CUMULATIVE RANK SUM TEST:
THEORY AND APPLICATIONS

by

MICHEAL KEVIN THRAN

A.A. in Liberal Arts and Sciences, Joliet Junior College, 1972
B.S. in Psychology, Northern Illinois University, 1976

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Statistics

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1981

Approved by:

R. M. Robison

Major Professor

TABLE OF CONTENTS

LIST OF TABLES

FIGURE

ACKNOWLEDGEMENTS

I would like to express my heartfelt appreciation to Mike Rubison, my major professor, for his assistance, guidance, and most of all, support. This report would not have been possible without him. I would also like to thank, for their assistance, my committee members: James L. Hess, Doris L. Grosh, and Arthur Dayton. I would also like to thank Gary Gaeth and Michael Frerichs for their assistance with the computer.

THE CUMULATIVE RANK SUM TEST:
Theory and Applications

INTRODUCTION

Process control has long been a source of problems in quality control. For the purposes of this paper, process control will refer to the means of insuring that a given procedure, device, or process, is producing within a desired set of limits. If the limits are exceeded, we want to be able to detect this as soon as possible. It is also desirable to be able to determine what, if any, minor changes have occured and when they occured - this is useful information; for example, a new technique might improve/worsen the process. Therefore, process control implies two functions; 1) to insure that an acceptable level of quality is maintained, and 2) to provide a history of the process so that the effects of minor changes may be analyzed. A series of solutions to the problem of process control have been proposed over the years. These include the Shewhart chart, the sequential probability ratio test and the cumulative sum control chart.

The first of the control charts, the Shewhart chart (named for its originator Dr. Walter A. Shewhart), has become one of the most widely accepted charts in industry and has served as the model for most of the later procedures. The Shewhart chart, despite being quite simple to use, has been shown to be a very reliable test (Page, 1961; Duncan, 1965). The actual chart is just a plot of the sample means[1]. The basic premise of the Shewhart chart is that even if a process is under control there will be some random variation in its output. Assuming normality, it is expected that at least 99% of the observations should lie within three standard deviations of the process mean (also called the target value, goal value, or reference value). Since these variations are plotted on a chart, any systematic variation will quickly become visible.

As originally introduced, the Shewhart chart was a simple affair. Two action lines at 3 standard deviations above and below the process mean line were drawn. The action lines were so named because if a sample point crossed the lines the process was to be stopped and the appropriate action taken (i.e., return the process to within the limits). The resulting chart was then watched as the sample points varied about the process mean line (Page, 1961; Duncan, 1965).

Later improvements further enhanced the performance of the Shewhart chart. Since it is often costly to wait until a

---

[1] X_{ij} = the jth observation from the ith sample, where $i = 1, 2, 3, \ldots, j = 1, 2, 3, \ldots n$. The chart is just the plot of $(i, \overline{X}_i)$.

process is out of control before any action is taken (Page, 1961; Gibra, 1975), warning lines were added to the chart, usually at the two standard deviation levels. The rules for interrupting the process were then; if either a point crosses the action lines or, if n_1 points cross the warning lines within n_2 samples then stop. Obviously, the location of the warning and action lines as well as the choice of n_1 and n_2 would determine the characteristics of this procedure.

The success of the Shewhart chart can be seen in the number of related charts and tests that have appeared. For example, (i) for the fraction defective there is a p-chart (with fixed or variable n) and its associated chi-squared test. (ii) for the number of defects per unit there is the c-chart and u-chart with their associated tests. (iii) there is an X chart for variations in the sample variances. (iv) and an R-chart for variations in the ranges of the samples (In short, if there is an interest in a parameter there is probably a Shewhart style chart for monitoring it.) (Page, 1961; Duncan, 1965; Gibra, 1975).

While it might seem that a solution to the problem had been found, all was not well. A Shewhart chart requires a relatively large number of samples and is slow to detect persistant changes (particularly if they are small). There are also problems if the changes were infrequent (Page, 1961; Duncan, 1965; Gibra, 1975; Lucas, 1976). Two further approaches to this problem of process control were developed in the late 1940's and the 1950's, the sequential probability ratio test (Wald, 1947) and the cumulative sum procedure (Page, 1954).

It can be seen in the later modifications of the Shewhart chart that previous samples began to contribute to the decision process, e.g., when n_1 sample points cross the warning lines within n_2 samples a corrective action is taken. The next logical step would be to make better use of previous information in the decision rule. Page (1954) proposed just such a procedure with his first version of the cumulative sum test. His procedure makes use of not only the present sample but also looks at the difference between the present and past samples to see what changes might have occured. It is interesting to note that his argument expands the Shewhart chart but also makes use of the work on sequential probability ratio tests by Wald.

The first procedure proposed by Page was based on the differences of the sums of all the observations and the smallest sum of the previous observations (min S_n) .

Let X_1 be the ith observation,
then for

$$S_n = \sum_{i=1}^{n} X_i$$

we reject the hypothesis of no shift if $S_n - \min_{1 \leq i \leq n} S_i > h$.
This procedure was converted to a one-sided procedure by defining:

$$S_n' = \max(S_i' + X_i, 0) \qquad (i \geq 1) ,$$

$$S_0' = 0$$

and modifying the rejection rule as:

take action if $S_n' > h$.

This procedure was later modified by Page (1955) to record the cumulative sums of the deviations of the observations from the expected process mean. Thus,

$$S_n = \sum_{i=1}^{h} (X_i - k) \ , \ S_0 = 0$$

and we take action when

$$S_n - \min_{1 \leq i \leq n} S_i > h \ ,$$

where k is the expected process mean.

