We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Simulation-Based Comparative Analysis of Nonparametric Control Charts with Runs-Type Rules

Ioannis S. Triantafyllou

Abstract

In this chapter, we study well-known distribution-free Shewhart-type monitoring schemes based on order statistics. In order to empower the in- and out-of-control performance of the control charts being under consideration, several runs-type rules are enhanced. The simulation-based experimentation carried out reveals that the proposed schemes achieve remarkable efficiency for detecting possible shifts in the distribution of the underlying process.

Keywords: distribution-free monitoring schemes, average run length, false alarm rate, Lehmann alternatives, statistical process control, order statistics

1. Introduction

Statistical process control is widely applied to monitor the quality of a production process, where no matter of how thoroughly it is maintained, a natural variability always exists. Control charts help the practitioners to identify assignable causes so that the state of statistical control can be accomplished. Generally speaking, when a cast-off shift in the process takes place, a control chart should detect it as quickly as possible and produce an out-of-control signal.

Shewhart-type control charts were introduced in the early work of Shewhart [1], and since then, several modifications have been established and studied in detail. For a thorough study on statistical process control, the interested reader is referred to the classical textbooks of Montgomery [2] or Qiu [3]. Most of the monitoring schemes are distribution-based procedures, even though this assumption is not always realized in practice. To overcome this obstruction without disrupting the primary formation of the traditional control charts, several nonparametric (or distribution-free) monitoring schemes have been proposed in the literature. The plotted statistics being utilized for constructing such type of control charts are related to well-known nonparametric testing procedures. Among others, a variety of distribution-free control charts appeared already in the literature are based on order statistics; see, e.g., Chakraborti et al. [4], Balakrishnan et al. [5], or Triantafyllou [6, 7]. For an up-to-date account on nonparametric statistical process control, the reader is referred to the review chapter of Koutras and Triantafyllou [8], the recent monograph of Chakraborti and Graham [9], or Qiu [10, 11].

In the present chapter, we study well-known distribution-free *Shewhart*-type monitoring schemes based on order statistics. The general setup of the control charts being in consideration is presented in Section 2, while their performance characteristics are investigated based on the algorithm described in Section 3. In order to enhance the ability of the proposed monitoring schemes for detecting possible shifts of the process distribution, some well-known runs-type rules are considered. In Section 4, we carry out extensive simulation-based numerical comparisons that reveal that the underlying control charts outperform the existing ones under several out-of-control scenarios.

2. The setup of well-known nonparametric control charts based on order statistics

Let us assume that a reference random sample of size m, say $X_1, X_2, \ldots, X_m$, is drawn independently from an unknown continuous distribution F, namely, when the process is in-control. The control limits of the distribution-free monitoring scheme are determined by exploiting specific order statistics of the reference sample. In the sequel, test samples are drawn independently of each other (and also of the reference sample) from a continuous distribution G, and the decision whether the process is still in-control or not rests on suitably chosen test statistics. The framework for constructing nonparametric control charts based on order statistics calls for the following step-by-step procedure.

Step 1. Draw a reference sample of size m, namely, $X_1, X_2, \ldots, X_m$, from the process when it is known to be in-control.

Step 2. Form an interval by choosing appropriately a pair of order statistics from the reference sample (say, e.g., (X_a, X_b), where $1 \leq a < b \leq m$).

Step 3. Draw independently future (test) samples of size n, namely, $Y_1, Y_2, \ldots, Y_n$, from the underlying process.

Step 4. Pick out l order statistics $(0 < l \leq n)$ from each test sample.

Step 5. Determine the number of observations of each test sample, say R that lie between the limits of the interval (X_a, X_b).

Step 6. Configure the signaling rule by utilizing both the statistics R and the l ordered test sample observations as monitoring statistics.

The implementation of the above mechanism does not require the assumption of any specific probability distribution for the underlying process (measurements). The reference sample (usually of large size) is drawn from the underlying in-control process, while test (Phase II) samples are picked out from the future process in order to decide whether the process remains in-control or it has shifted to an out-of-control state. The proposed monitoring scheme is likely to possess the robustness feature of standard nonparametric procedures and is, consequently, less likely to be affected by outliers or the presence of skewed or heavy-tailed distributions for the underlying populations.

It is straightforward that the proposed framework requires the construction of more than one control charts, which monitor simultaneously the underlying process. In fact, the design parameter l is connected to the number of the control charts which are needed to be built for trading on the aforementioned mechanism. Indeed, for each one of the l order statistics from the test sample, a separate two-sided control chart should be constructed.

The family of distribution-free monitoring schemes presented earlier includes as special cases some nonparametric control charts, which have been already established in the literature. For example, the monitoring scheme established by Balakrishnan et al. [5] calls for the following plotted statistics:

- A quantile $Y_{j:n}$ of the test sample which is compared with the control limits (X_a, X_b)

- The number of observations from the test sample that lie between the control limits

It goes without saying that the control chart introduced by Balakrishnan et al. [5] (*Chart* 1, hereafter) belongs to the family of monitoring schemes described previously. In fact, the *BTK* chart could be seen as a special case of the aforementioned class of distribution-free control schemes with $l = 1$. According to *Chart* 1, the process is declared to be in-control, if the following conditions hold true

$$X_{a:m} \leq Y_{j:n} \leq X_{b:m} \quad \text{and} \quad R \geq r, \tag{1}$$

where r is a positive integer.

