

## A COMPARISON OF PHASE I CONTROL CHARTS

M.L.I Coelho<sup>1</sup>, S. Chakraborti<sup>2,3</sup> & M.A. Graham<sup>4</sup>

<sup>1,2,4</sup>Department of Statistics  
University of Pretoria, South Africa  
<sup>1</sup>margicoelho@gmail.com, <sup>4</sup>marien.graham@up.ac.za

<sup>3</sup>Department of Information Systems, Statistics and Management Science  
University of Alabama, U.S.A.  
schakrab@cba.ua.edu

### ABSTRACT

A carefully done Phase I analysis is a vital part of an overall statistical process control and monitoring regime. Distribution-free control charts can play a useful role in this analysis, as a parametric model assumption often cannot be adequately verified. The performance of two distribution-free charts for the location - the mean-rank and the median chart - are compared in this paper. For benchmarking reasons, the parametric  $\bar{X}$ -chart is included in the comparisons. It is seen that the distribution-free charts are in-control robust, whereas the  $\bar{X}$ -chart is not, and they both perform similarly to the  $\bar{X}$ -chart for normally distributed data. However, for non-normal data, they both outperform the  $\bar{X}$ -chart. The results provide evidence in favour of using distribution-free Phase I charts in practice. Concluding remarks and suggestions for future research are given.

### OPSOMMING

'n Versigtige analise van Fase I is 'n belangrike deel van 'n algehele statistiese proseskontrole en monitering skema. Verdelingsvrye kontrolekaarte kan 'n nuttige rol in hierdie fase speel, aangesien die model aannames nie voldoende geverifieer kan word nie. Die prestasie van twee verdelingsvrye kontrolekaarte - die gemiddelde-rank en die mediaan kontrolekaarte - word vergelyk in hierdie artikel. Vir vergelykingsredes, is die parametriesse  $\bar{X}$ -kontrolekaart ingeluit in die vergelykings. Dit is gevind dat die verdelingsvrye kontrolekaarte robuuste eienskappe het, terwyl die  $\bar{X}$ -kontrolekaart dit nie het nie, en beide presteer op soortgelyke wyse as die  $\bar{X}$ -kontrolekaart vir normaalverdeelde data. Egter, vir nie-normale data vaar albei beter as die  $\bar{X}$ -kontrolekaart. Hierdie resultate bied bewyse ten gunste van die gebruik van verdelingsvrye Fase I kontrolekaarte in die praktyk. Slotopmerkings en voorstelle vir toekomstige navorsing word gegee.

## 1 INTRODUCTION

In Statistical Process Control (SPC) there are generally two phases or stages (see, for example, [1]). The initial phase is called Phase I, in which a retrospective analysis is performed on existing (historical) data in order to establish that a process is stable, i.e., in-control (IC), and thus to find an IC (or a reference) sample. This is an extremely important stage of process control. When process parameters are unknown or unspecified, a Phase I analysis must be done before process monitoring can begin, prospectively, in Phase II. Thus, most Phase II control charts assume that a reference sample is available from a corresponding Phase I analysis, from which the control limits can be estimated. Thus the success of prospective process monitoring in Phase II depends critically on the success of the Phase I analysis.

According to Montgomery [1], Shewhart-type control charts are particularly useful for Phase I analysis because they are easy to construct and interpret and they are effective in detecting large, sustained shifts in the process parameters and outliers, measurement errors, data recording, and/or transmission errors and the like, which are perhaps expected in the early part of a process development. For a thorough account of the Phase I control charting literature, see Chakraborti et al. [2]. Recognising and highlighting the importance of Phase I analysis, a more recent comprehensive account of it is given in Jones-Farmer et al. [3]. This reinforces the need for an effective Phase I analysis using appropriate methods and metrics. Many of the available control charts in the literature are parametric control charts, based on an assumption of a particular underlying process distribution, such as the normal. The performance of these charts generally deteriorates when the underlying assumptions are not met. In fact, it is well-known that the Shewhart charts are very sensitive to the normality assumption, in that their false alarm rates are inflated, often significantly, when the underlying distribution deviates from normality. Moreover, it is often not possible to know much about the underlying distribution in the early stages of process development and study (i.e., Phase I), so that a specific distributional assumption (such as normality) cannot be reasonably justified; see for example Woodall [4]. In these settings, nonparametric control charts can provide a useful and robust alternative to the practitioner. Two major advantages of the nonparametric charts are that one does not have to make an assumption about the form and the shape of the underlying process distribution, and that the IC properties of nonparametric (or distribution-free) charts (such as the false alarm probability (*FAP*)) remain known and constant (stable) over all continuous distributions. This makes the nonparametric charts particularly suitable for Phase I applications, since if the IC properties of a chart are not stable and robust - its further applications in Phase II become somewhat meaningless. For a list of advantages and a thorough account of the fast-growing nonparametric control charts literature, see Chakraborti et al. [5].

Although a lot of work has been done in nonparametric SPC over the last few years, much of this has been for prospective or Phase II process monitoring, which assumes the availability of an IC reference sample from a Phase I analysis. Although some work has been done in monitoring scale using Phase I charts (see, for example, Jones-Farmer and Champ [6]), in this paper our focus is on the Phase I analysis and monitoring of location parameters so as to find the reference sample. As we noted earlier, the success of an overall SPC monitoring and control regime depends on a successful Phase I analysis, and nonparametric control charts can be especially useful in this setting. In this paper, we compare the performance of two of the available nonparametric Shewhart-type Phase I control charts according to their IC and out-of-control (OOC) performance based on the *FAP*, which is the probability of at least one false alarm out of  $m$  Phase I samples (where a false alarm is the event that a single charting statistic plots on or outside the control limits when the process is IC). Note that the *FAP* is the metric usually used in Phase I, as the signalling events are dependent.

Of the two nonparametric charts included in our study, the first chart is the mean-rank chart, which is based on the standardised subgroup mean-rank, as proposed by Jones-

Farmer et al. [7]. This chart is distribution-free and, consequently, it maintains a constant and known  $FAP$  for all continuous distributions. The second nonparametric chart in our study is the one proposed by Graham et al. [8], which is based on the subgroup precedence counts from the pooled median. This chart is also distribution-free and thus maintains a stable and known IC performance for all continuous distributions. Both of these distribution-free charts are effective at detecting shifts in the location of the process and useful to establish control in Phase I. We emphasise that, to date, a head-to-head performance comparison of the mean-rank and median charts is not available; we intend, therefore, to fill that gap in this study. This is important information, since practitioners need to know which of these nonparametric charts should be used, and when. However, although the focus is on nonparametric charts, we include a normal theory (parametric) chart in our study for benchmarking purposes. This is the Phase I  $\bar{X}$ -chart for the mean, proposed by Champ and Jones [9], which is based on the assumption of normality. When the underlying process is indeed normally distributed, this chart is expected to achieve an  $FAP$  very close to or equal to the nominal  $FAP_0$ , and should be more effective in detecting shifts in the process mean with a higher signalling probability. However, when the process distribution deviates from the normal distribution, its efficacy becomes questionable. This is typical of parametric charts; and this is why their IC robustness is always a practical concern.