If the reference value (k) is chosen so that it is slightly larger than the expected level then the cumulative sum will have a negative trend under the null condition. Page (1955) suggested that when S_n goes below 0 set $S_n = 0$, thereby producing a one-sided test. Setting the reference value smaller than the expected level and making h negative with the corresponding change in the decision rule will result in a CUSUM test for a decrease in the process mean. Page attempted to make a two-sided test out of his procedure by combining a test for a decrease in the process mean (lower tail test) with a test for an increase in the process mean (upper tail test). His attempt did not work out since he could not find a suitable method for determining the average run length.

It should be noted that as with any new procedure, new terms are generated. The Wald sequential probability ratio test has its A.S.N. (average sample number) and the cumulative sum charts (CUSUM charts) have their A.R.L. (average run length). For the purposes of this paper, A.R.L. will be defined as the average number of samples taken before rejection occurs.

Computing the A.R.L. for a CUSUM procedure initially caused problems, particularly for the two-sided case (Page, 1961). However, Ewan and Kemp (1960) presented both tables and a nomogram for calculating an A.R.L. based on the solutions of integral equations that came out of the Wald test. Their method of calculating an A.R.L. depends on selecting an appropriate A.R.L.(L_0) under the null hypothesis and an A.R.L.(L_1) under the alternate hypothesis. If a reference value is chosen appropriately close to q (where q is the average of the expected process mean and an unacceptable process mean) then the A.R.L. for the null hypothesis is near its maximum for a given L_1 . This empirical approximation derived by Ewan and Kemp could also be used to chose the defining characteristics of a one-sided CUSUM scheme[2]. Another method for approximating the A.R.L. was derived by Reynolds (1975). This approximation was derived both from the similarity between the sequential probability ratio test and CUSUM and by using a Brownian motion approximation to the CUSUM. An interesting side note is that the Brownian motion approximation does not require the usual normality assumption.

In a later review paper by Ewan (1963), a method for computing the A.R.L. (L) was proposed. This method consisted of summing the A.R.L. (L_1) for a upper level test and the A.R.L. (L_2) for a lower level test; i.e., $L = L_1 + L_2$. Two years before Kemp presented his version of the CUSUM test, Barnard (1959) introduced his modifications to the CUSUM. Barnard is generally given credit for introducing the V-mask to the CUSUM. The V-mask is a V shaped rejection region which is placed over the CUSUM. In application, the V-mask is a cutout that is physically placed over the graph. The vertex of the mask is placed a distance d from the last point. If any of the previous CUSUM points are covered by the mask, the procedure is considered out of control. Obviously, the angle of the mask and the distance d are the factors that determine the A.R.L. for this procedure. Barnard suggested that the choice of d and the angle (2θ) should be determined by intelligent "cut and try". He went on to suggest finding a value, c, which would immediately signal that the process was out of control and using this term to assist in choosing d and θ;

$$c = p(d + a)\tan\theta$$

---

[2] $L_1 = 1 + h/(\mu_1 - k)$, where h is the maximum distance the CUSUM should be above its minimum, μ_1 is the unacceptable process mean, and k is the reference value. Thus for μ_1 given, L_1 h and k can be defined.

where a is the horizontal distance between succesive points and p is the scale for converting vertical distance on the chart to x-units (CUSUMS).

Until the early 1960's, most of the work on the CUSUM procedures, including the V-mask, had been done in England. At that time, the CUSUM technique was a popular procedure in industry primarily because of the advantages it had over the Shewhart procedure. However, there was a minor problem with the CUSUM which was its dependence on the assumption of normality[3]. Johnson and Leone (1962) published a series of papers on the CUSUM in the United States. The major importance of their work was the relationship shown between the Wald sequential probability ratio test and the CUSUM procedure. In particular, they developed a series of CUSUM tests for non-normal cases. The CUSUM V-mask procedures they developed include a procedure for Poisson variables as well as one for binomial variables. They also suggested some procedures for monitoring process variability (based on the sample variance and range).

The V-mask CUSUM test was not limited to just watching for small sudden changes in the process mean. Both Lucas (1973) and Barnard (1959) mention modifying the V-mask so it becomes more responsive to large changes. The modification was to make the shape of the V-mask parabolic at its vertex

---

[3]Ewan and Kemp (1962) also developed a CUSUM procedure for a Poisson variable.

since this would allow a large change to be detected more quickly. Barnard also brought forth the idea of using the CUSUM procedure to estimate parameters of the process. While he did suggest this concept, it did not seem to catch on.

By the late 1960's the CUSUM procedure was widely used and little new work was being done. Texts on industrial statistics, such as Duncan (1965), covered the subject thoroughly. Van dobben de Bruyn (1968) wrote a very comprehensive monograph on the CUSUM, which includes theoretical as well as applied derivations. The CUSUM had developed from its early stages in 1954 to a rather sophisticated procedure by the middle 1960's. It incorporated the best features of its predecessor, the Shewhart chart. Such features as a physical mask that allowed for testing by people not explicitly trained in process control[4] and an accesible history of the process made both the Shewhart chart and the CUSUM test very popular. A good review of the various control procedures is contained in a paper by Gibra (1975). Numerous tables and nomograms are available for choosing the parameters of the V-mask with a desired A.R.L. (Duncan, 1965; Beyer, 1966). It seemed that the CUSUM procedure had been developed to its fullest potential.

[4] It can be argued that the action limits and warning lines of the Shewhart chart are identical in function to the V-mask of the CUSUM test with respect to the ease of training people in their proper use.

In addition to the CUSUM test, the sequential probability ratio test, and the Shewhart chart, other procedures have been proposed. For instance, Sen and Srivastava (1975) proposed a one-sided test for the change in the process mean. Their method used a Bayesian test statistic that seemed to perform as well, if not better, than a more conventional test based on the maximum likelihood estimator. A more innovative and practical approach to the problem has been the introduction of nonparametrics. Rhodes (1960) developed two nonparametric tests corresponding to the sequential probability ratio test. As a nonparametric procedure, his was not encumbered by potentially unrealistic assumptions and under real world conditions performed more economically than its parametric counterpart.