In addition, the monitoring scheme introduced by Triantafyllou [6] (*Chart* 2, hereafter) takes into account the location of two order statistics of the test sample drawn from the process along with the number of its observations between the control limits. In other words, the aforementioned control chart could be viewed as a member of the general class of nonparametric monitoring schemes with $l = 2$. According to *Chart* 2, the process is declared to be in-control, if the following conditions hold true

$$X_{a:m} \leq Y_{j:n} \leq Y_{k:n} \leq X_{b:m} \quad \text{and} \quad R \geq r, \tag{2}$$

where r is a positive integer.

In a slightly different framework, Triantafyllou [7] proposed a distribution-free control chart based on order statistics (*Chart* 3, hereafter) by taking advantage of the position of single ordered observations from both test and reference sample. More precisely, *Chart* 3 asks for an order statistic of each test sample (say $Y_{j:n}$) to be enveloped by two prespecified observations $X_{a:m}$ and $X_{b:m}$ of the reference sample, while at the same time an ordered observation of the reference sample (say $X_{i:m}$) be enclosed by two predetermined values of the test sample ($Y_{c:n}, Y_{d:n}$). *Chart* 3 makes use of an in-control rule, which embraces the following three conditions:

Condition 1. The statistic $Y_{j:n}$ of the test sample should lie between the observations $X_{a:m}$ and $X_{b:m}$ of the reference sample, namely, $X_{a:m} \leq Y_{j:n} \leq X_{b:m}$.

Condition 2. The interval ($Y_{c:n}, Y_{d:n}$) formulated by two appropriately chosen order statistics of the test sample should enclose the value $X_{i:m}$ of the reference sample, namely, $Y_{c:n} \leq X_{i:m} \leq Y_{d:n}$.

Condition 3. The number of observations of the Y-sample that are placed enclosed by the observations $X_{a:m}$ and $X_{b:m}$ should be equal to or more than r, namely, $R \geq r$.

3. The simulation procedure and some results

In the present section, we describe the step-by-step procedure which has been followed in order to determine the basic performance characteristics of monitoring schemes mentioned previously. Two well-known runs-type rules are implemented in order to improve the performance of the control charts being considered. More precisely, if we denote by *LCL* and *UCL* the lower and the upper control limit of the underlying monitoring scheme, we apply the following runs rules

- The 2-of-2 rule. Under this scenario, an out-of-control signal is produced from the control chart, whenever two consecutive plotted points fall all of them

either on or above the *UCL* or all of them fall on or below the *LCL* (see, e.g., Klein [12]).

- The 2-of-3 rule. Under this scenario, an out-of-control signal is produced from the control chart, whenever two out of three consecutive plotted points fall outside the control limits (*LCL*, *UCL*) of the corresponding scheme.

We next illustrate the detailed procedure for determining the performance of *Chart* 2 enhanced with the 2-of-3 rule. It goes without saying that a similar algorithm has been constructed in order to study the corresponding characteristics of the remaining control schemes, namely, *Chart* 1 and *Chart* 3 enhanced with either the 2-of-2 or the 2-of-3 runs rule.

Step 1. Generate a reference sample of size m from the in-control distribution F and k_2 test samples of size n from the out-of-control distribution G.

Step 2. Determine the control limits of the monitoring scheme *Chart* 2, by selecting appropriately the parameters a, b, r.

Step 3. Calculate the test statistics Y_j, Y_k, R for each test sample, and examine whether *Chart* 2 produces an out-of-control signal or not, namely, whether at least one of the conditions mentioned in (2) is violated.

Step 4. Define a dummy variable T_i, $i = 1, 2, \ldots, k_2$ for each test sample separately. The variable T_i takes on the value 0 when all conditions in (2) are satisfied, while it takes on the value 1 otherwise.

Step 5. Determine all consecutive (uninterrupted) triplets consisting of $T_i's$ elements, namely, all triplets $(T_i, T_{i+1}, T_{i+2}), i = 1, \ldots, k_2 - 2$. Define the dummy variable D_j, $j = 1, 2, \ldots, k_2 - 2$ for each triplet separately. The variable D_j takes on the value 0 when the triplet consists of at least two 0s, while it takes on the value 1 otherwise.

Step 6. Calculate the alarm rate of the monitoring scheme as $AR = \sum_{j=1}^{k_2-2} D_j /$ $(k_2 - 2)$. When $F = G$, the aforementioned probability indicates the false alarm rate of the monitoring scheme, while in case of different distributions F, G the AR corresponds to its out-of-control alarm rate.

Step 7. Define a variable $RL_h, h = 1, 2, \ldots, H$ which counts the number of $D_j's$ elements, till the first appearance of a D_j equal to 1. The so-called average run length of the monitoring scheme is calculated as $ARL = \sum_{h=1}^{H} RL_h / H$. When $F = G$ ($F \neq G$), the aforementioned quantity indicates the in-control (out-of-control) average run length of the monitoring scheme.

All steps 1–7 are repeated k_1 times and the performance characteristics of the proposed *Chart* 2 enhanced with 2-of-3 runs rule, namely, the false alarm rate (*FAR*, hereafter), the out-of-control alarm rate (*AR_{out}*, hereafter), the in-control average run length (*ARL_{in}*, hereafter), and the out-of-control run length (*ARL_{out}*, hereafter) are estimated as the mean values of the corresponding k_1 results produced by steps 6 and 7, respectively.