We study the performance of all three charts in a simulation study for different combinations of subgroup size  $m$  and sample size  $n$ . First, we introduce the Phase I control charts, starting with the two nonparametric charts.

## 2 PHASE I CONTROL CHARTS

### 2.1 Phase I mean-rank chart

A nonparametric Phase I Shewhart-type control chart was proposed by Jones-Farmer et al. [7]. This chart is based on the well-known Kruskal-Wallis test (see Gibbons and Chakraborti [10]). We begin by combining the observations from the  $m$  Phase I samples in a single pooled sample of size  $N = m \times n$  and ordering the observations from the lowest to the highest. Then ranks ( $R_{ij}$  where  $i = 1, 2, \dots, m$  and  $j = 1, 2, \dots, n$ ) are assigned to each observation of this pooled sample. The average rank for the  $i^{\text{th}}$  sample is given by

$$\bar{R}_i = n^{-1} \sum_{j=1}^n R_{ij} \quad \text{for } i = 1, 2, \dots, m. \quad (1)$$

The charting statistics for the mean-rank chart are the standardised sample mean-ranks, given by

$$Z_i = \frac{\bar{R}_i - E(\bar{R}_i|IC)}{\sqrt{\text{var}(\bar{R}_i|IC)}} \quad \text{for } i = 1, 2, \dots, m. \quad (2)$$

Since the expected value and the variance of the ranks, when the process is IC, are given by  $E(R_{ij}) = \frac{N+1}{2}$  and  $\text{var}(R_{ij}) = \frac{(N-1)(N+1)}{12}$  (see Gibbons and Chakraborti [10], page 192), the expected value and the variance of  $\bar{R}_i$  for an IC process are given by  $E(\bar{R}_i) = \frac{N+1}{2}$  and  $\text{var}(\bar{R}_i) = \frac{(N-n)(N+1)}{12n}$ , respectively. Thus,  $Z_i$  in Equation (2) simplifies to

$$Z_i = \frac{\bar{R}_i - \frac{N+1}{2}}{\sqrt{\frac{(N-n)(N+1)}{12n}}} \quad \text{for } i = 1, 2, \dots, m. \quad (3)$$

Jones-Farmer et al. [7] studied two choices of the control limits: the simulated and the approximate normal theory control limits. For the latter choice, note that by the central limit theorem for large  $n$ , the standardised mean rank  $Z_i$  approximately follows a standard normal distribution when the process is IC. However, the charting statistics  $Z_1, Z_2, \dots, Z_m$  are dependent random variables, and this dependence needs to be properly accounted for. Appealing to available results in the literature, Jones-Farmer et al. [7] claimed that, asymptotically, the joint distribution of  $Z_1, Z_2, \dots, Z_m$  can be approximated by a multivariate normal distribution with means equal to zero, standard deviations equal to one, and a

common pairwise correlation  $\rho_{ij} = \sqrt{1/(m-1)}$ . Since this correlation tends to zero as  $m$  increases (for equal subgroup sizes), the authors suggested that the lower control limit (*LCL*) and the upper control limit (*UCL*) can be approximated using the quantiles of the (univariate) standard normal distribution. The control limits are chosen such that the attained *FAP* does not exceed a given  $FAP_0 = \alpha$ . Using Monte Carlo simulations, the authors provided tables for the control limits using the normal approximation for  $\alpha = 0.05$  or  $0.10$ . The centre line of the proposed chart, *CL*, is equal to zero. It was recommended that the simulated limits, and not the approximate normal theory tables, be used, since the latter were more conservative (see Jones-Farmer et al. [7]). In short, these control limits are based on the simulated empirical distribution of the standardised mean rank,  $Z_i$ , and vary based on the combination of  $m$  and  $n$ . Although exact limits are possible to be obtained by enumerating the distribution of the standardised mean rank, doing so is impractical for moderate values of  $m$ . Accordingly, the authors provided control limits, for specific  $(m, n)$  combinations using 100,000 Monte Carlo simulations. They recommended that a conservative approach be used so that the attained *FAP* does not exceed a given  $FAP_0 = \alpha$ ; i.e., the control limits are found using Monte Carlo simulations so that  $FAP \leq \alpha$ . Using simulations, we extended the tables of the control limits in Jones-Farmer et al. [7] by considering some of their combinations of  $m$  and  $n$  and incorporating other combinations of  $m$  and  $n$ . This should be of benefit in practice.

We describe the second nonparametric Phase I control chart next.

## 2.2 Phase I median chart

An alternative to the mean-rank chart, proposed by Graham et al. [8], is called the median chart. This chart is based on the well-known Mood test (see Gibbons and Chakraborti [10], page 323). Suppose that  $M$  denotes the median of the pooled sample of size  $N$ . The charting statistic for the median chart is the subgroup precedence statistic - i.e., the number of observations that precede  $M$  - and is given by  $U_i = \sum_{j=1}^n I(X_{ij} < M)$  for  $i = 1, 2, \dots, m$ , where  $M$  denotes the pooled median and  $I(A) = 1$  or  $0$  when  $A$  is true or false.

As in the case of the mean-rank statistics, the charting statistics  $U_1, U_2, \dots, U_m$  are dependent, and so the IC joint distribution of  $U_1, U_2, \dots, U_m$  is necessary in order to calculate the *FAP* and the charting constant  $a$ . For a more detailed derivation of the IC joint distribution and of the *FAP*, refer to Graham et al. [8]. Since the IC distribution is symmetrical about its mean, the authors propose using symmetrically placed control limits with  $LCL = a$ ,  $UCL = n - a$  and  $CL = 0$ .

Since the joint distribution of the  $U_i$ 's is discrete, a conservative approach is used where  $a$  is found as the largest positive integer, such that the attained *FAP* is less than or equal a specified  $FAP_0$ . It can be shown that

$$FAP = FAP(m, n, UCL = n - a, LCL = a) = 1 - \sum_{u_1=a+1}^{n-a-1} \sum_{u_2=a+1}^{n-a-1} \dots \sum_{u_m=a+1}^{n-a-1} f_{U_1, U_2, \dots, U_m}(u_1, u_2, \dots, u_m). \quad (4)$$

Thus  $a$  is the largest positive integer, so that

$$a = \max\{c: FAP_0 \geq 1 - \sum_{u_1=c+1}^{n-c-1} \sum_{u_2=c+1}^{n-c-1} \dots \sum_{u_m=c+1}^{n-c-1} f_{U_1, U_2, \dots, U_m}(u_1, u_2, \dots, u_m)\}. \quad (5)$$

Graham et al. [8] found that some of the attained *FAP* values can be very conservative, meaning that they may be much smaller than the  $FAP_0$  value. They provided tables for the control limits for several combinations of subgroup size  $m$  of size  $n$ , for  $FAP_0$  values 0.01, 0.05, and 0.10 respectively. They noted that the sample size  $n$  needs to be much larger than the number of Phase I sample  $m$  in order to achieve industry-standard small *FAP* values.

