GESTÃO & PRODUÇÃO

ISSN 0104-530X (Print)
ISSN 1806-9649 (Online)

THEMATIC SECTION: STATISTICAL PROCESS MONITORING AND CONTROL

Monitoring multinomial processes based on a weighted chi-square control chart

Monitoramento de processos multinomiais com base em um gráfico de controle qui-quadrado ponderado

Achouri Ali[1], Emira Khedhiri[2], Ramzi Talmoudi[3], Hassen Taleb[1]

[1]University of Carthage, ARBRE Laboratory, Tunis, Tunisia.
[2]University of Tunis, ARBRE Laboratory, Tunis, Tunisia.
[3]University of Carthage, ENVIE Laboratory LR18ES48, Nabeul, Tunisia.

How to cite: Achouri, A., Khedhiri, E., Talmoudi, R., & Taleb, H. (2021). Monitoring multinomial processes based on a weighted chi-square control chart. *Gestão & Produção*, 28(3), e43. https://doi.org/10.1590/1806-9649-2021v28e43

Abstract: Interpreting an out-of-control signal is a crucial step in monitoring categorical processes. For the Chi-Square Control Chart (CSCC), an out-of control situation does not specify if it was a process deterioration or a process improvement. For this reason, a weighted chi-square statistical control chart WSCC is proposed with different weighting categories in order to enable an accelerated disclosure of a control situation after a shift due to a deterioration of quality and on the other hand, decelerate an out of control situation after a shift due to a quality improvement. Furthermore, in comparison with Marcucci's method, the new procedure provides an accurate and easier way to interpret several signals. In other words, the WSCC allows a faster detection of an out-of control situation in the case of a quality deterioration, however, an out-of control situation is not quickly detected in the case of a quality improvement. Indeed, comparative studies have been performed to find the best control chart for each combination. Concluding remarks with comments and recommendations are given based on Average Run Length (ARL) and standard deviation run length (SDRL).

Keywords: Multinomial processes; Categorical processes; Chi-square control chart; Weighted chi-square statistic; ARL and SDRL.

Resumo: Interpretar um sinal fora de controle é uma etapa crucial no monitoramento de processos categóricos. Para o gráfico de controle do qui-quadrado (CSCC), uma situação fora de controle não especifica se foi uma deterioração do processo ou uma melhoria do processo. Por esta razão, um gráfico de controle estatístico qui-quadrado ponderado WSCC é proposto com diferentes categorias de ponderação, a fim de permitir uma divulgação acelerada de uma situação de controle após uma mudança devido a uma deterioração da qualidade e, por outro lado, desacelerar uma situação fora de controle após um turno devido a uma melhoria da qualidade. Além disso, em comparação com o método de Marcucci, o novo procedimento fornece uma maneira precisa e mais fácil de interpretar vários sinais. Em outras palavras, o WSCC permite uma detecção mais rápida de uma situação fora de controle no caso de uma deterioração da qualidade, entretanto, uma situação fora de controle não é detectada rapidamente no caso de uma melhoria da qualidade. Na verdade, estudos comparativos foram realizados para encontrar o melhor gráfico de controle para cada combinação. As observações finais com comentários e recomendações são fornecidas com base no comprimento médio de execução (ARL) e comprimento de execução de desvio padrão (SDRL).

Received July 30, 2020 – Accepted Oct. 16, 2020
Financial support: None.

Palavras-chave: Processos multinomiais; Processos categóricos; Gráfico de controle qui-quadrado; Estatística qui-quadrado ponderada; ARL e SDRL.