In order to ascertain the validity of the proposed simulation procedure described above, we shall first apply the algorithm without embodying any runs-type rule and compare the simulation-based outcomes to the corresponding results produced by the aid of the theoretical approximation appeared in Triantafyllou [6]. The simulation study has been accomplished based on the R software environment and involves 10.000 replications. **Table 1** displays several designs of the monitoring scheme mentioned as *Chart* 2 with a nominal level of in-control performance. Since we consider the same designs as those presented by Triantafyllou [6], the exact *FAR*s have been taken from his **Table 1**. As it is easily observed, the simulation-based results seem to be quite close to the exact values in all cases considered. For example, let us assume that we draw a reference sample of size m = 60 and test

samples of size n = 5. In order to achieve a prespecified in-control performance level, namely, *FAR* equal to 1%, the remaining parameters are determined as a = 1, j = 2, k = 4, and r = 1. Under the aforementioned design, Triantafyllou [6] computed the exact *FAR* equal to 0.0096, while the simulation-based procedure proposed in the present chapter gives a corresponding *FAR* value equal to 0.0116.

A different approach for appraising the ability of a monitoring scheme to detect a possible shift in the underlying distribution is based on its run length. We next focus on the waiting time random variable N, which corresponds to the amount of random test samples up to getting the first out-of-control signal from the monitoring scheme, in order to evaluate its performance. **Table 2** displays the exact and the simulation-based average run length for several designs of *Chart* 2 that meet a desired nominal level of in-control performance. The exact values of *ARL* needed for building up **Table 2**, have been picked up from Triantafyllou [6] and more specifically from **Table 2** therein.

As it is readily observed, the simulation-based results seem to be quite close to the exact values in all cases considered. For example, let us assume that we draw a reference sample of size m = 400 and test samples of size n = 5. In order to achieve a prespecified in-control performance level, namely, ARL_{in} equal to 370, the remaining parameters are determined as a = 5, b = 379, j = 2, k = 3, and r = 2. Under

		Reference	sample			
		40			**60**	
n	(a, j, k, r)	*Exact FAR*	*Simulated FAR*	(a, j, k, r)	*Exact FAR*	*Simulated FAR*
5	(1, 2, 3, 3)	0.0114	0.0119	(1, 2, 4, 1)	0.0096	0.0116
	(3, 2, 3, 2)	0.0615	0.0632	(4, 3, 4, 2)	0.0478	0.0507
	(4, 3, 4, 2)	0.0999	0.0897	(6, 3, 4, 2)	0.0969	0.1044
11	(2, 4, 8, 4)	0.0116	0.0158	(4, 4, 7, 5)	0.0108	0.0098
	(1, 2, 4, 5)	0.0431	0.0551	(7, 4, 7, 4)	0.0518	0.0573
	(1, 6, 10, 5)	0.0432	0.0437	(6, 3, 6, 5)	0.1030	0.1084
25	(4, 8, 12, 4)	0.0139	0.0150	(9, 10, 14, 5)	0.0097	0.0100
	(5, 14, 17, 5)	0.0137	0.0126	(17, 12, 14, 5)	0.1031	0.1131
	(10, 11, 14, 4)	0.0928	0.1049	(15, 11, 15, 4)	0.1037	0.1130
		Reference	sample			
		100			**200**	
n	(a, j, k, r)	*Exact FAR*	*Simulated FAR*	(a, j, k, r)	*Exact FAR*	*Simulated FAR*
5	(3, 3, 4, 3)	0.0119	0.0105	(7, 2, 3, 2)	0.0132	0.0159
	(5, 2, 4, 3)	0.0505	0.0566	(10, 2, 4, 2)	0.0479	0.0473
	(1, 1, 5, 3)	0.0934	0.1076	(21, 3, 4, 2)	0.1008	0.0937
11	(4, 3, 8, 5)	0.0146	0.0107	(8, 3, 6, 4)	0.0105	0.0103
	(7, 3, 7, 6)	0.0460	0.0506	(16, 4, 6, 3)	0.0106	0.0097
	(7, 3, 8, 5)	0.0524	0.0603	(26, 4, 6, 3)	0.0498	0.0492
25	(19, 11, 14, 4)	0.0113	0.0122	(39, 11, 14, 4)	0.0093	0.0090
	(24, 11, 14, 4)	0.0490	0.0463	(45, 13, 16, 6)	0.0479	0.0445
	(19, 14, 17, 5)	0.0504	0.0518	(45, 14, 17, 4)	0.0995	0.0992

Table 1.
Exact and simulation-based FAR *for a given design of* Chart 2.

ARL_0	m	n	(a, b)	(j, k, r)	Exact ARL_{in}	Simulated ARL_{in}
370	200	5	(2, 188)	(2, 3, 2)	376.4	369.9
		11	(3, 180)	(2, 3, 3)	367.7	377.7
		25	(14, 178)	(6, 9, 14)	369.9	369.2
	300	5	(5, 287)	(2, 3, 2)	358.9	365.0
		11	(19, 285)	(4, 7, 4)	372.6	374.8
		25	(20, 213)	(6, 9, 7)	367.2	373.8
	400	5	(5, 379)	(2, 3, 2)	378.7	379.1
		11	(12, 366)	(3, 6, 4)	384.8	378.4
		25	(33, 296)	(7, 10, 9)	373.0	377.9
	500	5	(6, 473)	(2, 3, 2)	369.8	372.6
		11	(15, 440)	(3, 5, 4)	369.1	374.8
		25	(32, 362)	(6, 9, 6)	369.4	365.1

Table 2.
Exact and simulation-based ARL_{in} *for a given design of* Chart 2.

the aforementioned design, Triantafyllou [6] computed the exact ARL_{in} equal to 378.7, while the simulation-based procedure proposed in the present chapter gives a corresponding ARL_{in} value equal to 379.1.