Graham et al. [8] also showed that the correlation between any two charting statistics, when the process is IC, is  $-1/(m-1)$ , which tends to zero as  $m$  increases. Consequently, we can approximate the control limits by ignoring the dependence among the charting statistics, simply using the marginal univariate hypergeometric distribution of the  $U_i$ 's.

Although our main purpose is to compare the two nonparametric Phase I charts head to head, for benchmarking and robustness considerations, we also include the normal theory Phase I  $\bar{X}$ -chart proposed by Champ and Jones [9]. This is briefly described in the next section.

### 2.3 Phase I $\bar{X}$ chart

A Phase I  $\bar{X}$ -chart was proposed by Champ and Jones [9], which assumed that the underlying distribution is normal. The charting statistic for this chart is the mean,  $\bar{X}$ , of each Phase I sample, and the control limits are given by

$$\bar{\bar{X}} \pm k \hat{\sigma} / \sqrt{n} \quad (6)$$

where  $\bar{\bar{X}}$  is the grand mean; i.e.,  $\bar{\bar{X}} = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n X_{ij}$  and  $X_{ij}$  denotes the  $j^{\text{th}}$  observation from the  $i^{\text{th}}$  subgroup. The unbiased estimator of the process standard deviation  $\sigma$  is based on the average sample variances, and is given by  $\hat{\sigma} = \frac{1}{c_4} \left[ \frac{1}{m} \sum_{i=1}^m S_i^2 \right]^{\frac{1}{2}}$ , where  $c_4$  is an unbiasing constant and  $S_i^2$  is the  $i^{\text{th}}$  sample variance. The constant  $k > 0$  is the chart design constant.

Champ and Jones [9] provided tables for the charting constant  $k$  for several combinations of  $m$  and  $n$ . They also studied the choice of the control limits for  $m \geq 20$ , using simulations, with (i) exact limits based on the multivariate  $t$  distribution, (ii) approximate limits using the univariate Student's  $t$  critical values, and (iii) approximate limits assuming that  $T_{v,i}$  approximately follows a standard normal distribution (this is only recommended where  $m \geq 30$ ). The reader is referred to their paper for more details. Briefly speaking, the charting constant  $k$  (i.e., the control limit) is found using 100,000 Monte Carlo simulations for a specific  $(m, n)$  combination so that the attained  $FAP$  does not exceed the given  $FAP_0 = \alpha$ . The authors recommended that when  $m < 20$  subgroups are available, the constants given for the multivariate  $t$  distribution are to be used to compute the limits. However, for  $m \geq 20$ , the authors suggested that the constants be obtained using the (univariate)  $t$  distribution with  $m(n-1)$  degrees of freedom.

## 3 COMPARATIVE STUDY

### 3.1 Implementation of the control charts

In order to compare the performance of the control charts, first the control limits need to be found; we discuss this here. To this end, note that the discreteness of the two nonparametric charting statistics causes the attained  $FAP$  values to be sometimes very conservative and often much smaller than a nominal value - say, 0.05. Thus we chose combinations of  $m$  and  $n$  values that yielded attained  $FAP$  values closest to the nominal value of 0.05. This is also one of the reasons why larger values of  $n$  were used, which is not the typical approach when implementing variables control charts. However, note that this (larger  $n > m$ ) is often necessary for the case of attributes control charts where the charting statistics have discrete distributions, as in the case of the nonparametric charts. Since exact calculations are tedious and involve multidimensional sums and integrals, simulation was used. In Table 1, the control limits for the two nonparametric charts are provided, along with the constant  $k$  for the  $\bar{X}$ -chart and the attained  $FAP$  values, for a  $FAP_0 = 0.05$  and a number of combinations of  $m$  and  $n$  values. When investigating which combinations of  $m$  and  $n$  to consider, recall that for the median chart, the sample size  $n$  needs to be much larger than the number of Phase I sample  $m$  in order to achieve industry-standard small  $FAP$  values. Accordingly, we considered smaller values of  $m$  ( $= 4, 5, 6, 7, 8$ ,

9, and 10) paired with larger values of  $n$  (=15(1)25), obtained the corresponding  $FAP$  values, and selected the pairs that yielded the attained  $FAP$  values closest to the nominal value of 0.05. We included one larger value of  $m$  (=50), together with the value of  $n$  (=15), which yielded an attained  $FAP$  value closest to the nominal value of 0.05. Table 1 will be useful to implement the Phase I control charts under consideration in practice.

**Table 1: The constant  $k$  for the  $\bar{X}$ -chart, the control limits for the mean-rank chart, and the constant  $a$  (LCL) for the median chart, for a  $FAP_0 = 0.05$ , along with the attained  $FAP$  values**

$m$	$n$	$\bar{X}$ -chart		Mean-rank chart		Median chart	
		$k$	Attained $FAP$	Control limits ( $\pm$ )	Attained $FAP$	$a = LCL$	Attained $FAP$
4	18	2.529	0.0496	2.452	0.0497	4	0.0479
5	16	2.608	0.0495	2.532	0.0499	3	0.0486
5	19	2.608	0.0496	2.547	0.0499	4	0.0434
7	17	2.722	0.0499	2.647	0.0499	3	0.0493
7	20	2.721	0.0493	2.650	0.0499	4	0.0450
8	23	2.749	0.0495	2.697	0.0499	5	0.0495
9	15	2.811	0.0498	2.718	0.0499	2	0.0406
9	18	2.809	0.0492	2.733	0.0499	3	0.0411
10	15	2.840	0.0498	2.755	0.0499	2	0.0472
10	18	2.841	0.0497	2.757	0.0499	3	0.0480
10	21	2.827	0.0497	2.761	0.0497	4	0.0456
10	24	2.826	0.0491	2.768	0.0497	5	0.0415
50	15	3.306	0.0499	3.196	0.0497	1	0.0429

### 3.2 Performance comparisons

A comparison of the chart performance was done by means of extensive Monte Carlo simulations using 50,000 iterations in SAS<sup>®</sup> v 9.3. For each iteration,  $m$  samples, each of size  $n$ , were generated from one of three distributions: the standard normal distribution ( $N(0,1)$ ), the symmetric but heavier-tailed Student's  $t$  distribution with three degrees of freedom ( $t(3)$ ), and the right-skewed Gamma distribution with a shape parameter of two and a scale parameter of one ( $Gamma(2,1)$ ). We compare both the IC and the OOC performance of the Phase I control charts. The IC performance indicates how robust the chart is with respect to the specified  $FAP_0$  value; the nonparametric charts are expected to be IC robust (because they are distribution-free), but since the charting statistics are discrete, it is of interest to see what the attained  $FAP$  values are and how far off they might be from the nominal values. On the other hand, the OOC comparison involves comparing the probabilities of alarm (at least one signal) under some 'OOC condition' when the charts have the same or roughly the same  $FAP_0$  (i.e. the same IC performance). The chart with the highest probability of at least one signal under the OOC condition is favoured.

