas.

Unfortunately, seasonality is not the only factor which must be taken into account, for the average revenue per customer will also be affected by inflation. Not only do the data have to be deseasonalized, they also need to be deflated and expressed in 'constant price' terms relative to some base year. The most readily attainable measure of inflation is provided by the General Index of Retail Prices, and as this Index stood at 100 in January 1974, all revenues were adjusted to January 1974 prices before deseasonalization.

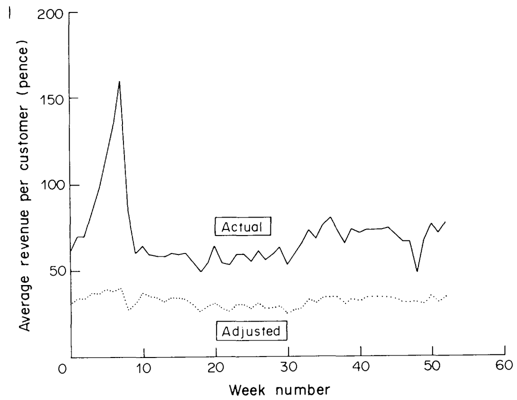

FIG. 1

REVENUE PER CUSTOMER

Having removed from the data the effects of inflation and seasonality, we begin to look for evidence of changes in the average revenue per customer. In order to signal the occurrence of likely change points, a cusum (cumulative sum) graph may be used. A cusum graph, as the name implies, is a plot of the sum of a set of observations x_i $(i = 1, \ldots. t)$ against the value t. The cusum is given by

$$S_t = \sum_{i=1}^{t} x_i,$$

and if the observations are drawn from a distribution with a fixed mean μ, then the expected value of S_t is $t\mu$, and so the graph of S_t against t can be expected to take the form of a straight line with slope μ. If, however, the mean of the distribution of x_i changes to $\mu + d$ at some point k, the expected value of S_t is now $k\mu + (t - k)(\mu + d) = t\mu + (t - k)d$, so that the slope of the line changes at $t = k$ from μ to $\mu + d$. In order to make any changes in slope more evident, it is usual to subtract a 'target value' X from each observation before the cusum is calculated, i.e.

$$S_t = \sum_{i=1}^{t} (x_i - X) = \sum_{i=1}^{t} x_i - tX.$$

The ideal value for X is the mean of the observations μ so that the cusum graph will have zero slope. If there is subsequently an increase in μ, the cusum graph will then have a

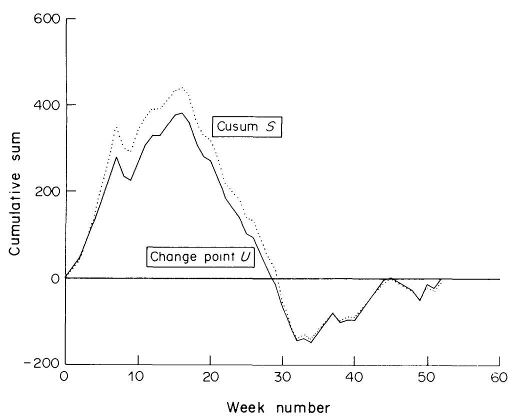

FIG. 2

positive slope, whereas with a decrease in μ, the slope would be negative. Figure 1 shows, for November 1978 onwards, the average revenue per customer in its raw form and also after deflation and deseasonalization. Figure 2 shows the cusum for the same period with a target value of $X = 32$. The change points are clearly seen in week 16 (week ending 24 February 1979) and week 32 (week ending 16 June 1979).

To corroborate the apparent changes in weeks 16 and 32, a non-parametric analysis due to Pettitt[1] was also performed. This calls for the computation of a Mann–Whitney type statistic $U_{t,T}$ for all partitions of the observations x_i ($i = 1, 2, \ldots, T$) into $x_1, x_2, \ldots, x_t$ and $x_{t+1}, x_{t+2}, \ldots, x_T$. $U_{t,T}$ is a measure of the 'precedence' of the first section of the data over the second in that it is the difference between $N_{1,t}$ and $N_{2,t}$, where $N_{1,t}$ is the number of times each observation in the first set $x_1, \ldots, x_t$ exceeds an observation in the second set $x_{t+1}, \ldots, x_T$, and $N_{2,t}$ is the number of times that each observation in the second set exceeds one in the first. For example, given the following five observations:

$$2, 4, 4, 3, 5,$$

partitioned into (2) and (4, 4, 3, 5), then $N_{11} = 0$ and $N_{21} = 4$ because the observation in the first set exceeds none of the observations in the second, whereas each of the four observations in the second set exceeds the one value in the first, i.e.