Bakir and Reynolds (1979) proposed a nonparametric test that could be used for process control that was based on within-group ranking. A CUSUM type decision rule was used and like its parametric counterpart, their test could be related to an earlier sequential test (in this case Rhodes' procedure). When used in the proper conditions (i.e., the observations must be taken in groups) this test performs well but these same conditions are its shortcoming. An interesting nonparametric CUSUM test was brought forth by McGilchrist and Woodyer (1975). While their procedure did not readily lend itself to monitoring an ongoing process, it does deserve mention. They used the same basic idea of summing the deviations from their expected value. In their case, the estimate of the expected value was the sample median. This forces the test to be used in a more postmortem fashion.

In his review paper, Gibra brought up some areas in process control that need further work. One of the points he made was the unrealistic distributional assumptions that are made in order to use some procedures. He also mentioned that there ought to be a simple technique that can be administered with ease. The nonparametric test that is introduced here is a simple procedure that is similar in form to the V-mask CUSUM. It will be shown to have similar qualities without some of the more restrictive assumptions of its parametric parent.

NOTATIONS AND ASSUMPTIONS

Throughout this paper the following notations and definitions will be used.

Let X_i denote the ith observation from some continuous distribution. Define S_n to be the sum of these observations, $X_1, X_2, X_3, \ldots, X_n$, i.e.,

$$S_n = \sum^{n} X_i \, .$$

In addition, S_r will denote a partial sum of the X_i (from 1 to r), where $1 \leq r \leq n$. The difference, $S_n - S_{n-r}$, is then

$$S_n - S_{n-r} = \sum_{n-r+1}^{n} X_i \, .$$

The following definitions describe the parameters associated with the V-mask CUSUM procedures.

n.....The sample size.

k.....The reference value. Traditionally, the value of k is chosen either slightly larger or smaller than the process mean. This will result in the CUSUM having, respectively, a negative or positive trend when no change in distribution has occurred.

d.....The distance from the last CUSUM coordinate plotted and the vertex of the V-mask.

θ.....The angle made by the upper (or lower) limb of the V-mask with the horizontal. Therefore 2θ is the full angle of the V-mask (see figure 1).

We assume that each X_i has a distribution function $F_i(X) = F(X - \mu_i)$ for some continuous distribution function $F(\ )$. The null hypothesis is that all μ_i are equal. We propose the CUSUM test of H_0 against the alternate hypothesis of a distributional shift of the form;

(i) $\mu_1 = \mu_2 = \mu_3 = \cdots\cdots = \mu_c < \mu_{c+1} < \cdots\cdots < \mu_n$

or

(ii) $\mu_1 = \mu_2 = \mu_3 = \cdots\cdots = \mu_c > \mu_{c+1} > \cdots\cdots > \mu_n$

for some unknown $c > 1$.

FIGURE 1. A CUSUM CHART WITH V-MASK

A (n, S_n)