We next focus on the ability of the distribution-free monitoring scheme defined in (1), under the assumption that the process has shifted to an out-of-control state. When the process has shifted from distribution F to G, then the ability of the scheme to detect the underlying alteration is associated with the function $G \circ F^{-1}$. For example, under the well-known Lehmann-type alternative (see, e.g., van der Laan and Chakraborti [13]), the out-of-control distribution function can be expressed as $G = F^{\gamma}$, where $\gamma > 0$. **Table 3** sheds light on the out-of-control performance of *Chart* 2 by offering the corresponding alarm rate of the proposed scheme under the Lehmann alternatives with parameter $\gamma = 0.2, 10$. Since we consider the same designs as those presented by Triantafyllou [6], the exact values of AR_{out} have been copied from his **Table 3**, while the simulated results have been produced by following the procedure described earlier.

		Reference		sample		
		40			**60**	
n	(a, j, k, r)	Exact AR_{out}	Simulated AR_{out}	(a, j, k, r)	Exact AR_{out}	Simulated AR_{out}
5	(2, 2, 4, 3)	0.8284	0.8209	(2, 2, 4, 4)	0.7890	0.7783
		0.5448	0.5244		0.3931	0.3589
	(3, 2, 4, 2)	0.8853	0.8758	(4, 2, 4, 3)	0.8800	0.8805
		0.7351	0.7403		0.7161	0.7095
11	(2, 4, 8, 4)	0.8883	0.8825	(7, 4, 7, 4)	0.9812	0.9737
		0.5357	0.5502		0.9076	0.9090
	(3, 5, 8, 4)	0.8483	0.8424	(6, 5, 9, 4)	0.9135	0.9278
		0.7556	0.7612		0.9678	0.9614
25	(9, 11, 14, 4)	0.9983	0.9978	(12, 12, 14, 4)	0.9938	0.9922
		0.9978	0.9978		0.9982	0.9970
	(10, 11, 14, 4)	0.9991	0.9991	(9, 10, 14, 5)	0.9971	0.9929
		0.9993	0.9994		0.9759	0.9763

		Reference		sample		
		100			**200**	
n	*(a, j, k, r)*	*Exact AR$_{out}$*	*Simulated AR$_{out}$*	*(a, j, k, r)*	*Exact AR$_{out}$*	*Simulated AR$_{out}$*
5	(5, 2, 4, 3)	0.8530	0.8513	(10, 2, 4, 2)	0.8609	0.8661
		0.6062	0.5815		0.6331	0.6414
	(3, 3, 4, 3)	0.4783	0.4854	(15, 2, 4, 2)	0.9056	0.8908
		0.3333	0.3867		0.8300	0.8237
11	(7, 3, 8, 5)	0.9888	0.9693	(25, 5, 8, 4)	0.9527	0.9573
		0.8133	0.8059		0.9936	0.9912
	(19, 5, 7, 4)	0.9812	0.9768	(26, 4, 8, 4)	0.9895	0.9835
		0.9982	0.9831		0.9953	0.9913
25	(19, 11, 14, 4)	0.9982	0.9974	(39, 11, 14, 4)	0.9989	0.9963
		0.9994	0.9962		0.9999	0.9963
	(24, 11, 14, 4)	0.9996	0.9989	(45, 13, 16, 6)	0.9942	0.9963
		0.9999	0.9989		0.9999	0.9962

Each cell contains the AR's attained for Y = 0.2 (upper entry) and Y = 10 (lower entry).

Table 3.
Exact and simulation-based AR$_{out}$ *for a given design of* Chart *2.*

Each cell contains the *AR*s attained for $\gamma = 0.2$ (upper entry) and $\gamma = 10$ (lower entry).
Based on the above table, it is evident that the proposed simulation algorithm seems to come to an agreement with the corresponding exact values of the out-of-control alarm rate of *Chart* 2. For example, for a design $(a, j, k, r) = (10, 2, 4, 2)$ with reference sample size $m = 200$ and test sample size $n = 5$, the exact alarm rate for a shift to Lehmann alternative with parameter $\gamma = 0.2$ (10) equals to 86.09% (66.31%), while the simulation-based alarm rate of *Chart* 2 is quite close to the exact one, namely, it equals to 86.61% (64.14%).