$$U_{1T} = 0 - 4 = -4.$$

Likewise,

$$N_{12} = 1, N_{22} = 4 \text{ and } U_{2T} = -3$$
$$N_{13} = 2, N_{23} = 4 \text{ and } U_{3T} = -2$$
$$N_{14} = 0, N_{24} = 4 \text{ and } U_{4T} = -4.$$

If $t = 0$ or T, this would result in a partition involving an empty set, so that in all cases

$$U_{0T} = U_{TT} = 0.$$

It may be shown that $U_{t,T}$ can be determined more simply by taking the rank of x_t, which we denote by R_t, and forming the quantity

$$V_{t,T} = 2R_t - (T + 1),$$

from which $U_{t,T}$ can be calculated using the formula

$$U_{t,T} = U_{t-1,T} + V_{t,T}.$$

From the above data, we obtain the following calculations:

t	x_t	R_t	$V_{t,T}$	$U_{t,T}$
				0
1	2	1	-4	-4
2	4	$3\frac{1}{2}$	1	-3
3	4	$3\frac{1}{2}$	1	-2
4	3	2	-2	-4
5	5	5	4	0

It can be seen that $U_{t,T}$ is itself a cusum of the quantity $V_{t,T}$, and to illustrate this, the graph of $U_{t,T}$ for the data of Figure 1 is also shown in Figure 2.

One of the main advantages of the non-parametric procedure over the straightforward cusum is that it permits the significance of any change point to be readily determined. It has been shown by Pettitt that the approximate probability (one-tailed) associated with either

$$K_+ = \max_{1 \leq t \leq T} (U_{t,T})$$

or

$$K_- = -\min_{1 \leqslant t \leqslant T} (U_{t,T})$$

is

$$P_K = \exp(-6K^2/T^2(T+1)). \tag{1}$$

For the data from November 1978, the values of K_+ and K_- are 381 and 148 respectively, occurring in weeks 16 and 32. Using equation (1), the approximate significance probabilities are 0.002 and 0.400, giving reasonable support to the occurrence of a decrease in revenue per customer in week 16. The apparent upward change in week 32 is not, however, shown up as being significant. With the occurrence of multiple change points, Pettitt[2] suggests a sequential procedure progressing from one change point to the next. Following this approach, if the change-point analysis is re-done with only the data from week 16 onwards, then a significant upward change occurs in week 32 at the 0.00005 level.

NUMBER OF CUSTOMERS

It was explained in the introduction that stealing by understriking often gives rise to an increase in the number of 'no sales' or 'zero totals' because of the attempt to disguise the incorrectly rung-in amount. All of the tills in use in the shop during the period of the investigations were such that whenever anything was rung into the till, it would register as a customer. This, then, would cause a large number of 'no sales' or 'zero totals' to inflate the number of customers figure.

To investigate whether or not the number of customers had changed significantly during the period of 1979 identified in the previous section, a second cusum and change-point analysis was applied to the deseasonalized number of customers. The cusum of the number of customers and the graph of $U_{t,T}$ for the period beginning November 1978 are shown in Figure 3.