B $(n - r, S_n)$

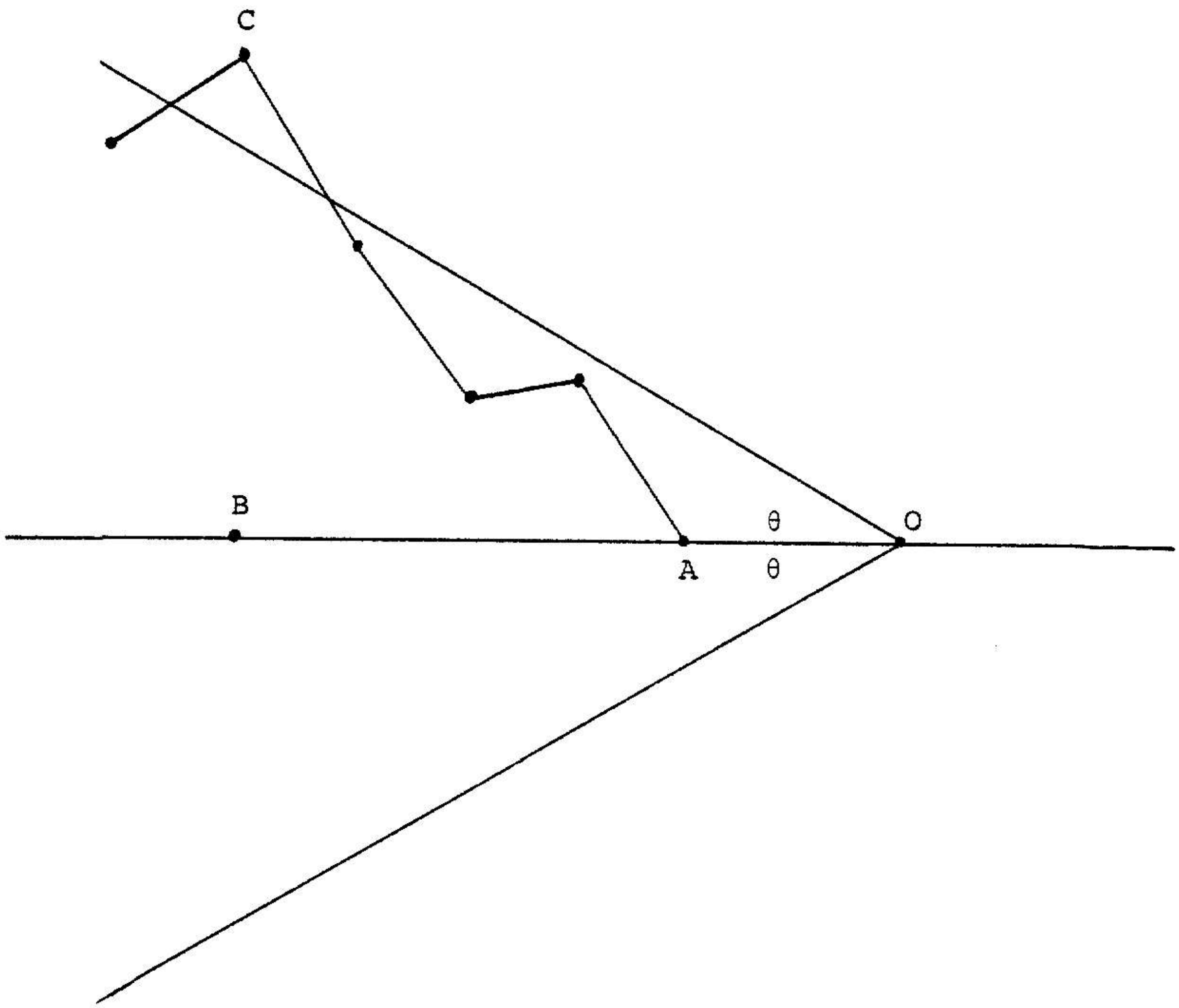
C $(n - r, S_n - r)$

O $(n + d, S_n)$

$\overline{BA} = r$

$\overline{AO} = d$

$h = d \tan \theta$



THE CUMULATIVE RANK SUM TEST

The cumulative rank sum test (CURSUM test), which we propose as the nonparametric equivalent of the CUSUM test, uses the following notations and definitions:

Let R_{ij} be the rank of X_i among $(X_1, X_2, X_3, \ldots, X_j)$ where $i \leq j$ and each X_i has a continuous distribution function $F(X - \mu_i)$ with $F(0) = 1/2$. Under H_0 $\mu_1 = \mu_2 = \mu_3 = \ldots$, the cumulative ranks R_{jj} are independent with mean

$$E_0(R_{jj}) = (j + 1)/2$$

and variance

$$Var_0(R_{jj}) = (j - 1)(j + 1)/12 .$$

Define S_n' to be the sum of the deviations of the R_{jj} from their mean i.e.,

$$S_n' = \sum^{n} \{R_{jj}/(j + 1) - 1/2\}.$$

under the null hypothesis,

$$E_0(S_n') = 0$$

with variance

$$Var_0(S_n') = \{n - 2 \sum^{n} (1/(j + 1))\}/12n .$$

The random variables $R_{jj}/(j + 1)$ are distributed as independent discrete uniform variables on the points

$\{1/(j+1)\ ,\ 2/(j+1),\ \ldots\ j/(j+1)\}$ under H_0. In effect, the standard CUSUM procedure based on the partial sums of

$$S_n = \sum^{n} (X_i - k)$$

has been transformed so that X_i has a discrete uniform (0,1) distribution on j points, with a reference value of $k = 1/2$.

The V-mask rejection region will reject H if for some n and $r \leq n$,

$$\sum_{n-r+1}^{n} \{R_{jj}/(j+1) - 1/2\} < -(r+d)\tan\theta$$

or

$$\sum_{n-r+1}^{n} \{R_{jj}/(j+1) - 1/2\} > (r+d)\tan\theta\ ,$$

where d is defined as with the CUSUM test. It should be noted that the proposed test will fail to reject the null hypothesis (regardless of the alternative hypothesis) if the slope θ used in the V-mask is greater than $\tan^{-1}(1/2)$.

Like its parametric counterpart, the A.R.L. is a measure of the performance of the CURSUM. The A.R.L. is defined as the average number of samples taken before the null hypothesis is rejected. Unlike the CUSUM, which has three parameters, the CURSUM test has only two parameters that affect the performance (i.e., the A.R.L.) of the test.

SIMULATIONS

A series of simulations were run in order to determine the performance characteristics of the CURSUM test. The Average Run Lengths (A.R.L.s) were determined under the null condition for a series of d lengths and selected θ values. These A.R.L.s are contained in Table 1 and were derived from 900[5] replications of sequences of standard normal variables. The A.R.L. and power for the CURSUM was also determined under two types of alternative hypotheses:

(1) A shift of Δ (= 1, 2 standard deviations) at a point c (= 10, 20, 30), i.e.,

$$\mu_i = \begin{cases} 0 & i = 1, 2, 3, \ldots c \\ \Delta & i = c + 1, c + 2, \ldots \end{cases} .$$

Tables 2a-2e[6] contain the A.R.L. and power for condition one with selected values of d and θ .

(2) A persistent change of Δ (= .1) in the mean at a point c (= 10, 20, 30), i.e.,

---

[5]The starred values were not derived from 900 replications since computer time considerations made this impossible.

[6]As the value of d was increased, changes of 1 standard deviation and the larger θ values were dropped from the simulation due to the impractically large A.R.L.s and the computer time required to generate them.

$$\mu_i = \begin{cases} 0 & i = 1, 2, 3, \ldots c \\ (i - c)\Delta & i = c + 1, c + 2, \ldots \end{cases}$$

Tables 3 and 3a display the A.R.L.s and power for condition two for selected values of θ and d .

When operating under the alternative hypotheses, the A.R.L. serves a different purpose than it does under the null hypotheses. When a change has occured in the sequence, the location of the change is usually desired. Furthermore, it is preferable that the test reject quickly after the shift has occured. We might note that some of the A.R.L.s are very close or "sensitive" to the change point, while others are less sensitive. It will be shown that the choice of theta and d will determine the sensitivity of the A.R.L. to c , though usually at the cost of power.

From Table 1 and 2 it can be seen that as a given V-mask is moved back from the last CURSUM coordinate (i.e., as d is increased) the A.R.L. increases. Thus, if a large A.R.L. is desired then a large value of θ and/or d should be selected. With d held constant (at 1) it can be seen that as θ increases the power improves though with some loss of sensitivity for small changes and for changes early in the sequence. The case of a θ of .30 with a 2 standard deviation change in the mean at the 30th sample has an A.R.L. of 33.945 with power .9308 , while an equal change in the mean with a θ of .27 yeilds an insensitive A.R.L. of power .6267 . Futhermore, with a θ of .25 and below and a d

of 1 , the power is nearly zero with an A.R.L. equal to the A.R.L. under H_0 .