1 Introduction

Throughout the years, the importance of measurement and improvement of quality was enhanced in the investigation of continuous improvement of products and services. The use of statistical process control (SPC) via control charts has proven its efficiency in monitoring and improving manufacturing processes. The most commonly used SPC charts are those of Shewhart (1925). Shewhart & Deming (1939) discussed Statistical method of quality control. There are two different types of control charts which differ depending on the nature of the control characteristic. If the quality dimension is measured through a numerical scale, then a control chart by variables is used. On the other hand, if the product can only be categorized as defective or non-defective, control chart by attributes is applied. Steiner et al. (1996) proposed control charts to detect mean and standard deviation shifts based on grouped data. For the second case, Duncan (1950) developed a chi-square chart for controlling a set of percentages. The main purpose of this work is to focus on the attribute control chart. The attribute control chart is applied if the quality cannot be measured with numerical scale, such as appearance, softness, color, etc. Nelson (1987) investigated chi-square control chart for several proportions and Woodall (1997) discussed construction methods of control charts based on attribute data. Product units are then classified as either conforming or nonconforming, depending upon whether or not they meet specifications. The binary classification used in the p-chart might not be applied in several situations where product quality does not transform suddenly from conforming to non-conforming, and there might be a number of intermediate states such as conforming, minor non-conforming and major non-conforming. Hence with many categories of classification, the process develops multinomial random variables. Several researches were conducted in order to monitor such processes like Marcucci (1985), Raz & Wang (1990) and Taleb & Limam (2002). A Control charts for process average and variability based on linguistic data was proposed by Kanagawa et al. (1993), on the other hand, Tucker et al. (2002) analyzed control chart method for ordinal data. Taleb et al. (2006) discussed methods based on multivariate fuzzy multinomial control charts. Topalidou & Psarakis (2009) reviewed multinomial and multi-attribute quality control charts. The previous studies do not provide an idea about which category is responsible for the out-of-control situation and could not detect if it is a process deterioration or not. For this reason, a new chart using a Weighted Sum of Chi-squares that express the relative importance of all categories and with known quality proportions is proposed. Consequently, an efficient way is presented to interpret out-of-control signals. Classical control chart for attribute processes is discussed in Section 2. The framework for the proposed Weighted Sum of Chi-Square chart is presented in Section 3. An experimental study and a sensitivity analysis are given in Section 4 and 5 to illustrate the effectiveness of the new approach and compare it to the classical one.

2 Control chart for attribute processes

Marcucci (1985) introduces data where samples are classified into 3 categories such as conforming, non-conforming type A and non-conforming type B, with baseline proportions 0:95, 0:03 and 0:02 respectively. Table 1 shows simulated data using the Marcucci example parameters. The Marcucci procedure uses as a test statistic the chi-square statistic and is defined in two cases as follows:

– If such proportion π is unknown, then a common statistical task is on homogeneity testing of proportions between the base period and each monitoring period (Duncan, 1974), and the correspondingly test statistic is expressed as follows (Equation 1):

$$S_i^2 = \sum_{k=i,0}\sum_{j=1}^{t}\frac{n_k\left(\frac{X_{kj}}{n_k}-\frac{\left(X_{ij}+X_{0j}\right)}{n_i+n_0}\right)^2}{\frac{X_{ij}+X_{0j}}{n_i+n_0}} = n_i n_0 \sum_{j=1}^{t}\frac{\left(\pi_{ij}-\pi_{0j}\right)^2}{X_{ij}+X_{0j}} \tag{1}$$

where X_{kj} is the number of items of category j in sample k, n_k is the size of sample k and π_{kj} is the proportion of category j in sample k.

– If the process proportions are known, then a common statistical task which is the Pearson goodness-of-fit statistic is applied and defined as follows (Equation 2):

$$Z_i^2 = \sum_{j=1}^{t}\frac{\left(X_{ij}-n_i\pi_j\right)^2}{n_i\pi_j} \tag{2}$$

where X_{ij} is the number of items of category j in sample i, n_i is the size of sample i and π_j is the proportion of category j, $Z_i^2 \to \chi^2\left(q-1\right)$, where q is the number of categories of quality.

The upper control limit for the CSCC is expressed as a level of percentile of the chi-square distribution as follows (Equation 3):

$$UCL = \chi_\alpha^2\left(q-1\right) \tag{3}$$

where α is the significance level.

Table 1. Marcucci's Example.