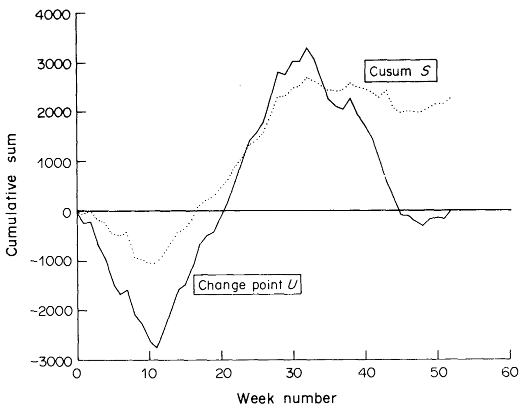

FIG. 3

The values of K_+ and K_- are 327 and 275 respectively, occurring in weeks 32 and 11. The first of these is significant at the 0.011 level, indicating a decrease in the number of customers in week 32. The second is significant at the 0.042 level, showing an increase in the number of customers during week 11. This is clearly a valuable corroboration of the previous change-point analysis, and the difference between this second analysis and the first in the timing of the upward shift in the number of customers is perhaps not surprising. At the best of times, the number of customers would be expected to be a quite volatile quantity, depending on such things as alterations in the pattern of trade generally within the town, the weather, the existence of special promotions and so on. Furthermore, if Mrs F had commenced stealing on a relatively small scale and gradually built up over time, the increase in the number of customers would almost certainly be apparent before the effect on revenue per customer.

THE EVIDENCE OF THE 'NO SALES'

The foregoing section has indicated quite conclusively that the number of customers decreased after week 32. As yet, however, no reason has been adduced for this, and the question that must be answered is whether or not this was due to a change in the number of no sales. This is very difficult to check upon, however, for the only record of no sales is provided by the till rolls, and these are normally kept for only a few weeks after they have been removed from the till and then thrown away. In this case it was possible to obtain six past till rolls, four from till A and two from till B. It turned out that three of the rolls came from before week 32, and the other three related to the period after week 32. A count of the number of 'no sales' and 'small totals' (i.e. 1p, 2p or zero) yielded the data of Tables 1 and 2.

TABLE 1. TILL ROLLS BEFORE WEEK 32

Till roll	Till	No sales	Zero and small totals	Days covered
1	A	—	92	10
2	A	—	136	12
3	B	270	48	29

'No sales' on till A register as 'zero totals'.

TABLE 2. TILL ROLLS AFTER WEEK 32

Till roll	Till	No sales	Zero and small totals	Days covered
4	A	—	24	8
5	A	—	51	13
6	B	141	11	29

Comparing the observed number of no sales and small totals with the expected number for each till produces the results in Table 3.

TABLE 3. ANALYSIS OF NO SALES AND SMALL TOTALS

	Till A		Till B	
	Observed	Expected	Observed	Expected
Before week 32	228	148.0	318	235.0
After week 32	75	155.0	152	235.0
Total	303	303	470	470
$\sum\frac{(O-E)^2}{E}$	70.34		56.83	

Both chi-squared statistics are overwhelmingly significant, which points to the conclusion that there were more no sales and small totals per day before week 32 than there were afterwards, which may well explain the change in the number of customers observed at or about week 32.

TILL DIFFERENCES

We have already commented on the fact that, with persistent stealing on a large scale, the thief can often lose control, and the amount of money subsequently removed from the till will be quite different from the amount actually understruck. Furthermore, if it is assumed that the thief will actually err on the side of caution, the amount of money actually in the till at the end of the day will more often than not be greater than the actual amount recorded; i.e. there will be a consistently positive till difference.

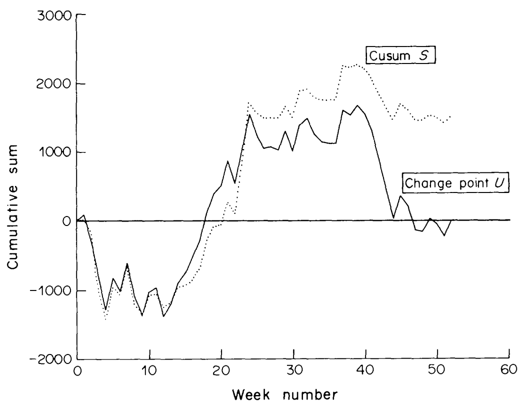

FIG. 4

Considering again the period from November 1978, a cusum of the till difference and a graph of $U_{t,T}$ is shown in Figure 4. From the graphs, the rise during the period of suspicion (weeks 16 to 32) is quite obvious, although from the change-point analysis, the change points (which occur in weeks 12 and 39) turn out to be insignificant. Taking a different view of the data, if the year from November 1978 is divided into the three periods which were identified in the foregoing analysis of average revenue per customer, i.e.