As d is increased (Tables 2a-2e) the performance of the smaller values of theta (.15 - .25) improves. While θ of .30 has power 1 whenever $d > 1.5$ its A.R.L. becomes very insensitive to the change point (c). If you compare various levels of d at a given $\underline{\theta}$ (say .20) for a set change ($c = 10$, $\Delta = 2$) there will be a slight increase in the A.R.L. but a large increase in power as d goes from 1 to 2.5 , i.e., A.R.L. goes from 9.554 to 15.344 but power increases from .4667 to .967 .

Another interesting result of the simulations was the detection of small changes (1 standard deviation) anywhere in the sequence. For values of θ of .25 and above, there was good power but very insensitive A.R.L.s which increased as d was increased. When a small theta was used with a large d value a test sensitive to those small changes was found. Large changes early in the sequence could be picked up with a medium θ (.25) and a low d ($d < 1.5$) or a small theta ($\theta < .25$) with a large d ($d > 2$) .

It appears that for the first type of alternative (1) the following rules apply;

(i) When keeping c and Δ constant for a given theta, power increases as d increases.

(ii) As d increases the A.R.L. will increase (again all else held constant) usually until the A.R.L. becomes so much larger than the shift point as to be of little use.

(iii) With d small ($d < 1.5$), a large value of theta ($\theta > .25$) is required to accurately detect large changes ($\Delta = 2$) in the sequence.

(iv) As d is increased, the smaller values of theta ($\theta < .20$ become sensitive to small changes in the sequence.

(v) For small and large changes in the sequence and ($c < 20$), the intermediate values of theta ($.20 < \theta < .25$) gives the best general performance.

When the alternative hypotheses is one of a persistent change (2), the CURSUM test performs very well. From Table 3 and 3a it can be seen that when the change occurs late in the sequence ($c = 30$) a large value of theta with small d works the best. As d is increased theta must be decreased to maintain power but there is some loss of sensitivity in the A.R.L.. When the change starts early in the sequence a small theta with a large d is required to adequately detect the shift. Similarly as theta is increased, d must be decreased, again with a loss of A.R.L. performance. If the shift occurs at a more moderate location in the sequence ($c = 20$) a theta in the middle range with a corresponding d performs nicely. For instance, $\theta = .25$ with a $d = 1.5$ has an A.R.L. of 22.205 and power of .9117 , while you must increase d for smaller values of theta ($\theta < .25$) and decrease d for larger values of theta ($\theta > .25$) to get an equivalent performance of the A.R.L..

It was mentioned earlier that estimating the shift point (c) in the sequence would be desirable. While some of the

A.R.L.s in the simulations were close to the shift point, most were not within five samples of the shift point. This is not to suggest that the A.R.L. is meant to be an estimator, but rather that it can serve in some cases as a very good upper bound. The method that seems to work the best is to use the smallest value which exceeds the test critical value as your estimate of c. This crossover point (cp) is $n - r$ in the partial sum;

$$S_n - S_{n-r} = \sum_{n-r+1}^{n} X_i \ .$$

Table 4 displays the frequency distribution of the crossover points for $\theta = .25$ and $d = 1$ under the two types of alternatives considered.[7] As can be seen from Table 4, the mode of the crossover point distribution is c. This holds true for all values of c, theta and d generated in this study. There is a distinct relationship between the size of delta and the percentage of cps for a given c. For $\Delta = 1$, the percentage of crossover points before c is 14.77% . When $\Delta = 2$, the percentage is 50% within one of c. For a persistent change the percentage is 50% within 2 samples; though this percentage is larger for $c = 20$. In particular, for test statistics with low power the estimate will, when it is in error, underestimate the shift point. The conditions of $d = 1, 2$ when

[7]This theta was chosen since it was representative of the A.R.L.s and powers that were encountered in the simulations.

$c = 20$ are examples of this. When power is in a more preferred state (> .9), the distribution of the crossover points tend to be more symmetrical, e.g., when $d = 1, 2$ and $c = 10$ are good examples. It should be noted that Leone and Johnson (1961) also recommend using this method to estimate the crossover or out-of-control point, while Lucas (1976) uses a variant of the procedure.

APPLICATIONS AND EXAMPLES

The use of a CURSUM test is quite simple. The preferred method is to make use of any knowledge there might be about possible changes in the process. That is, if information on the size of the shift and/or the location of the suspected shift is available then this information may be useful in selecting theta and d . However, if little is known about the process then θ and d may be selected on the basis of the A.R.L. of the test under H_0 .

When selecting θ and d with some information about the process, the relative merits of A.R.L. and power must be weighed. For instance, suppose there was reason to believe that a shift might occur around the 20th sample and with a 2 standard deviation change in magnitude. If the process is expensive to stop and/or some error is tolerable, then large power is preferable to low A.R.L.. In this type of situation, the selection of (θ = .25, d = 2) or (θ = .20, d = 4) has power and A.R.L. given respectively by (.9966, 67.063) and (1.00, 26.04). However, if errors are expensive to make then the A.R.L. becomes more important relative to power. In this case, either of the selections of (θ = .20, d = 2.5) or (θ = .20, d = 3) yields good A.R.L. with reasonable power (A.R.L.s are 22.011 and 22.94 with power of .86 and .90 respectively). These are by no means the only values of theta and d to be considered for these cases, but rather, just examples of how theta and d might be selected under different considerations of A.R.L. and power.