#### 3.2.1 IC performance

The control limits for the two nonparametric charts were found from Table 1. For the  $\bar{X}$ -chart, the control limits were set up using the constant  $k$  from Table 1. Then, using simulations, the attained  $FAP$  values were found and compared with the  $FAP_0$  of 0.05. Figures 1(a) to 1(c) show the IC performance of the three charts for the  $N(0,1)$ ,  $t(3)$ , and  $Gamma(2,1)$  distribution respectively. Since the  $\bar{X}$ -chart assumes normality, it is expected that this chart will perform well in the case of  $N(0,1)$  distribution. From Figure 1(a) it can be seen that that is the case. However, the mean-rank chart also has attained  $FAP$  values very near the nominal level of 0.05. The median chart also performs well; but its attained  $FAP$  values are more conservative. Thus, in the IC case, all three charts seem to perform well in the case of the normal distribution, performing at or close to the  $FAP_0$  values.

When the underlying distribution is symmetric but more heavy-tailed, such as the  $t(3)$  distribution, as shown in Figure 1(b), the mean-rank chart still provides attained  $FAP$  values very close to the  $FAP_0$ . The median chart also performs well, but its attained  $FAP$  values are again seen to be conservative. The  $\bar{X}$ -chart, however, performs very poorly, with attained

*FAP* values larger than the nominal level and in fact increasing as  $m$  increases; for  $m = 50$  it even reaches values larger than 0.25, which is unacceptable.

Finally, when the underlying distribution is skewed, such as the *Gamma*(2,1) distribution, the  $\bar{X}$ -chart performs very poorly, with attained *FAP* values larger than the nominal level. In addition, the attained *FAP* values increase as  $m$  increases, as in the case of the  $t(3)$  distribution; for  $m = 50$  it reaches values just under 0.10. This is shown in Figure 1(c). However, the mean-rank chart is seen to provide attained *FAP* values very close to the nominal value. The median chart also performs well, but the attained *FAP* values are again shown to be conservative.

In summary: from Figures 1(a) to 1(c) it is seen that the mean-rank chart has the best overall IC performance; however, the median chart is competitive, and performs well although somewhat conservatively (in terms of the attained *FAP* values). Although the  $\bar{X}$ -chart performs well when the data follows a standard normal distribution, it performs quite poorly in the non-normal cases - i.e., for the  $t(3)$  and the *Gamma*(2,1) distributions. Thus the lack of IC robustness of the Phase I  $\bar{X}$ -chart should be a serious concern for its implementation in practice.

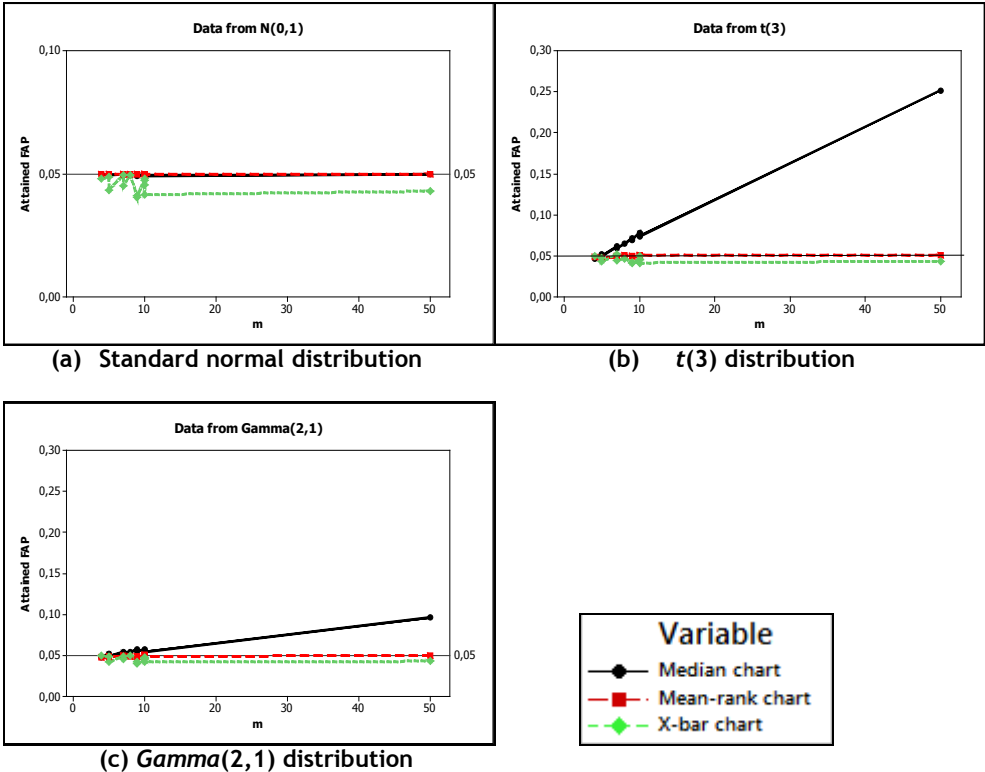


Figure 1: The attained *FAP* values for various distributions when  $FAP_0 = 0.05$  and the process is IC (see online for colour version)

### 3.2.2 OOC performance

As noted earlier, the OOC comparison involves comparing the probabilities of alarm (at least one signal) for the charts under different distributions and shift configurations. Following Jones-Farmer et al. [7], two configurations were considered: (i) an isolated mean shift (a shift in only one subgroup mean and in none of the other subgroup means) of size  $\delta$  (in units of standard deviations), with the shift introduced in the first subgroup without loss of generality; and (ii) a sustained mean shift (a shift in one of the subgroup means and sustained throughout the remaining subgroup means) of size  $\delta$  (in units of standard

deviations), with the shift introduced two-thirds of the way into the sample and sustained until the end of the sample. This means that the mean stayed IC until two-thirds of the way into the sample, of size  $m$ , at which point a constant shift was observed in the mean and was sustained until the end of the sample. For example, say  $m = 30$ , then the mean stayed IC for the first 20 observations with a constant shift occurring at observation number 21 which was then sustained through samples 21 to 30. Decreases as well as increases in the mean were considered with shifts of size  $\delta = -3(0.5)3$ . Since the shift is expressed in units of the process standard deviation  $\sigma$ , the amount of the shift in the mean,  $\sigma\delta$ , is given by  $1\delta$ ,  $\sqrt{3}\delta$  and  $\sqrt{2}\delta$ , for the standard normal,  $t(3)$ , and  $Gamma(2,1)$  distributions respectively. The OOC mean is the IC mean plus  $\sigma\delta$ , with  $\delta \neq 0$ .