A week 1–week 16
B week 17–week 32
C week 33–week 52,

the mean and standard deviation of the weekly till differences are as shown in Table 4. The t statistics, also shown in Table 4, indicate that during periods A and C, the mean weekly till difference is not significantly different from zero, whereas during period B, the difference is significantly positive at the 0.025 level. To put this significance more into context, 20 other 16-week periods were chosen randomly from the remaining six years data, and the mean till difference was tested for significance. In no case was the mean value significantly different from zero. This seems to imply that, in what one can assume to be normal working, the till difference figures do tend to average zero, and apparently the only time when this was not so was again during the previously identified period of suspicion, when the mean difference was significantly positive.

TABLE 4. TILL DIFFERENCES, NOVEMBER 1978–OCTOBER 1979

Period	No. of weeks	Mean (£)	Standard deviation (£)	t
A	16	−0.54	3.40	0.635
B	16	+1.73	3.12	2.218
C	20	−0.20	1.64	0.545

'INTERVENTION' ANALYSIS

The analyses of the number of customers and the average revenue per customer have both indicated quite conclusively that clearly identifiable changes occurred during the week ending 16 June 1979 (week 32). This was, in fact, the week during which Mrs F was arrested, and so the analysis can be seen as definite confirmation of the known intervention which took place during that week. This, however, opens up the question of whether a change-point analysis is strictly appropriate in this case. There are, in fact, two quite different types of change point present. The first, the time at which the stealing started, is unknown; whereas the second, when Mrs F was arrested, is clearly known. Furthermore, the changes themselves may also be different. There is some evidence that the first change point was 'gradual', with possibly a build-up in the amount being stolen; whereas the second change point would almost certainly be abrupt. A change-point analysis would seem quite appropriate for the first (unknown) change, but the second change should probably more realistically be viewed as an example of an 'intervention' situation, as described by Box and Tiao.[3] The question posed by Box was: "Given a known intervention, is there evidence that a change in the series of the kind expected actually occurred and, if so, what can be said of the nature and magnitude of the change?"

The model proposed by Box assumes that any value in a series x_i ($i = 1, 2, \ldots, T$) can be represented as

$$x_i = f(\theta, \delta, i) + N_i \tag{2}$$

where f is some function of a set of unknown parameters (θ), the deterministic effects of time (i) and the effects of exogenous variables (δ), which may be interventions. N_i represents the stochastic background variation or noise. In particular, we suppose that the noise may be modelled by a mixed autoregressive moving average process of the form

$$\phi_1(B)\, N_i = \phi_2(B) a_i, \tag{3}$$

where B represents the backshift operator such that $Bx_i = x_{i-1}$, $\phi_1(B)$ and $\phi_2(B)$ are the autoregressive and moving average polynomials in B, and a_i ($i = 1,2, \ldots, T$) are a set of independently distributed normal variables with zero mean and variance σ^2 (white noise).

If we consider a single intervention at time t in the form of a pulse input δ_i such that

$$\delta_i = \begin{cases} 0 & i \neq t \\ 1 & i = t, \end{cases}$$

this will produce an output response in the form of a step change of size θ if the function f takes the following form

$$f(\theta, \delta, i) = \frac{\theta B}{1 - B} \delta_i. \tag{4}$$

This, then, leads to the following model for x_i:

$$x_i = \frac{\theta B}{1 - B} \delta_i + \frac{\phi_2(B)}{\phi_1(B)} a_i. \tag{5}$$