The first example uses data from a paper by Lucas (1980). His paper compares the standard CUSUM procedure with an improved method. The improvement he suggests is a "fast initial response" (FIR) which is to simply start S_n at a value other than zero. This would accelerate the CUSUM procedure. Table 5 shows the original data as well as the values of R_{jj} and S_m. The lower limb critical points were calculated from the following formulas:

$$H_L = (r*\tan\theta) + S_n' - ((n + d)\tan\theta)$$

or

$$H_U = -(r*\tan\theta) + S_n' + ((n + d)\tan\theta)$$

The data was shifted at the 11th observation by a delta of 1 standard deviation.

The standard CUSUM with an A.R.L. of 22 rejected H_0 with a run length of 16. The FIR-CUSUM had an A.R.L. of 16 and rejected H_0 at 13. The CURSUM had an A.R.L. of 15.07 ($d = 1.5$, $\theta = .20$). The run length was 13 for the CURSUM test with the cp at 10, which was the last point before the shift was started. Unfortunately neither the cp nor the power of the two CUSUM tests were given so that a complete comparison could not be made.

Table 5. Example of CURSUM test versus FIR and standard CUSUM on data from Lucas (1980).

Table 5

	X_i	R_{jj}	S_r	H_L
1	1.0	1	0	-2.5982
2	-.5	1	-.1666	-2.3955
3	0	2	-.1666	-2.1928
4	-.8	1	0.4666	-1.990
5	-.8	1.5	-.7166	-1.7874
6	-1.2	1	-1.0737	-1.5847
7	1.5	7	-.6987	-1.3820
8	-.6	4	-.7542	-1.1793
9	1.0	7.5	-.5042	-.9766
10	-.9	2	-.8223	-.7738*
11	1.2	11	-.4056	
12	.5	8	-.2902	
13	2.6	13	.1383**	
14	.7	8	.1716	
15	1.1	12	.4216	
16	2.0	15	.8039***	

* CURSUM crossover point

** CURSUM and FIR run length

*** CUSUM run length

The second example uses real data. The modified data comes from Leone and Johnson (1962), the modification being that .003 was subtracted from the last six observations. The observations are the means ($n = 5$) of heights of fragmentation bomb bases.

Table 6 displays the data with the CURSUM test. Only the upper limb of the test is listed since the direction of the change was known. The test statistic used ($\theta = .25$, $d = 1$) has an A.R.L. of 21.843 with power .9323 under the alternative hypothesis of 1 standard deviation change at $c = 10$. The CUSUM had an A.R.L. of 23.2 for the same alternative.

It can be seen that the CURSUM rejects at a run length of 13 . It is interesting to note that even though both procedures have the same cps, the CURSUM has a run length of 13 while the CUSUM has a run length of 16 .

Table 6. CURSUM test (θ = .25, d = 1) on data from Leone and Johnson (1962).

Table 6

Run Length

	X_i	R_{jj}	S_n	H_U
1	.8324	1	0	2.1600
2	.8706	1	-.1667	1.9047
3	.8262	1	-.4166	1.6493
4	.8326	4	-.1166	1.394
5	.8290	2	-.2833	1.1386
6	.8316	4	0.2119	.8833
7	.8336	7	-.1631	.6279
8	.8310	4	0.2186	.3726
9	.8336	9	.1813	.1173*
10	.8306	3	-.0459	-.1380*
11	.8302	2	-.3792	-.3933*
12	.8258	1	-.8023	-.6487
13	.8280	2	-1.1594**	
14	.8264	1	-1.6529	
15	.8292	6	-1.7779	
16	.8228	1	-2.219***	

* CURSUM crossover points

** CURSUM run length

*** CUSUM run length

CONCLUSION

The development of control chart procedures for process control has centered on modifications of the rejection regions and stopping rules for the normal distribution theory case. Even though some work has been done on nonnormal distributions, all previous CUSUM procedures assume that the underlying distribution of the observations are known.

Recently, work on procedures which are independent of distributional assumptions have been reported. While these nonparametric procedures perform well, they do not have the same ease of application of the CUSUM tests. The cumulative rank sum test proposed in this paper is shown to be a distribution free procedure with the same type of V-mask rejection region as the CUSUM.

Since the CURSUM procedure has fewer parameters it is easier to implement than the CUSUM. In addition, the CURSUM does not require an estimate of the process variance, therefore scaling factors for graphical displays are not needed. This ease of operation, however, does not sacrifice the ability of the CURSUM to closely monitor a process or to detect a change in the process. The CURSUM procedure also provides an accurate estimate of the location of the out-of-control point.

As with the CUSUM, further development of the nonparametric control chart needs to be continued. Several possible areas of future work are;

1) Development of a CURSUM procedure which varies values of both theta and d as the samples are taken. This should increase the sensitivity of the test to early shifts in the process while providing both good power and A.R.L.s if the shift should occur later in the process.

2) A distribution-free Shewhart type control chart for data collected in groups could be developed. While it would use a cumulative rank procedure the rejection region would differ significantly from the CURSUM.

3) A CURSUM procedure for detecting shifts in variability could be developed, though it should be noted that the CURSUM can, in its present form, be used for such a test.

REFERENCES

Bakir, S.T. and M.R. Reynolds Jr., (1979) 'A nonparametric procedure for process control based on within group ranking', Technometrics, Vol. 21, No. 2, pp. 175-183.

Barnard, G.A., (159) 'Control charts and stochastic processes', Journal of the Royal Statistical Society, B, Vol. 21, No. 2, pp. 238-257.

Beyer, W.H., Editor, (1966) Handbook of Tables for Probability and Statistics. CRC Press, Cleveland.

Duncan, A.J., (1965) Quality Control and Industrial Statistics, Richard D. Irwin, Inc. Homewood, Illinois.

Ewan, W.D., (1963) 'When and how to use the CUSUM charts', Technometrics, Vol. 5, No. 1, pp. 1-22.