For the OOC comparisons, it was necessary to re-calculate the chart design constant  $k$  for the  $\bar{X}$ -chart for the  $t(3)$  and the  $Gamma(2,1)$  distributions, so that all three charts have the same or nearly the same  $FAP$  value (when the process is IC) to ensure a fair OOC performance comparison. Table 2 shows these 'adjusted'  $k$  values, which were obtained via simulation. Note that this was not necessary for the nonparametric charts, as their IC performance remains the same for all continuous distributions.

**Table 2: Adjusted  $k$  values for the  $\bar{X}$ - chart for non-normal distributions**

$m$	$n$	Adjusted $k$	
		$t(3)$	$Gamma(2,1)$
4	18	2.503	2.522
5	16	2.503	2.623
5	19	2.625	2.619
7	17	2.622	2.745
7	20	2.793	2.735
8	23	2.853	2.789
9	15	2.955	2.865
9	18	2.941	2.849
10	15	3.029	2.908
10	18	3.009	2.891
10	21	2.989	2.874
10	24	2.984	2.864
50	15	4.555	3.591

### 3.2.2.1 Isolated shifts

The OOC performance for the isolated shift case was studied for all combinations of  $m$  and  $n$  in Table 1. However, owing to space constraints we report the results for the combinations:  $(m, n) = (4, 18)$ ,  $(10, 24)$ , and  $(50, 15)$ , for the standard normal,  $t(3)$ , and the  $Gamma(2,1)$  distributions respectively in Tables 3 to 5. The chart that performs the best - i.e., detects shifts with the highest probability - is indicated with the use of grey shading. If more than one cell is highlighted, it indicates that the charts are performing similarly. The remaining results are similar. The sample size  $n$  did not seem to have much of an effect on the performance, especially since the sample sizes were close to each other. It can be seen that the shift is detected with a smaller probability as the number of subgroups,  $m$ , increases. This could be due to the fact that the isolated shift is only observed in the first subgroup. As the subgroup size increases, the impact of the shift becomes watered down as there are more IC observations. Next, for the two symmetric distributions ( $N(0,1)$  and  $t(3)$ ), the direction of the shift did not seem to make a difference with respect to the alarm probability of the charts. From Table 3, it can be seen that the  $\bar{X}$ -chart performs best under the normal distribution, detecting shifts with the highest probability. However, as in the isolated shift case, the mean-rank chart's performance is again seen to be very close to that of the  $\bar{X}$ -chart. For example, for a shift of  $\delta = 0.5$ , the mean-rank chart's performance is only 4.83 per cent below that of the  $\bar{X}$ -chart, and for a larger shift of  $\delta = 1.0$ , the mean-rank chart's performance is only 2.37 per cent below that of the  $\bar{X}$ -chart. The median chart detects the shifts with a lower probability. Overall, all three charts seem to perform better for a smaller subgroup size than for a larger subgroup size.



**Table 3: Attained alarm probability values for the standard normal distribution for  $FAP_0 = 0.05$  and for an isolated shift**

$\delta$	$m = 4$ and $n = 18$			$m = 10$ and $n = 21$			$m = 50$ and $n = 15$		
	$\bar{X}$	Mean-rank	Median	$\bar{X}$	Mean-rank	Median	$\bar{X}$	Mean-rank	Median
-3.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998
-2.5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9953
-2.0	1.0000	1.0000	0.9990	1.0000	1.0000	0.9996	1.0000	0.9999	0.9519
-1.5	0.9985	0.9975	0.9609	0.9999	0.9997	0.9771	0.9931	0.9836	0.7328
-1.0	0.8802	0.8594	0.6669	0.9384	0.9194	0.6954	0.7164	0.6574	0.3036
-0.5	0.2939	0.2787	0.1915	0.2971	0.2803	0.1652	0.1267	0.1176	0.0693
0.0	<b>0.0494</b>	<b>0.0496</b>	<b>0.0476</b>	<b>0.0500</b>	<b>0.0500</b>	<b>0.0459</b>	<b>0.0500</b>	<b>0.0500</b>	<b>0.0424</b>
0.5	0.2924	0.2783	0.1923	0.2934	0.2808	0.1643	0.1265	0.1199	0.0679
1.0	0.8802	0.8598	0.6685	0.9393	0.9179	0.6894	0.7141	0.6549	0.3000
1.5	0.9983	0.9969	0.9613	0.9999	0.9998	0.9772	0.9935	0.9847	0.7308
2.0	1.0000	1.0000	0.9991	1.0000	1.0000	0.9996	1.0000	0.9999	0.9532
2.5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9960
3.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998

For the  $t(3)$  distribution, the control limits for the  $\bar{X}$ -chart were obtained using Table 2. From Table 4, it can be seen that the mean-rank chart detects the mean shifts with the highest probability for a small as well as a large subgroup size. The median chart performs the second best, and the  $\bar{X}$ -chart shows the poorest results. For  $m = 50$  and  $n = 15$ , we see that the median chart at first outperforms the  $\bar{X}$ -chart, but for larger shifts, the  $\bar{X}$ -chart seems to reach an alarm probability of 1 faster than the median chart. Again, it is observed that the smaller the subgroup size, the easier it is for the chart to detect the isolated shift.