Time	Conforming	Non-conforming Type A	Non-conforming Type B	Total	Zi^2	Yi^2
1	242	8	4	254	0.25	0.23
2	199	5	3	207	0.58	0.81
3	228	10	5	243	1.05	0.78
4	193	5	3	201	0.46	0.65
5	214	15	3	232	10.05	3.11
6	132	4	2	138	0.22	0.35
7	206	7	5	218	0.13	0.2
8	146	5	4	155	0.3	0.79
9	207	7	7	221	1.57	4.28
10	174	24	8	206	57.44	75.93
11	223	12	10	245	8.66	26.09
12	204	12	5	221	4.59	2.34
13	196	8	8	212	3.9	14.79
14	225	10	10	245	6.52	24.67
15	225	7	5	237	0.02	0.02
16	141	2	5	148	2.75	12.83

Consequently, the UCL control limit for the Marcucci's example is equal to 5:99. The CSCC provides an out-of-control signal when a single plotted point exceeds UCL. Therefore, it relies only on a single value to take decision and it is relatively insensitive to small process shift. Hence if we consider the additional data given in Table 2, we can notice that it is difficult to interpret the results. Both observations are significantly outside the control limit with a false alarm rate of 0.05 as used by Marcucci (1985). For observation 18, though we do not have non-conforming items, the process detects an out-of-control signal.

Table 2. Additional data for Marcucci Example.

Time	Conforming	Non-conforming Type A	Non-conforming Type B	Total	Z_i^2
17	200	10	50	260	395.5
18	260	0	0	260	13.7

Consequently, we can explain that the statistical test used by Marcucci, i.e., the Pearson statistic, is unable to detect whether we actually have a process deterioration or a process improvement. Then, as a conclusion, for the CSCC, an out-of control situation does not specify if it was a process deterioration or a process improvement. For this reason, a WSCC is proposed with different weighting categories in order to enable an accelerated disclosure of a control situation after a shift due to a deterioration of quality and on the other hand, decelerate an out-of-control situation after a shift due to a quality improvement. Furthermore, in comparison with Marcucci's method, the new procedure provides an accurate and easier way to interpret several signals. In other words, the WSCC allows a faster detection of an out-of control situation in the case of a quality deterioration, however, an out-of control situation is not quickly detected in the case of a quality improvement. Taleb & Limam (2002) proposed another approach which is based on membership functions, all categories are represented with only one representative value according to fuzzy theories and fuzzy sets. Fuzzy sets are composed by objects with different degrees of membership that varies in a range between 0 and 1, membership functions are associated to each object and their values come from the fact that it is impossible to precisely identify the class to which belongs objects in the set; the vagueness; the lack of clearly defined criteria for classifying and the presence of random variables makes it necessary to resort to fuzzy theory and membership functions to solve problem of identification. The classic function takes only 2 values 0 and 1 which indicates the certainty of the class that owns the object. If the value assigned is 0 so object does not belong to the set, else (value = 1) it is certain that the object is part of the set. However, the membership function can cover a multitude of values between 0 and 1 according to the degree of membership of an object to a class. If the value is different from 1 or 0, so we are doubting about the membership of the object to the class. The major difference between a characteristic function and a membership function is that the first is unique while the second is infinite and can generate an infinite number of membership functions. In addition, it consists only of weighting all the categories according to their proportions into a single representative value which, using fuzzy operations, makes it possible to derive a representative value capable of emphasizing the improvement or the deterioration of the quality. However, fuzzy theory is based on uncertainty and hesitation in the assignment due to human subjectivity and the shape of the membership function will change. Fuzzy sets are then drawn with intersection areas between categories. This problem of undefined limits makes better to privilege the probability theory and propose our new approach. In the following section, a new approach that offers an easier way to interpret an out-of-control signal is presented. The new method is called the Weighted Sum of Chi-squares method (WSCC).