Ewan, W.D. and K.W. Kemp, (1960) 'Sampling inspection of continuous processes with no auto-correlation between successive results', Biometrika, Vol. 47, No. 3, pp. 363-380.

Gibra, I.N., (1975) 'Recent developments in control chart techniques', Journal of Quality Technology, Vol. 7, No. 4, pp. 183-192. Sum control charts' Technometrics, Vol. 17, No. 1, February.

Johnson, N.L. and Leone F.C., (1962) 'Cumulative sum control charts', Part I, Industrial Quality Control, Vol. 18, No. 12, June 1962, pp. 15-21.

Johnson, N.L. and Leone F.C., (1962) 'Cumulative sum control charts', Part II, Industrial Quality Control, Vol. 19, No. 1, July 1962, pp. 29-36.

Johnson, N.L. and Leone F.C., (1962) 'Cumulative sum control charts', Part III, Industrial Quality Control, Vol. 19, No. 2, August 1962, pp. 22-28.

Johnson, N.L. and Leone F.C., (1977) Statistics and Experimental Design, Vol. 2, second ed., John Wiley & Sons New York.

Kemp, K.W., (1961) 'The avarage run length of a cumulative sum chart when a V-mask is used', Journal of the Royal Statistical Society, B, Vol. 23, No. 2, pp. 149-153.

Lucas, J.M., (1973) 'A modified V-mask control scheme', Technometrics, Vol. 15, pp. 833-847.

Lucas, J.M., (1980) 'Fast initial response (FIR) for cumulative sum quality control schemes', ASA meeting, Houston, Texas, August, 1980.

Lucas, J.M., (1976) 'The design and use of V-mask control schemes', Journal of Quality Technology, Vol. 8, No. 1, pp. 1-12.

McGilchrist, C.A. and K.D. Woodyer, (1975) 'Notes on a distribution free cusum technique', Technometrics, Vol. 17, No. 3, August, pp. 321-325.

Page, E.S., (1954) 'Continuous inspection schemes', Biometrika, Vol. 41, pp. 100-115.

Page, E.S., (1955) 'A test for a change in a parameter occurring at an unknown point', Biometrika, Vol. 42, pp. 523-527.

Page, E.S., (1957) 'On problems in which a change in a parameter occurs at an unknown point', Biometrika, Vol. 44, pp. 248-252.

Page, E.S., (1961) 'Cumulative Sum Charts', Technometrics, Vol. 3, No. 1, February, pp. 1-9.

Reynolds, M.R., (1975) 'Approximation to the average run length in cumulative sum control charts', Technometrics, Vol. 17, No. 1, February.

Rhodes, L.J., (1962) Two Sequential Rank Tests, Master's Thesis, Florida State University.

Sen, A. and M.S. Srivastava, (1975) 'Some one sided tests for change in level', Technometrics, Vol. 17, No. 1, pp. 61-71.

Van Dubben de Bruyn, C.S., (1968) Cumulative Sum Tests, Hafner, New York.

Wald, A., (1947) Sequential Analysis. Wiley, New York.

Table 1. Average Run Length for the CURSUM test under H_0 .

θ	d	A.R.L.	θ	d	A.R.L.
0.10	1	3.691	0.20	1	11.292
	1.5	4.366		1.5	20.213
	2	5.013		2	32.480
	2.5	6.670		2.5	64.894
	3	7.933		3	112.87
	4	10.433	0.25	1	30.422
0.15	1	5.993		1.5	59.018
	1.5	8.123		2	124.21
	2	11.773		2.5	130.303
	2.5	15.286	0.27	1	44.646
	3	19.986		1.5	103.51
	4	35.316	0.30	1	110.15

Table 2. A.R.L. and power for CURSUM with $d = 1$ and $\Delta = 1, 2$.

θ	C	Δ	A.R.L.	power
0.20	10	1	10.354	0.4812
	20	1	11.342	0.4756
	10	2	9.554	0.4667
	20	2	11.182	0.0433
0.25	10	1	21.843	0.9323
	20	1	72.654	0.6423
	10	2	14.164	0.9367
	20	2	19.490	0.6300
0.27	10	1	33.118	0.9856
	20	1	30.134	0.8078
	30	1	32.184	0.6232
	10	2	19.651	0.9889
	20	2	21.844	0.8045
	30	2	28.205	0.6267
.30	10	1	106.37*	1.00
	20	1	72.77*	1.00
	30	1	59.878	0.9339
	10	2	61.142	1.00
	20	2	29.704	0.9772
	30	2	33.945	0.9308

Table 2a. A.R.L. and power for CURSUM with $d = 1.5$ and $\Delta = 1, 2$.

θ	C	Δ	A.R.L.	power
0.20	10	1	15.706	0.7834
	20	1	17.486	0.407
	10	2	11.896	0.7834
	20	2	16.433	0.4068
0.25	10	1	41.736	0.9519
	20	1	36.010	0.9174
	30	1	39.265	0.7915
	10	2	22.613	0.9831
	20	2	22.905	0.9091
	30	2	30.524	0.7878
0.30	20	2	67.19*	1.00
	30	2	38.703	1.00

Table 2b. A.R.L. and power for CURSUM with $d = 2.0$ and $\Delta = 1, 2$.

θ	C	Δ	A.R.L.	power
0.20	10	1	23.45	0.93
	20	1	22.856	0.667
	30	1	26.243	0.445
	10	2	13.547	0.923
	20	2	19.743	0.668
	30	2	24.856	0.446
0.25	10	1	63.06*	1.00
	20	1	53.325*	1.00
	30	1	50.395*	1.00
	10	2	31.23*	1.00
	20	2	26.063	0.9966
	30	2	35.83	0.9765
0.30	20	2	97.2*	1.00
	30	2	66.04*	1.00

Table 2c. A.R.L. and power for CURSUM with $d = 2.5$ and $\Delta = 1, 2$.