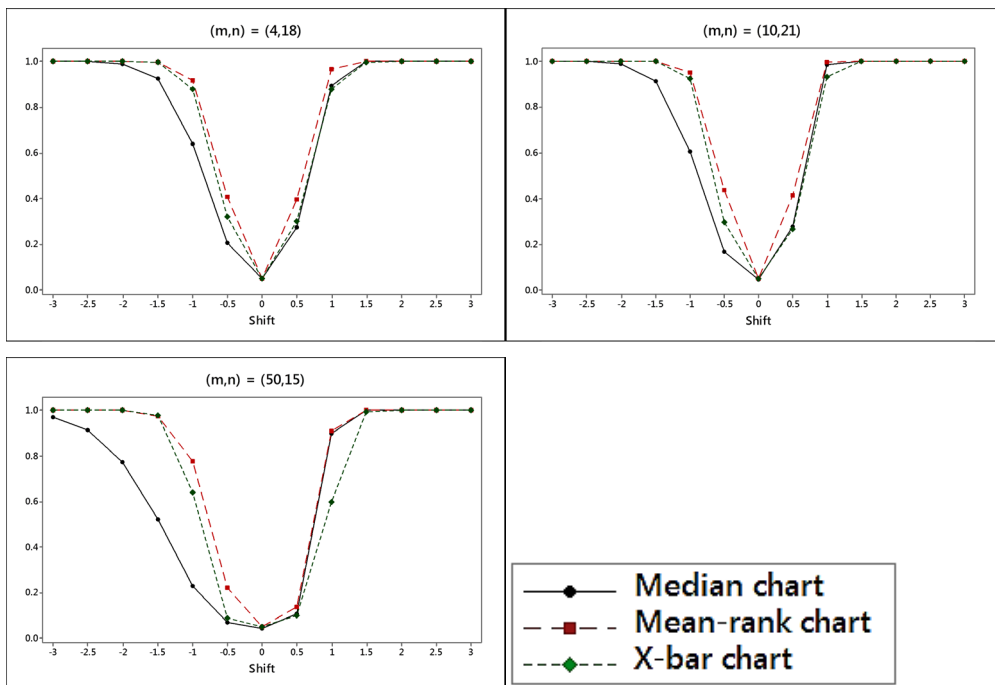
**Table 4: Attained alarm probability values for the  $t(3)$  distribution for  $FAP_0 = 0.05$  and for an isolated shift**

$\delta$	$m = 4$ and $n = 18$			$m = 10$ and $n = 21$			$m = 50$ and $n = 15$		
	$\bar{X}$	Mean-rank	Median	$\bar{X}$	Mean-rank	Median	$\bar{X}$	Mean-rank	Median
-3.0	0.9987	1.0000	1.0000	0.9996	1.0000	1.0000	0.9998	1.0000	0.9955
-2.5	0.9980	1.0000	1.0000	0.9993	1.0000	1.0000	0.9993	1.0000	0.9877
-2.0	0.9956	1.0000	1.0000	0.9983	1.0000	0.9999	0.9946	0.9999	0.9643
-1.5	0.9859	0.9999	0.9979	0.9932	1.0000	0.9981	0.9105	0.9985	0.8794
-1.0	0.8943	0.9825	0.9443	0.9165	0.9949	0.9488	0.3224	0.9290	0.6031
-0.5	0.3756	0.5197	0.4334	0.2935	0.5676	0.4121	0.0577	0.2697	0.1456
0.0	<b>0.0505</b>	<b>0.0470</b>	<b>0.0484</b>	<b>0.0494</b>	<b>0.0510</b>	<b>0.0474</b>	<b>0.0477</b>	<b>0.0492</b>	<b>0.0437</b>
0.5	0.3742	0.5126	0.4319	0.2944	0.5637	0.4126	0.0578	0.2674	0.1468
1.0	0.8949	0.9814	0.9455	0.9177	0.9952	0.9483	0.3182	0.9277	0.6047
1.5	0.9861	0.9999	0.9984	0.9928	1.0000	0.9984	0.9102	0.9984	0.8811
2.0	0.9955	1.0000	1.000	0.9980	1.0000	0.9999	0.9950	1.0000	0.9638
2.5	0.9981	1.0000	1.0000	0.9990	1.0000	1.0000	0.9993	1.0000	0.9883
3.0	0.9989	1.0000	1.0000	0.9996	1.0000	1.0000	0.9997	1.0000	0.9954

For the  $\text{Gamma}(2,1)$  distribution, the control limits for the  $\bar{X}$ -chart were again obtained using Table 2. The three charts perform almost equally well for positive shifts when  $m$  is small. However, when  $m = 50$ , the  $\bar{X}$ -chart has slightly lower alarm probabilities than the other two charts, which seem to perform almost identically. For negative shifts, however, the mean-rank chart performs the best, with the median chart showing rather low alarm probabilities. The  $\bar{X}$ -chart performs almost as well as the mean-rank chart. The three charts again perform better when  $m$  is small, especially in the case of the median chart, where the performance drops drastically when  $m = 50$ . The mean-rank chart shows the overall best performance for an isolated shift, whereas both the median chart and the  $\bar{X}$ -chart show some dissatisfactory results. Since the  $\text{Gamma}(2,1)$  distribution is right-skewed and the interpretations might not be straightforward using Table 5, the attained alarm probabilities are illustrated in Figure 2. Recall that for the two symmetric distributions ( $N(0,1)$  and  $t(3)$ ), the direction of the shift does not make a difference regarding the alarm probability and, accordingly, interpreting Tables 3 and 4 is more straightforward and thus no figures are provided.

**Table 5: Attained alarm probability values for the  $\text{Gamma}(2, 1)$  distribution for  $FAP_0 = 0.05$  and for an isolated shift**

$\delta$	$m = 4$ and $n = 18$			$m = 10$ and $n = 21$			$m = 50$ and $n = 15$		
	$\bar{X}$	Mean-rank	Median	$\bar{X}$	Mean-rank	Median	$\bar{X}$	Mean-rank	Median
-3.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.000	1.0000	1.0000	0.9686
-2.5	1.0000	1.0000	0.9988	1.0000	1.0000	0.9991	1.0000	1.0000	0.9116
-2.0	1.0000	1.000	0.9905	1.0000	1.0000	0.9902	0.9999	0.9987	0.7716
-1.5	0.9952	0.9964	0.9228	0.9992	0.9992	0.9142	0.9778	0.9749	0.5212
-1.0	0.8766	0.9151	0.6396	0.9242	0.9520	0.6038	0.6398	0.7755	0.2264
-0.5	0.3180	0.4064	0.2054	0.2976	0.4376	0.1688	0.0879	0.2198	0.0676
0.0	<b>0.0500</b>	<b>0.0496</b>	<b>0.0485</b>	<b>0.0492</b>	<b>0.0502</b>	<b>0.0458</b>	<b>0.0502</b>	<b>0.0505</b>	<b>0.0434</b>
0.5	0.3003	0.3968	0.2740	0.2678	0.4135	0.2784	0.0993	0.1382	0.1066
1.0	0.8786	0.9677	0.8938	0.9337	0.9964	0.9834	0.5973	0.9094	0.8991
1.5	0.9980	1.0000	0.9994	1.0000	1.0000	1.0000	0.9918	1.0000	1.0000
2.0	1.0000	1.0000	0.9999	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2.5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000



**Figure 2: Attained alarm probability values for the  $\text{Gamma}(2, 1)$  distribution for  $FAP_0 = 0.05$  and for an isolated shift (see online for colour version)**

### 3.2.2.2 Sustained shifts

The OOC performance for the sustained shift case was studied for all combinations of  $m$  and  $n$  in Table 1. Again, as in the case of the isolated shifts, we report the representative results for the combinations:  $(m, n) = (4, 18)$ ,  $(10, 24)$ , and  $(50, 15)$ , for the three distributions under consideration, in Tables 6 to 8. It is seen that for the two symmetric distributions  $N(0, 1)$  and  $t(3)$ , the direction of the shift does not make a difference regarding the alarm probability.