4. The proposed control charts enhanced with runs-type rules

In this section, we carry out an extensive numerical experimentation to appraise the ability of the distribution-free monitoring schemes *Chart* 1, *Chart* 2, and *Chart* 3

							Chart 1				Chart 1 with 2-of-3 runs rule		
ARL$_0$	*m*	*n*	*(a, b)*	*j*	*r*	*Exact ARL$_{in}$*	ARL$_{out}$	*(a, b)*	*j*	*r*	*Exact ARL$_{in}$*	ARL$_{out}$
370	100	5	(2, 96)	2	2	363.7	15.77	(3, 71)	2	2	378.01	11.73
		11	(6, 83)	5	5	363.06	6.76	(8, 71)	4	5	364.79	6.59
	500	5	(7, 473)	3	3	382.01	10.88	(9, 352)	2	2	380.5	10.27
		11	(29, 444)	4	5	373.14	16.71	(17, 303)	4	4	360.9	9.29
500	100	5	(3, 96)	3	3	499.88	14.94	(2, 71)	2	2	497.66	12.69
		11	(6, 84)	5	5	485.85	7.79	(4, 70)	3	6	511.12	7.13
	500	5	(7, 476)	3	3	503.77	12.62	(8, 352)	2	2	488.61	11.56
		11	(28, 454)	4	5	499.79	27.63	(17, 309)	4	4	489.45	5.62

Underlying distributions: Exponential with mean equal to 2 (in-control) and 1 (out-of-control) respectively.

Table 4.
Comparison of the ARL$_{out}$*s with the same* ARL$_{in}$ *for* Chart *1.*

enhanced with runs-type rules for detecting possible shifts of the underlying distribution. The computations have been made by the aid of the simulation procedure presented in Section 3. **Tables 4** and **5** display the improved out-of-control performance of *Chart* 1, when the 2-of-3 runs-type rule is activated. We first compare the performance of the control charts by using a common ARL_{in} and then evaluating the respective ARL_{out} for specific shifts. Consider the case of a process with underlying in-control exponential distribution with mean equal to 2 and out-of-control exponential distribution with mean equal to 1. In **Table 4**, we present the ARL_{out} values of *Chart* 1 and the proposed *Chart* 1 enhanced with 2-of-3 runs-type rule, for $ARL_{in} = 370, 500$, $m = 100, 500$, and $n = 5, 11$. The remaining design parameters a, b, j, r, were determined appropriately, so that ARL_{in} takes on a value as close to the nominal level as possible. It is evident that the proposed monitoring scheme performs better than the one established by Balakrishnan et al. [5] for all cases considered. The fact that the ARL_{out}s that exhibit *Chart* 1 with 2-of-3 runs-type rule

		Chart 1	*Chart* 1 with 2-of-3 runs rule
θ	*δ*	ARL	ARL
0	1	499.88	517.21
0.25	1	562.33	127.89
0.5	1	386.09	26.63
1	1	61.22	3.83
1.5	1	8.01	1.44
2	1	2.23	1.06
0.25	1.25	49.69	35.07
0.5	1.25	39.96	13.72
1	1.25	14.83	3.85
1.5	1.25	4.76	1.66
2	1.25	2.08	1.14
0.25	1.5	13.91	17.77
0.5	1.5	12.32	9.43
1	1.5	7.07	3.65
1.5	1.5	3.53	1.85
2	1.5	1.98	1.26
0.25	1.75	6.55	10.24
0.5	1.75	6.11	7.64
1	1.75	4.40	3.49
1.5	1.75	2.83	1.92
2	1.75	1.88	1.35
0.25	2	4.04	6.93
0.5	2	3.87	5.49
1	2	3.17	3.09
1.5	2	2.37	2.03
2	2	1.77	1.45

Table 5.
Comparison of the ARL_{out}s *with the same* ARL_{in} *for* Chart *1 under normal distribution* (θ, δ).

				Chart 2				Chart 2 with 2-of-2 runs rule			
FAR	m	n	(a, b)	(j, k)	r	Exact FAR	AR$_{out}$	(a, b)	(j, k)	r	Exact FAR	AR$_{out}$
0.01	50	11	(3, 45)	(4, 6)	4	0.0103	0.4575 0.9239	(14, 35)	(4, 6)	1	0.0125	0.5527 0.9463
		25	(4, 45)	(7, 9)	7	0.0131	0.7756 0.9977	(11, 40)	(7, 9)	1	0.0110	0.9194 0.9989
	100	11	(4, 92)	(3, 5)	3	0.0126	0.5832 0.9661	(23, 75)	(3, 5)	1	0.0124	0.6771 0.9782
		25	(6, 65)	(6, 8)	4	0.0092	0.8181 0.9992	(20, 90)	(6, 8)	1	0.0135	0.9563 0.9992
	500	11	(19, 405)	(3, 5)	4	0.0099	0.5976 0.9736	(122, 350)	(3, 5)	1	0.0130	0.7278 0.9846
		25	(61, 420)	(8, 10)	5	0.0092	0.9023 0.9998	(130, 430)	(8, 10)	1	0.0089	0.9551 0.9998
	1000	11	(80, 894)	(4, 6)	5	0.0101	0.6106 0.9715	(300, 550)	(4, 6)	1	0.0110	0.6234 0.9715
		25	(125, 810)	(8, 10)	8	0.0098	0.9125 0.9998	(27, 800)	(8, 10)	1	0.0094	0.9644 0.9999
0.005	50	11	(2, 45)	(4, 6)	4	0.0054	0.3250 0.8627	(2, 45)	(4, 6)	4	0.0056	0.4229 0.9990
		25	(4, 44)	(8, 10)	8	0.0049	0.6577 0.9935	(4, 44)	(8, 10)	8	0.0057	0.8633 0.9979
	100	11	(2, 83)	(3, 5)	3	0.0046	0.7508 0.9983	(2, 83)	(3, 5)	3	0.0055	0.7620 0.9983
		25	(5, 88)	(6, 8)	4	0.0044	0.7508 0.9983	(5, 88)	(6, 8)	4	0.0054	0.9169 0.9989
	500	11	(15, 420)	(3, 5)	4	0.0050	0.5234 0.9618	(15, 420)	(3, 5)	4	0.0054	0.6611 0.9799
		25	(54, 380)	(8, 10)	5	0.0046	0.8624 0.9997	(54, 380)	(8, 10)	5	0.0056	0.9288 0.9997
	1000	11	(66, 900)	(4, 6)	4	0.0049	0.5357 0.9594	(66, 900)	(4, 6)	4	0.0047	0.5667 0.9596
		25	(87, 830)	(7, 9)	7	0.0050	0.8832 0.9998	(87, 830)	(7, 9)	7	0.0058	0.9646 0.9998
0.0027	50	11	(2, 49)	(4, 6)	4	0.0026	0.3249 0.8627	(2, 49)	(4, 6)	4	0.0026	0.4352 0.9081
		25	(6, 44)	(10, 11)	15	0.0036	0.6179 0.9908	(6, 44)	(10, 11)	15	0.0031	0.7553 0.9947
	100	11	(2, 94)	(3, 5)	3	0.0028	0.3722 0.9040	(2, 94)	(3, 5)	3	0.0025	0.5758 0.9670
		25	(4, 67)	(6, 8)	4	0.0025	0.6595 0.9960	(4, 67)	(6, 8)	4	0.0029	0.8722 0.9987
	500	11	(12, 440)	(3, 5)	4	0.0025	0.4559 0.9479	(12, 440)	(3, 5)	4	0.0031	0.6099 0.9724
		25	(49, 445)	(8, 10)	5	0.0026	0.8250 0.9995	(49, 445)	(8, 10)	5	0.0033	0.9318 0.9995
	1000	11	(55, 900)	(4, 6)	4	0.0027	0.4673 0.9452	(55, 900)	(4, 6)	4	0.0033	0.4673 0.9529
		25	(77, 800)	(7, 9)	7	0.0027	0.8435 0.9997	(77, 800)	(7, 9)	7	0.0032	0.9356 0.9997