3 The Weighted Sum of the Chi-square Control Chart (WSCC)

The WSCC consists in assigning a weight for each category as defined in the below Statistic (Equation 4):

$$Y_i^2 = \sum_{j=1}^{q} W_j Z_{i,j}^2 \tag{4}$$

In the test statistic, Wj is an appropriate way to determine the weights and shows a new method in order to find the relative weights of each category. This weight could be defined according to the following Equation 5:

$$W_j = e^{\left(j-\frac{q}{2}-1\right)\left(\frac{X_j - np_j}{\sqrt{np_j}}\right)} \tag{5}$$

Where:
q is the number of categories.
j = 1; …; q.
n is the sample size.
p_j is the proportion for each category.

for instance, if the number of categories is equal to q = 4, we will find the following results

$$j=1 \rightarrow j-\frac{q}{2}-1=-2$$

$$j=2 \rightarrow j-\frac{q}{2}-1=-1$$

$$j=3 \rightarrow j-\frac{q}{2}-1=0$$

$$j=4 \rightarrow j-\frac{q}{2}-1=1$$

on the other side of the weighting formula, and for the baseline proportions (p = [0:6; 0:25; 0:1; 0:05]) and the same number of categories (q = 4), we find for n = 1000 observation the following results:

- p = [0.6; 0.25; 0.1; 0.05] $\rightarrow$ we have 600 for C_1, 250 for C_2, 100 for C_3 and 50 for C_4

- if X_j gives 608 for C_1, then $\left(\frac{X_j - np_j}{\sqrt{np_j}}\right) = 0.32$

- if X_j gives 580 for C_1, then $\left(\frac{X_j - np_j}{\sqrt{np_j}}\right) = -0.81$

Table 3. An Illustration of the Effectiveness of CSCC.

	C_1			…	C_q				
	$j-\frac{q}{2}-1$	$\left(\frac{X_j - np_j}{\sqrt{np_j}}\right)$	W_j	…	$j-\frac{q}{2}-1$	$\left(\frac{X_j - np_j}{\sqrt{np_j}}\right)$	W_j	Y_i^2	
C_1 ↗ C_q ↘	-	+	$W_j \to 0$	…	+	-	$W_j \to 0$	↘	In control
C_1 ↘ C_q ↗	-	-	$W_j \to +\infty$	…	+	+	$W_j \to +\infty$	↗	Out of control

- if X_j gives 56 for C_4, then $\left(\frac{X_j - np_j}{\sqrt{np_j}}\right) = 0.84$

- if X_j gives 48 for C_4, then $\left(\frac{X_j - np_j}{\sqrt{np_j}}\right) = -0.28$

The aim of the new approach is to improve the sensitivity of Y_i^2 only on the shift affecting the worst categories, which leads to improve the sensitivity of the chart in detecting the process deterioration. Table 3 is provided in order to better assimilate the effectiveness of the WSCC. With C_1 and C_q represent respectively the best and the worst categories. In fact, in case of quality improvement, i.e, the number of observations in best category (C_1) increased, the corresponding weigh decreases, hence the process is in control. However, in case of quality deterioration, i.e, the number of observations in worst category (C_q) increased, the corresponding weigh increases, hence the process is out-of control. A lot of researches proved that the density function for the distribution of a weighted sum of independent chi-square random variables cannot be represented by elementary analytic functions. However, in many cases, the feasibility of approximating the distribution of Y_i^2 by a gamma distribution is proven feasible (Equation 6, 7 and 8), the first two moments are equal to the first two moments of Y_i^2 (Feiveson & Delaney, 1968). Thus:

$$G_{Y_i^2}(x) = \int_0^x \frac{\alpha^\lambda}{\Gamma(\lambda)} e^{-\alpha t} t^{\lambda - 1} dt \quad (6)$$

With:

$$\alpha = \left(\frac{1}{2}\right)\left(\frac{\sum W_i}{\sum W_i^2}\right) \tag{7}$$

$$\lambda = \left(\frac{1}{2}\right)\left(\frac{(\sum W_i)^2}{\sum W_i^2}\right) \tag{8}$$

Consequently, the UCL is equal to a level of percentile of the Gamma distribution expressed as follows (Equation 9):

$$UCL = \Gamma\left(p, \alpha, \lambda\right) \tag{9}$$

with p is the false alarm rate. However, determining the statistical test distribution of Y_i^2 with the new weight W_j is not an easy task due to the unequally weights, and we are not certain if it follows Gamma distribution or not, future research may focus on this distribution. Hence, in this paper, the UCL is computed using Simulation.