From Table 6, it can be seen that when the underlying process distribution is  $N(0, 1)$ , the  $\bar{X}$ -chart displays the highest alarm probabilities, with the mean-rank chart performing similarly. On the other hand, the median chart detects the shift with a slightly lower probability than the other two charts. For all three charts, it was observed that the probabilities of detecting a shift were higher when  $m = 50$ .

**Table 6: Attained alarm probability values for the standard normal distribution for  $FAP_0 = 0.05$  and for a sustained shift**

$\delta$	$m = 4$ and $n = 18$			$m = 10$ and $n = 21$			$m = 50$ and $n = 15$		
	$\bar{X}$	Mean-rank	Median	$\bar{X}$	Mean-rank	Median	$\bar{X}$	Mean-rank	Median
-3.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
-2.5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
-2.0	1.0000	1.0000	0.9987	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
-1.5	0.9985	0.9975	0.9603	1.0000	1.0000	0.9988	1.0000	1.0000	0.9994
-1.0	0.8822	0.8587	0.6668	0.9869	0.9781	0.8488	0.9953	0.9734	0.8162
-0.5	0.2957	0.2752	0.1913	0.4156	0.3899	0.2457	0.4095	0.3517	0.2103
0.0	0.0492	0.0481	0.0469	0.0482	0.0490	0.0455	0.0492	0.0486	0.0442
0.5	0.2965	0.2775	0.1916	0.4150	0.3840	0.2449	0.4050	0.3480	0.2120
1.0	0.8790	0.8558	0.6679	0.9870	0.9766	0.8491	0.9955	0.9746	0.8156
1.5	0.9982	0.9972	0.9618	1.0000	1.0000	0.9988	1.0000	1.0000	0.9994
2.0	1.0000	1.0000	0.9990	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2.5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

In the case of the  $t(3)$  distribution, the control limits for the  $\bar{X}$ -chart were again adjusted using Table 2. As shown in Table 7, all three charts perform similarly for  $m = 4$  to  $m = 10$ , with the mean-rank showing the highest alarm probabilities and the  $\bar{X}$ -chart having the lowest alarm probabilities; but as  $m$  increases, the performance of the  $\bar{X}$ -chart becomes much weaker. The mean-rank and median charts show the highest alarm probabilities for  $m = 50$ , although the mean-rank chart still outperforms the median chart.

**Table 7: Attained alarm probability values for the  $t(3)$  distribution for  $FAP_0 = 0.05$  and for a sustained shift**

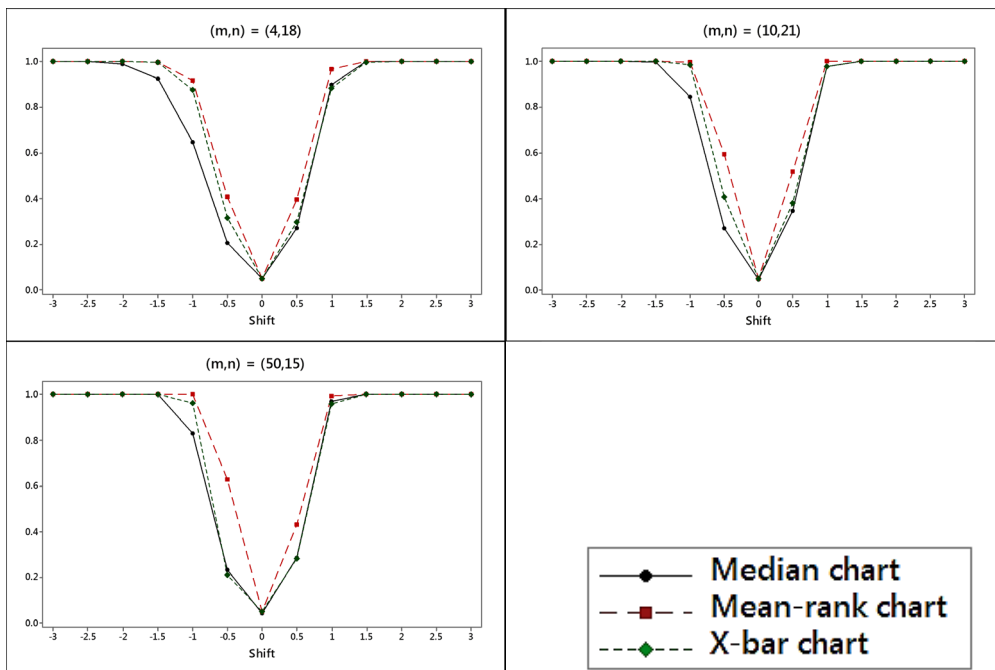
$\delta$	$m = 4$ and $n = 18$			$m = 10$ and $n = 21$			$m = 50$ and $n = 15$		
	$\bar{X}$	Mean-rank	Median	$\bar{X}$	Mean-rank	Median	$\bar{X}$	Mean-rank	Median
-3.0	0.9990	1.0000	1.0000	0.9996	1.0000	1.0000	0.9995	1.0000	1.0000
-2.5	0.9981	1.0000	1.0000	0.9991	1.0000	1.0000	0.9986	1.0000	1.0000
-2.0	0.9957	1.0000	0.9999	0.9983	1.0000	1.0000	0.9960	1.0000	1.0000
-1.5	0.9858	0.9999	0.9987	0.9949	1.0000	1.0000	0.9721	1.0000	1.0000
-1.0	0.8974	0.9815	0.9437	0.9591	0.9999	0.9981	0.5224	1.0000	0.9990
-0.5	0.3746	0.5184	0.4293	0.3949	0.7002	0.5869	0.0993	0.6488	0.5296
0.0	0.0506	0.0482	0.0478	0.0505	0.0528	0.0458	0.0466	0.0501	0.0421
0.5	0.3737	0.5146	0.4331	0.3942	0.7014	0.5888	0.0996	0.6473	0.5297
1.0	0.8951	0.9822	0.9443	0.9578	0.9999	0.9981	0.5241	1.0000	0.9994
1.5	0.9856	0.9998	0.9980	0.9958	1.0000	1.0000	0.9725	1.0000	1.0000
2.0	0.9957	1.0000	0.9999	0.9987	1.0000	1.0000	0.9965	1.0000	1.0000
2.5	0.9983	1.0000	1.0000	0.9993	1.0000	1.0000	0.9983	1.0000	1.0000
3.0	0.9988	1.0000	1.0000	0.9995	1.0000	1.0000	0.9994	1.0000	1.0000

As can be seen from Table 8 and Figure 3 for the  $Gamma(2,1)$  distribution, for which the control limits for the  $\bar{X}$ -chart were again adjusted using Table 2, the mean-rank chart shows the highest alarm probabilities for a negative as well as a positive shift. The  $\bar{X}$ -chart and the median chart perform almost identically for positive shifts, but for negative shifts the median chart shows lower probabilities than the  $\bar{X}$ -chart. Since the  $Gamma(2,1)$  distribution is right-skewed and the interpretations might not be straightforward using Table 8, the attained alarm probabilities are illustrated in Figure 3. Recall that for the two symmetric distributions ( $N(0,1)$  and  $t(3)$ ), the direction of the shift does not make a difference regarding the alarm probability and, accordingly, interpreting Tables 6 and 7 are more straightforward and thus no figures are provided.