Table 6.
Comparison of the AR$_{out}$*s with the same* FAR *for* Chart 2.

are smaller than the respective ones of *Chart* 1 indicates its efficacy to detect faster the shift of the process from the in-control distribution.

Underlying distributions: exponential with mean equal to 2 (in-control) and 1 (out-of-control), respectively.

It goes without saying that the nonparametric control charts are robust in the sense that their in-control behavior remains the same for all continuous distributions. However, it is of some interest to check over their out-of-control performance for different underlying distributions. We next study the performance of the proposed *Chart* 1 enhanced with the 2-of-3 runs-type rule under normal distribution (θ, δ).

More specifically, the in-control reference sample is drawn from the standard normal distribution, while several combinations of parameters θ, δ have been examined. **Table 5** reveals that the proposed *Chart* 1 with 2-of-3 runs-type rule is superior compared to the existing *Chart* 1 for almost all shifts of the location parameter θ and the scale parameter δ considered.

We next study the out-of-control performance of *Chart* 2 presented in Section 2. In **Table 6**, three different *FAR* levels and several values of the parameters m, n have been considered. For each choice, the *AR* values under two specific Lehmann alternatives corresponding to $\gamma = 0.4$ and $\gamma = 0.2$ are computed via simulation for both *Chart* 2 and *Chart* 2 enhanced with 2-of-2 runs-type rule.

Table 6 clearly indicates that, under a common *FAR*, the proposed *Chart* 2 with the 2-of-2 runs-type rule performs better than *Chart* 2, with respect to *AR* values, in all cases considered. For example, calling for a reference sample of size $m = 500$, test samples of size $n = 11$, and nominal $FAR = 0.0027$, the proposed *Chart* 2 with the 2-of-2 runs-type rule achieves alarm rate of 0.6099 (0.9724) for $\gamma = 0.4$ ($\gamma = 0.2$), while the respective alarm rate for *Chart* 2 is 0.4559 (0.9479).

Tables 7 and **8** shed more light on the out-of-control performance of the proposed *Chart* 2 enhanced with appropriate runs-type rules. More specifically, the schemes being under consideration are designed such as a nominal in-control *ARL* performance is attained. From the numerical comparisons carried out, it is straightforward that *Chart* 2 with 2-of-2 rule becomes substantially more efficient than *Chart* 2. Under Lehmann alternative with parameter $\gamma = 0.5$, the proposed chart exhibits smaller out-of-control *ARL* than *Chart* 2, and therefore it seems more capable in detecting possible shift of the process distribution.

				Chart 2				Chart 2 with 2-of-2 runs rule				
ARL_0	m	n	(a, b)	(j, k)	r	Exact ARL_{in}	ARL_{out}	(a, b)	(j, k)	r	Exact ARL_{in}	ARL_{out}
100	100	5	(4, 99)	(2, 3)	2	115.1	4.7	(10, 80)	(3, 5)	1	116.1	4.6
		11	(5, 78)	(3, 4)	3	88.9	2.8	(10, 83)	(2, 4)	1	83.2	2.7
	500	5	(15, 470)	(2, 3)	2	109.8	5.2	(50, 435)	(2, 3)	1	101.7	4.9
		11	(20, 410)	(4, 6)	4	119.7	6.8	(50, 400)	(2, 3)	1	122.5	2.2
200	100	5	(3, 98)	(2, 3)	2	185.3	6.7	(10, 85)	(2, 5)	1	204.8	1.5
		11	(4, 80)	(3, 4)	3	187.7	3.5	(8, 95)	(2, 3)	1	221	3.3
	500	5	(10, 482)	(2, 3)	2	208.9	7.2	(15, 435)	(3, 5)	1	188.4	5.4
		11	(18, 420)	(4, 6)	4	191.1	7.5	(50, 420)	(2, 3)	1	234.8	2.2

Table 7.
Comparison of the ARL_{out}*s with the same* ARL_{in} *for* Chart 2.