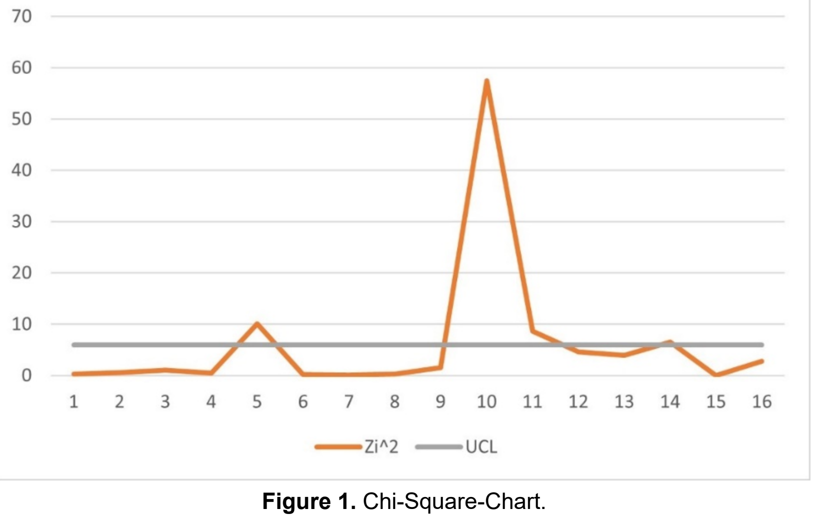

Figure 1. Chi-Square-Chart.

4 Experimental study

To illustrate the application of the proposed WSCC and the interpretation of its out-of-control signal, The Marcucci example given in Table 1 is considered. In our experimental study, some assumptions have been introduced; we set false alarm rate on α = 0.05. The corresponding upper control limit for the WSCC is calculated using simulation and it is equal to UCL = 5.02. Moreover, large sample sizes were investigated for this study, i.e n = 100; 300; 500; 1000. Then, for the chi-square chart, the UCL = 5.99 and it's equal to the percentile level of the chi-square distribution. Figure 1 and Figure 2 represent the resulting CSCC and the WSCC: As highlighted on Figure 1 and Figure 2, 4 samples are out-of control for the CSCC however 5 samples are out-of control for the WSCC . A close examination of the data related to those samples, indicated that despite sample 5 represents an improvement of the quality, however the CSCC consider it as out of control, but the proposed WSCC was able to differentiate that it was an improvement of the quality and consequently interpreted the signal as in-control. On the other side, if we analyze the data related to sample 13, we found out that this sample represents a deterioration of the quality. Consequently, the WSCC has succeeded in interpreting it as an out-of control situation. After this analysis, we can conclude, as a general rule, that the WSCC outperforms the CSCC in differentiating between the improvement and the deterioration of the quality.

The sensitivity analysis is determined in terms of the ARL and SDRL. In the field of Statistical Process Control, the random variable generally used to evaluate the performance of a control chart (based on the T_i statistic) is the Run Length defined by Equation 10:

$$RL = inf\left\{i = 1,2, \ldots \mid T_i \notin \left[LCL, UCL\right]\right\}. \tag{10}$$

When possible, it is important to evaluate its probability density function $f_{RL}(l)$, its cumulative distribution function $F_{RL}(l)$, its mean value ARL =E(RL) (Equation 11) and its standard deviation SDRL = δ(RL) (Standard-deviation Run Length), (Equation 12). Moreover, the number of samples needed to report an out of control situation (which equals the length of the sequence) is a geometric random variable RL (the support of the random variable Ω(RL) = (1; 2; 3; …; ∞). The ARL of this geometric variable RL with parameter p is is given by its first moment.