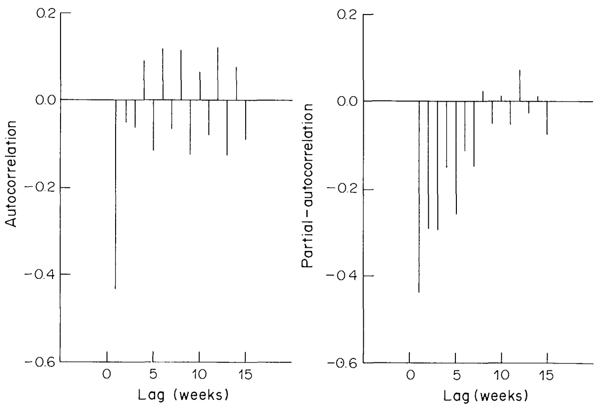

FIG. 5. *Autocorrelation and partial autocorrelation functions.*

If we consider now the data relating to the average revenue per customer, the autocorrelation and partial autocorrelation functions of the first differences for the 328 observations prior to the first change point are shown in Figure 5. This suggests an integrated moving average model for the noise of the form

$$(1 - B)\, N_i = (1 - \phi B) a_i, \tag{6}$$

leading to the following maximum likelihood estimates of the parameter ϕ and the residual standard deviation σ:

$$\hat{\phi} = 0.84 \quad \hat{\sigma} = 2.79$$

For an intervention of the type postulated, it may be shown (see Box and Tiao[3]) that a maximum likelihood estimator of the step change θ is obtained from the difference between two weighted averages, one of the observations up to the time of the pulse and the other afterwards, namely

$$\hat{\theta} = (1-\phi)\sum_{i=0}^{\infty} \phi^i x_{t+1+i} - (1-\phi)\sum_{i=0}^{\infty} \phi^i x_{t-i}, \tag{7}$$

where

$$\operatorname{var}(\hat{\theta}) = (1-\phi^2)\,\sigma^2. \tag{8}$$

There are 16 observations prior to the intervention (weeks 17 to 32 inclusive), and 20 observations were available for the period after the intervention (weeks 33 to 52). Making the appropriate corrections to the two weighted averages, the estimated increase in average revenue per customer is

$$\hat{\theta} = \frac{0.16}{1-0.84^{20}}\sum_{i=0}^{19} 0.84^i x_{33+i} - \frac{0.16}{1-0.84^{16}}\sum_{i=0}^{15} 0.84^i\, x_{32-i} = 4.83.$$

The estimated standard error of this change is

$$\mathrm{SE} = \sqrt{(1-0.84^2)} \times 2.79 = 1.51,$$

showing that there has been a real increase of about 4.8p in the average revenue per customer after week 32, which amounts to about £150 per week at June 1979 prices.

CONCLUSIONS

Considering overall the evidence of the previous four sections, there seems to be incontrovertible evidence that something quite dramatic happened during the week ending 16 June 1979 (week 32). This was the week Mrs F was arrested, and after that week the revenue per customer rose, the number of customers fell, the number of no sales recorded on the tills decreased, and the till difference figures seemed to return to normality. All the evidence then seems to point to the fact that Mrs F was stealing on a considerable scale prior to week 32, and when this was brought to an end there is a very clear indication in the data.

What perhaps is more questionable is when the stealing actually commenced. The revenue per customer showed a significant downturn following the week ending 24 February 1979 (week 16), but the increase in the number of customers seems to possibly precede this by one or two weeks. The most obvious explanation of this would seem to be that stealing on a small enough scale not to have been noticeable in the revenue figures was taking place during the early weeks of 1979, and then during the latter part of February 1979, the amount being stolen was increased to such an extent that it became evident in the customer revenue figures. It could, of course, be the case that stealing had been taking place for a number of years but that the scale of things has never been sufficient to make a noticeable difference to the shop's takings. This can only be surmise, but what seems quite clear is that by February 1979, what might have started on a very small scale had built up to such an extent that the shop's trading performance was being adversely affected.

ACKNOWLEDGEMENTS

The author is extremely grateful to Dr Tony Pettitt for many valuable comments and suggestions made during the preparation of this paper. Thanks are also due to Mr John Greenlees for work undertaken while he was an undergraduate student in the Department of Management Studies, Loughborough University of Technology.

REFERENCES

[1] A. N. Pettitt (1979) A non-parametric approach to the change point problem. *Appl. Statist.* **28,** 126–135.
[2] A. N. Pettitt (1980) A simple cumulative sum type statistic for the change point problem with zero–one observations. *Biometrika* **67,** 79–84.
[3] G. E. P. Box and G. C. Tiao (1975) Intervention analysis with applications to economic and environmental problems. *J. Am. Statist. Ass.* **70,** 70–79.