θ	δ	*Chart* 2	*Chart* 2 with 2-of-3 runs rule
0	1	446.6	502.1
0.25	1	163.9	127.9
0.5	1	51.64	26.6
1	1	7.4	3.8
1.5	1	2.1	1.4
2	1	1.2	1.1
0.25	1.25	35.7	35.1
0.5	1.25	17.9	13.7
1	1.25	5.0	3.9
1.5	1.25	2.1	1.7
2	1.25	1.3	1.1
0.25	1.5	15.0	17.7
0.5	1.5	9.8	9.4
1	1.5	4.1	3.6
1.5	1.5	2.1	1.8
2	1.5	1.4	1.3
0.25	1.75	8.5	10.6
0.5	1.75	6.5	7.6
1	1.75	3.5	3.4
1.5	1.75	2.1	1.9
2	1.75	1.5	1.3
0.25	2	5.7	6.9
0.5	2	4.8	5.4
1	2	3.1	3.1
1.5	2	2.0	2.0
2	2	1.5	1.4

Table 8.
Comparison of the $\text{ARL}_{out}s$ *with the same* ARL_{in} *for* Chart *2 under normal distribution* (θ, δ).

In addition, **Table 8** depicts the out-of-control *ARL* performance of *Chart* 2 under normal distribution. More precisely, several shifts of both location and scale parameter have been considered, and *Chart* 2 with 2-of-3 rule detects the underlying shift sooner than *Chart* 2 does in almost all cases examined.

Finally, **Tables 9** and **10** present simulation-based comparisons of the nonparametric monitoring scheme *Chart* 3 with 2-of-2 runs rule versus *Chart* 3 established by Triantafyllou [7]. For delivering the numerical results displayed in **Tables 9** and **10**, the Lehmann alternatives have been considered as the out-of-control distribution.

Each cell contains the *AR*s attained for $\gamma = 0.5$ (upper entry) and $\gamma = 0.2$ (lower entry).

5. Conclusions

In the present chapter, we investigate the in- and out-of-control performance of distribution-free control charts based on order statistics. Several runs-type rules are

				Chart 3				Chart 3 with 2-of-2 runs rule				
FAR	m	n	(a, b)	(i, c, j, d)	r	Exact FAR	AR_{out}	(a, b)	(i, c, j, d)	r	Exact FAR	AR_{out}
0.01	100	11	(6, 93)	(7, 2, 4, 7)	4	0.0104	0.5363 0.9537	(15, 75)	(7, 2, 4, 7)	1	0.0086	0.7064 0.9825
		25	(5, 86)	(6, 2, 6, 12)	10	0.0102	0.8181 0.9992	(10, 79)	(6, 2, 6, 12)	1	0.0108	0.9103 0.9992
	500	11	(31, 484)	(33, 2, 4, 7)	6	0.0108	0.5335 0.9583	(80, 300)	(33, 2, 4, 7)	1	0.0088	0.7326 0.9876
		25	(30, 450)	(49, 2, 7, 9)	8	0.0101	0.9122 0.9999	(75, 300)	(49, 2, 7, 9)	1	0.0089	0.9585 0.9999
	1000	11	(60, 967)	(65, 2, 4, 7)	6	0.0099	0.5311 0.9584	(155, 690)	(65, 2, 4, 7)	1	0.0096	0.7161 0.9873
		25	(30, 960)	(99, 2, 7, 9)	8	0.0098	0.9181 0.9999	(165, 600)	(99, 2, 7, 9)	1	0.0129	0.9778 0.9999
0.005	100	11	(4, 93)	(5, 2, 4, 7)	4	0.0051	0.4134 0.9198	(14, 75)	(5, 2, 4, 7)	1	0.0054	0.6312 0.9787
		25	(4, 88)	(5, 2, 6, 12)	10	0.0051	0.7508 0.9983	(10, 81)	(5, 2, 6, 12)	1	0.0051	0.9227 0.9989
	500	11	(26, 478)	(31, 2, 4, 7)	5	0.0052	0.5094 0.9536	(70, 300)	(31, 2, 4, 7)	1	0.0049	0.6653 0.9827
		25	(30, 450)	(43, 2, 7, 9)	8	0.0052	0.8755 0.9998	(70, 314)	(43, 2, 7, 9)	1	0.0046	0.9485 0.9998
	1000	11	(60, 985)	(65, 2, 4, 7)	6	0.0050	0.5299 0.9583	(150, 700)	(65, 2, 4, 7)	1	0.0047	0.7051 0.9870
		25	(30, 960)	(87, 2, 7, 9)	8	0.0050	0.8832 0.9998	(150, 600)	(87, 2, 7, 9)	1	0.0056	0.9645 0.9998
0.0027	100	11	(3, 94)	(4, 2, 4, 7)	4	0.0029	0.3410 0.8905	(11, 75)	(4, 2, 4, 7)	1	0.0032	0.5283 0.9660
		25	(5, 86)	(6, 2, 7, 9)	7	0.0025	0.6951 0.9971	(9, 80)	(6, 2, 6, 9)	1	0.0024	0.8803 0.9988
	500	11	(24, 480)	(28, 2, 4, 7)	4	0.0028	0.4715 0.9453	(70, 345)	(28, 2, 4, 7)	1	0.0024	0.6624 0.9834
		25	(30, 450)	(38, 2, 7, 9)	8	0.0027	0.8338 0.9996	(70, 320)	(38, 2, 7, 4)	1	0.0030	0.9497 0.9996
	1000	11	(53, 990)	(55, 2, 4, 7)	6	0.0027	0.4673 0.9452	(135, 705)	(65, 2, 4, 7)	1	0.0023	0.6445 0.9823
		25	(30, 960)	(78, 2, 7, 9)	8	0.0028	0.8479 0.9997	(145, 600)	(78, 2, 7, 9)	1	0.0029	0.9587 0.9997