$$E\left(RL\right) = \frac{1}{p} = ARL \tag{11}$$

The variance of the geometric run length variable RL is

$$V\left(RL\right) = \frac{1-p}{p^2} = \left(1-p\right)ARL^2 \tag{12}$$

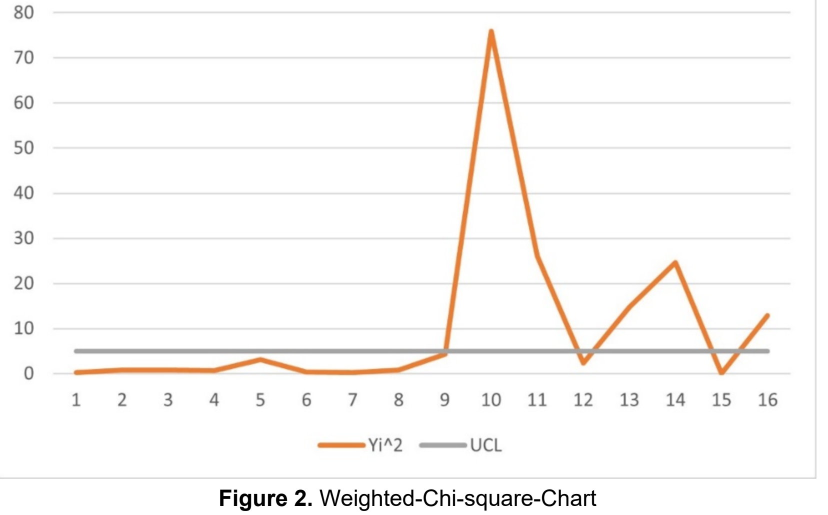

Figure 2. Weighted-Chi-square-Chart

Table 4. ARL for Known Proportions with q=5 and n=1000.

Baseline proportions	n = 1000 q=5			
P = [0.50 0.30 0.10 0.07 0.03]	UCL(Z_i^2)	SDRL(Z_i^2)	UCL(Y_i^2)	SDRL(Y_i^2)
	9.487729	19.1013	91.036737	19.453656
	ARL(Z_i^2)		ARL(Y_i^2)	
	19.607843		19.96008	
Improvement	ARL(Z_i^2)	SDRL(Z_i^2)	ARL(Y_i^2)	SDRL(Y_i^2)
P = [0.52 0.30 0.10 0.07 0.01]	1.0486577	0.2258877	136.9863	136.48538
P = [0.55 0.30 0.10 0.02 0.03]	1	0	35.460993	34.957417
P = [0.55 0.30 0.05 0.07 0.03]	1.0002	0.0141435	1.0316723	0.1807635
P = [0.58 0.22 0.10 0.07 0.03]	1.0007005	0.0264762	1.0001	0.0100005
P = [0.60 0.27 0.09 0.03 0.01]	1	0.0264762	2.7785496	2.2230133
Deterioration	ARL(Z_i^2)	SDRL(Z_i^2)	ARL(Y_i^2)	SDRL(Y_i^2)
P = [0.48 0.30 0.10 0.07 0.05]	1.2227929	0.5219479	1.0887316	0.3108133
P = [0.43 0.30 0.17 0.07 0.03]	1	0.0283062	1.0053282	0.0731887
P = [0.43 0.37 0.10 0.07 0.03]	1.0091836	0.0962701	1.0085729	0.0929860
P = [0.40 0.32 0.12 0.10 0.08]	1	0	1	0

Hence the standard deviation (SD) of the run length is

$$SD(RL) = \sqrt{1-p}\ \text{ARL} \quad (13)$$

5 Sensitivity analysis

In this section we summarize the results of a simulation study comparing the ARL and SDRL performance of the previously discussed CSCC and WSCC where the sample sizes are fixed, without loss of generality at n= 100, 300, 500, 1000, and a fixed number of categories at q= 3, 4, 5 (see from Table 4 to Table 15). This study provided a baseline information about the performance of the weighted sum of chi-square approach. In this simulation study, 10.000 samples are used to estimate each ARL and SDRL values. The Simulation study was done on two phases as below:

- Phase I
  - Step1: Generate 10:000 Samples with fixed size n=100, 300, 500, 1000, using known proportions p_0 = [p_{01}; p_{02}; …; p_{0q}] with p_{0j} is a specified proportion associated with each category j = 1; …; q.
  - Step 2: Calculate for each sample the corresponding Y_i^2 and Z_i^2
  - Step 3: Calculate the UCL_0 using a fixed Type I error for each chart.
- Phase II
  - Step 4: For fixed shifts of proportions vectors, generate 10:000 samples and calculate the corresponding Y_i^2 and Z_i^2
  - Step 5: Calculate the number of samples plotted outside the UCL_0 for both charts.
  - Step 6: Calculate the ARL and SDRL for each chart.

The following remarks can be deduced as conclusions from Table 4 to Table 15 below:

- Process Improvement:

As the number of categories increases, the values of ARL (Y_i^2) and SDRL (Y_i^2) increase:

P= [0.52, 0.30, 0.10, 0.07, 0.01] $\rightarrow$ ARL (Y_i^2) = 136.9863 and SDRL (Y_i^2) = 136.48538,

P= [0.64, 0.25, 0.10, 0.01] $\rightarrow$ ARL (Y_i^2) = 98.039216 and SDRL (Y_i^2) = 97.537934,

P= [0.75, 0.20, 0.05] $\rightarrow$ ARL (Y_i^2) = 67.567568 and SDRL (Y_i^2) = 67.065704,

As the sample size increases, the values of ARL (Y_i^2) and SDRL (Y_i^2) increase, for instance:

n= 100 $\rightarrow$ ARL (Y_i^2) = 99.009901 and SDRL (Y_i^2) = 98.508632

n= 300 $\rightarrow$ ARL (Y_i^2) = 21.716467 and SDRL (Y_i^2) = 19.96008

n= 500 $\rightarrow$ ARL (Y_i^2) = 112.35955 and SDRL (Y_i^2) = 111.85843

n= 1000 $\rightarrow$ ARL (Y_i^2) = 136.9863 and SDRL (Y_i^2) = 136.48538,

As the proportion in the best category increases, the ARL (Y_i^2) exceeds the ARL (Z_i^2):

P= [0.52, 0.30, 0.10, 0.07, 0.01] the ARL of the CSCC is equal to 1.0486577 however the ARL of the WSCC is equal to 136.9863.

- Process Deterioration:

As the sample size increases, the values of ARL (Y_i^2) and SDRL (Y_i^2) decreases:

n= 100 $\rightarrow$ ARL (Y_i^2) = 6.7204301 and SDRL (Y_i^2) = 6.2003025

n= 300 $\rightarrow$ ARL (Y_i^2) = 3.8774719 and SDRL (Y_i^2) = 3.3402569

n= 500 $\rightarrow$ ARL (Y_i^2) = 2.6226069 and SDRL (Y_i^2) = 2.0628766

n= 1000 $\rightarrow$ ARL (Y_i^2) = 1.6697278 and SDRL (Y_i^2) = 1.0574796,

As the number of categories increases, the values of ARL (Y_i^2) and SDRL (Y_i^2) decreases:

It is interesting to note that in most samples, the WSCC outperforms the CSCC in detecting a process deterioration comparing to CSCC.

Table 5. ARL for Known Proportions with q=4 and n=1000.