θ	C	Δ	A.R.L.	power
0.15	10	1	13.03	0.6565
	20	1	14.87	0.29
	10	2	10.893	0.666
	20	2	14.423	0.31
0.20	10	1	32.73	0.965
	20	1	28.806	0.851
	30	1	32.773	0.7
	10	2	15.344	0.967
	20	2	22.016	0.86
	30	2	29.213	0.71
0.25	10	2	76.11*	1.00
	20	2	35.5	1.00
	30	2	37.77	1.00

Table 2d. A.R.L. and power for CURSUM with $d = 3.0$ and $\Delta = 1, 2$.

θ	C	Δ	A.R.L.	power
0.15	10	1	12.846	0.7767
	20	1	17.64	0.4367
	10	2	11.756	0.7676
	20	2	16.623	0.4360
0.20	10	2	21.5*	.98
	20	2	22.94*	.90
	30	2	31.42*	.84

Table 2e. A.R.L. and power for CURSUM with $d = 4.0$ and $\Delta = 1, 2$.

θ	C	Δ	A.R.L.	power
0.15	10	1	22.096	0.93
	20	1	22.98	0.66
	30	1	27.073	0.51
	10	2	13.643	0.927
	20	2	20.066	0.6867
	30	2	25.549	0.512
0.20	10	2	31.23*	1.00
	20	2	26.04*	1.00
	30	2	34.6*	1.00

Table 3. A.R.L. for CURSUM with d = 1 and persistent change.

θ	C	Δ	A.R.L.	power
0.20	10	0.1	9.975	0.6423
	20		11.127	0.9377
0.25	10	0.1	15.475	0.9389
	20		19.310	0.620
.027	10	0.1	18.382	0.990
	20		21.416	0.8167
	30		27.892	0.5978
0.30	10	0.1	24.733	1.00
	20		24.23	0.9778
	30		31.991	0.9443

Table 3a. A.R.L. and power for CURSUM with selected d and persistent change.

θ	d	C	A.R.L.	power	θ	d	C	A.R.L.	power
.15	2.5	10	11.706	0.6534	.25	2	30	32.6866	0.9766
		20	15.29	0.29		2.5	10	26.325	1.00
	3	10	14.396	0.7767			20	25.435	0.99
		20	16.583	0.4367			30	33.7933	1.00
	4	10	15.183	0.93		3	10	29.345	1.00
		20	19.903	0.6867			20	26.680	0.995
.20	1.5	10	13.150	0.7834			30	26.553	0.996
		20	16.76	0.5533		4	10	34.21	1.00
	2	10	15.11	0.93			20	29.13	1.00
		20	19.563	0.667			30	36.11	1.00
		30	24.566	0.445	.30	1.5	10	33.513	1.00
	2.5	10	16.756	0.9667			20	26.665	1.00
		20	49.815	0.8537			30	33.890	1.00
		30	49.474	0.71		2	10	39.486	1.00
	3	10	18.45	0.9966			20	29.47	1.00
		20	23.2266	0.9666			30	35.125	1.00
		30	31.63	0.8966		2.5	10	43.1433	1.00
	4	10	21.506	1.00			20	32.47	1.00
		20	24.673	0.9966			30	36.64	1.00
		30	33.603	0.9866		3	10	51.644	1.00
.25	1.5	10	19.028	0.9845			20	36.13	1.00
		20	22.205	0.9117			30	38.315	1.00
		30	29.752	0.7856		4	10	51.644	1.00
	2	10	22.577	1.00			20	42.885	1.00
		20	23.967	0.98			30	42.075	1.00

Table 4. Location of crossover points for theta of .25 .

θ C	1 10	2 10	.1 10	1 20	2 20	.1 20
1	1	1	1	1	1	1
2	2	2	2	2	2	3
3	2	0	1	2	0	3
4	14	12	6	13	7	8
5	2	12	8	2	3	4
6	8	15	10	12	15	12
7	25	34	29	14	21	25
8	56	64	39	27	19	25
9	51	133	77	19	23	15
10	118*	421*	213*	21	28	25
11	61	67	100	37	37	32
12	53	31	82	28	25	23
13	38	24	60	19	24	27
14	34	9	56	26	27	24
15	39	15	47	32	36	36
16	28	7	35	21	30	33
17	34	10	43	27	41	50
18	25	3	21	40	62	66
19	25	7	16	70	121	133
20	19	7	21	147*	324*	322*
21	20	3	8	66	27	22
22	6	1	8	42	14	10
23	19	2	5	24	5	0
24	20	1	4	32	3	1
25	14	4	2	26	0	0
26	16	3	0	19	0	0
27	11	1	0	10	2	0
28	6	2	2	8	1	0
29	7	3	1	17	0	0
30	13	2	1	8	0	0
31	9			11		
32	7			4		
33	7			10		
34	8			3		
35	7			4		
36	4			3		
37	5			9		
38	4			4		
39	2			3		
40	4			1		
41	5			3		
42	2			0		
43	5			1		
44	3			3		
45	2			2		

CUMULATIVE RANK SUM TEST:
THEORY AND APPLICATIONS

by

MICHEAL KEVIN THRAN

A.A. in Liberal Arts and Sciences, Joliet Junior College, 1972
B.S. in Psychology, Northern Illinois University, 1976

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Statistics

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1981

ABSTRACT

A distribution-free test is proposed for detecting shifts in a location parameter in sequential procedures. This Cumulative Rank Sum test (CURSUM) has wide applications as a process control procedure because it is computationally simpler than the standard Cumulative Sum test (CUSUM) since it depends on fewer parameters. Empirically derived Average Run Lengths are generated via simulations for both the null hypothesis and selected alternatives. This paper also includes a concise history of process control procedures including CUSUM tests as well as the Shewart chart.