Over all the combinations of  $m$  and  $n$  for the sustained shift, the mean-rank chart shows the highest alarm probabilities, indicating that it will detect a shift faster, with the median chart and  $\bar{X}$ -chart showing less satisfactory results in some instances.

**Table 8: Attained alarm probability values for the  $\text{Gamma}(2, 1)$  distribution for  $FAP_0 = 0.05$  and for a sustained shift**

$\delta$	$m = 4$ and $n = 18$			$m = 10$ and $n = 21$			$m = 50$ and $n = 15$		
	$\bar{X}$	Mean-rank	Median	$\bar{X}$	Mean-rank	Median	$\bar{X}$	Mean-rank	Median
-3.0	1.0000	1.0000	0.9999	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
-2.5	1.0000	1.0000	0.9993	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
-2.0	0.9999	0.9999	0.9903	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
-1.5	0.9947	0.9960	0.9251	1.0000	1.0000	0.9963	1.0000	1.0000	0.9985
-1.0	0.8740	0.9154	0.6455	0.9843	0.9962	0.8453	0.9632	0.9993	0.8302
-0.5	0.3162	0.4063	0.2062	0.4085	0.5942	0.2691	0.2104	0.6258	0.2336
0.0	<b>0.0499</b>	<b>0.0496</b>	<b>0.0490</b>	<b>0.0502</b>	<b>0.0496</b>	<b>0.0472</b>	<b>0.0501</b>	<b>0.0480</b>	<b>0.0433</b>
0.5	0.2958	0.3974	0.2712	0.3803	0.5171	0.3445	0.2809	0.4293	0.2833
1.0	0.8819	0.9662	0.8965	0.9772	0.9983	0.9789	0.9591	0.9912	0.9708
1.5	0.9981	1.0000	0.9992	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2.5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000



**Figure 3: Attained alarm probability values for the  $\text{Gamma}(2, 1)$  distribution for  $FAP_0 = 0.05$  and for a sustained shift (see online for colour version)**

#### 4 REMARKS AND CONCLUSIONS

In this paper we compared the performances of two available nonparametric Phase I charts: the mean-rank and the median chart. Careful consideration was given to the discreteness of the corresponding charting statistics, which can cause the attained  $FAP$  values sometimes to be very conservative and much smaller than the nominal  $FAP$  value for some number of subgroup-sample size  $(m, n)$  combinations. We included the normal theory Phase I  $\bar{X}$ -chart for comparison purposes. It was found that in terms of both IC- and OC performance, the mean-rank and the median charts perform similarly and satisfactorily to the  $\bar{X}$ -chart when the underlying distribution is normal. However, for heavy-tailed or skewed distributions, the two nonparametric charts both outperform the  $\bar{X}$ -chart, and there are serious concerns about the IC performance (highly inflated false alarm probabilities) of the  $\bar{X}$ -chart in these cases. In summary, the considerably better IC performance of the median chart compared with that of the  $\bar{X}$ -chart for non-normal data outweighs its slightly worse performance

compared with the  $\bar{X}$ -chart in some OOC cases, as the IC performance is deemed to be more crucial in Phase I applications. The mean-rank chart showed the best results overall, and would be the recommended chart to use in practice where the need for a nonparametric chart arises.

Our results make a strong case for using nonparametric Phase I charts in practice, particularly with the computing resources available today. We believe that Phase I control chart research has not received enough attention, since the vast majority of the SPC literature concerns Phase II control charting techniques. Thus further work on Phase I nonparametric control charts would be welcome. More research needs to be done regarding Phase I analysis for small subgroup sizes and even for individual observations because, although traditional SPC applications of control charts involve sub-grouped data, recent advances have led to more and more instances where individual measurements are collected over time.

## ACKNOWLEDGEMENTS

This research constitutes part of the first author's Master's dissertation at the Department of Statistics, University of Pretoria. The authors acknowledge support from a South African Research Chair Initiative (SARChI) award. Research of the third author was supported in part by a National Research Foundation grant (Thuthuka programme: TTK20100724000013247, Grant number: 76219).

## REFERENCES

- [1] **Montgomery, D.C.** 2013. *Statistical quality control: A modern introduction*. 7th edition, John Wiley & Sons.
- [2] **Chakraborti, S., Human, S.W. & Graham, M.A.** 2009. Phase I statistical process control charts: An overview and some results. *Quality Engineering*, 21(1), pp. 52-62.
- [3] **Jones-Farmer, L.A., Woodall, W.H., Steiner, S.H. & Champ, C.W.** 2014. An overview of Phase I analysis for process improvement and monitoring. *Journal of Quality Technology*, 46(3), pp. 265-280.
- [4] **Woodall, W.H.** 2000. Controversies and contradictions in statistical process control. *Journal of Quality Technology*, 32(4), pp. 341-350.
- [5] **Chakraborti, S., Human, S.W. & Graham, M.A.** 2011. Nonparametric (distribution-free) quality control charts. In: Balakrishnan, N. (ed.), *Handbook of methods and applications of statistics: Engineering, quality control, and physical sciences*. New York: John Wiley & Sons, pp. 298-329.
- [6] **Jones-Farmer, L.A. and Champ, C.W.** 2010. A distribution-free Phase I control chart for subgroup scale. *Journal of Quality Technology*, 42(4), pp. 373-387.
- [7] **Jones-Farmer, L.A., Jordan, V. & Champ, C.W.** 2009. Distribution-free phase i control charts for subgroup location. *Journal of Quality Technology*, 41(3), pp. 304-316.
- [8] **Graham, M.A., Human, S.W. & Chakraborti, S.** 2010. A Phase I nonparametric Shewhart-type control chart based on the median. *Journal of Applied Statistics*, 37(11), pp. 1795-1813.
- [9] **Champ, C.W. & Jones, L.A.** 2004. Designing Phase I  $\bar{X}$  charts with small sample sizes. *Quality and Reliability Engineering International*, 20(5), pp. 497-510.
- [10] **Gibbons, J.D. & Chakraborti, S.** 2010. *Nonparametric statistical inference*. 5th edition, Boca Raton, FL: Taylor and Francis.