Table 9.
Comparison of the $AR_{out}s$ *with the same* FAR *for* Chart 3.

employed in order to enhance the ability of the aforementioned nonparametric monitoring schemes to detect possible shifts in distribution process. The *AR* and the *ARL* behavior of the underlying control charts is studied under several out-of-control situations, such as the so-called Lehmann alternatives and the exponential or the normal distribution model. The numerical experimentation carried out depicts the melioration of the proposed schemes with the runs-type rules. It is of some research interest to branch out the incorporation of such runs rules (or even more complicated) to additional nonparametric control charts based on well-known test statistics.

								Chart 3 with 2-of-2 runs rule				
ARL_0	m	n	(a, b)	(i, c, j, d)	r	*Exact* ARL_{in}	ARL_{out}	(a, b)	(i, c, j, d)	r	*Exact* ARL_{in}	ARL_{out}
100	50	10	(1, 45)	(2, 1, 2, 4)	4	101.06	2.3 1.02	(3, 40)	(2, 1, 2, 4)	1	111.4	2.3 1.02
	100	5	(3, 100)	(4, 1, 2, 4)	3	123.51	4.73 1.22	(13, 78)	(4, 1, 2, 4)	1	102.4	4.72 1.14
	100	10	(2, 89)	(3, 1, 2, 4)	3	101.82	2.31 1.02	(12, 81)	(5, 5, 3, 5)	1	114.0	2.28 1.02
200	50	10	(1, 48)	(3, 1, 2, 4)	3	213.49	2.30 1.02	(3, 45)	(2, 1, 2, 4)	1	187.5	2.23 1.02
	100	5	(1, 100)	(3, 1, 2, 4)	2	244.88	6.59 1.29	(10, 80)	(3, 1, 2, 4)	1	230.9	5.62 1.24
	100	10	(4, 94)	(5, 5, 3, 5)	4	222.69	3.20 1.04	(10, 82)	(5, 5, 3, 5)	1	192.2	3.10 1.04

Each cell contains the AR's attained for Y = 0.5 (upper entry) and Y = 0.2 (lower entry).

Table 10.
Comparison of the ARL_{out}*s with the same* ARL_{in} *for* Chart *3.*

Author details

Ioannis S. Triantafyllou
Department of Computer Science and Biomedical Informatics, University of
Thessaly, Greece

*Address all correspondence to: itriantafyllou@uth.gr

IntechOpen

References

[1] Shewhart WA. Quality control charts. Bell System Technical Journal. 1926;**2**:593-603

[2] Montgomery DC. Introduction to Statistical Quality Control. 6th ed. New York: John Wiley; 2009

[3] Qiu P. Introduction to Statistical Process Control. New York: CRC Press, Taylor & Francis Group; 2014

[4] Chakraborti S, van der Laan P, van der Wiel MA. A class of distribution-free control charts. Journal of the Royal Statistical Society: Series C: Applied Statistics. 2004;**53**:443-462

[5] Balakrishnan N, Triantafyllou IS, Koutras MV. A distribution-free control chart based on order statistics. Communications in Statistics: Theory and Methods. 2010;**39**:3652-3677

[6] Triantafyllou IS. Nonparametric control charts based on order statistics: Some advances. Communications in Statistics: Simulation and Computation. 2018;**47**:2684-2702

[7] Triantafyllou IS. A new distribution-free monitoring scheme based on order statistics. Journal of Nonparametric Statistics. 2019;**31**:1-30

[8] Koutras MV, Triantafyllou IS, editors. Recent advances on univariate Shewhart type distribution free control charts. In: Distribution-free Methods for Statistical Process Monitoring and Control. Springer; 2019. DOI 10.1007/978-3-030-25081-2

[9] Chakraborti S, Graham M. Nonparametric Statistical Process Control. USA: John Wiley & Sons; 2019

[10] Qiu P. Some perspectives on nonparametric statistical process control. Journal of Quality Technology. 2018;**50**:49-65

[11] Qiu P. Some recent studies in statistical process control. In: Statistical Quality Technologies. 2019. pp. 3-19

[12] Klein M. Two alternatives to the Shewhart X control chart. Journal of Quality Technology. 2000;**32**:427-431

[13] van der Laan P, Chakraborti S. Precedence tests and Lehmann alternatives. Statistical Papers. 2001;**42**:301-312