Baseline proportions	**n = 1000**			
	q=4			
P = [0.60 0.25 0.10 0.05]	UCL(Z_i^2)	SDRL(Z_i^2)	UCL(Y_i^2)	SDRL(Y_i^2)
	7.8147279	20.327185	31.619325	19.453656
	ARL(Z_i^2)		ARL(Y_i^2)	
	20.833333		19.96008	
Improvement	ARL(Z_i^2)	SDRL(Z_i^2)	ARL(Y_i^2)	SDRL(Y_i^2)
P = [0.64 0.25 0.10 0.01]	1	0	98.039216	97.537934
P = [0.65 0.25 0.05 0.05]	1.0002	0.0141435	2.9664788	2.4152676
P = [0.69 0.25 0.05 0.01]	1	0	3.866976	3.3296437
P = [0.70 0.15 0.10 0.05]	1	0	1	0
P = [0.75 0.15 0.05 0.05]	1	0	1	0
Deterioration	ARL(Z_i^2)	SDRL(Z_i^2)	ARL(Y_i^2)	SDRL(Y_i^2)
P = [0.55 0.25 0.10 0.10]	1	0	1	0
P = [0.52 0.25 0.15 0.08]	1	0	1	0
P = [0.50 0.35 0.10 0.05]	1	0	1.0001	0.0100005
P = [0.50 0.30 0.15 0.05]	1	0	1	0

Table 6. ARL for Known Proportions with q=3 and n=1000.

Baseline proportions	n = 1000			
	q=3			
P = [0.7 0.2 0.1]	UCL(Z_i^2)	SDRL(Z_i^2)	UCL(Y_i^2)	SDRL(Y_i^2)
	5.9914646	19.374265	11.225403	
	ARL(Z_i^2)		ARL(Y_i^2)	
	19.880716		1.96008	
Improvement	ARL(Z_i^2)	SDRL(Z_i^2)	ARL(Y_i^2)	SDRL(Y_i^2)
P = [0.75 0.20 0.05]	1	0	67.567568	67.065704
P = [0.75 0.15 0.10]	1.0306091	0.1776120	1.0291242	0.1731254
P = [0.76 0.20 0.04]	1	0	67.567568	67.065704
P = [0.78 0.16 0.06]	1.0001	0.0100005	1.1419436	0.4026060
P = [0.80 0.10 0.10]	1	0	1	0
Deterioration	ARL(Z_i^2)	SDRL(Z_i^2)	ARL(Y_i^2)	SDRL(Y_i^2)
P = [0.68 0.20 0.12]	2.1838829	1.6079371	1.6697278	1.0574796
P = [0.65 0.23 0.12]	1.1384335	0.3969853	1.0887316	0.3108133
P = [0.60 0.20 0.20]	1	0	1	0
P = [0.55 0.25 0.20]	1	0	1	0

Table 7. ARL for Known Proportions with q=5 and n=500.

Baseline proportions	n = 500			
	q=5			
P = [0.50 0.30 0.10 0.07 0.03]	UCL(Z_i^2)	SDRL(Z_i^2)	UCL(Y_i^2)	SDRL(Y_i^2)
	9.487729	19.1013	95.113769	
	ARL(Z_i^2)		ARL(Y_i^2)	
	19.607843		19.96008	
Improvement	ARL(Z_i^2)	SDRL(Z_i^2)	ARL(Y_i^2)	SDRL(Y_i^2)
P = [0.52 0.30 0.10 0.07 0.01]	1.7969452	1.19669	112.35955	111.85843
P=[0.55 0.30 0.10 0.02 0.03]	1.0031096	0.0558504	27.624309	27.1197
P = [0.55 0.30 0.05 0.07 0.03]	1.0654166	0.2639999	2.2727273	1.7007534
P = [0.58 0.22 0.10 0.07 0.03]	1.0659844	0.2652138	1.0539629	0.2384846
P = [0.60 0.27 0.09 0.03 0.01]	1.00040020	0.0200090	6.6269052	6.1064693
Deterioration	ARL(Z_i^2)	SDRL(Z_i^2)	ARL(Y_i^2)	SDRL(Y_i^2)
P = [0.48 0.30 0.10 0.07 0.05]	1.866368	1.271598	1.406272	0.7558630
P = [0.43 0.30 0.17 0.07 0.03]	1.0094892	0.0978736	1.1591515	0.4295122
P = [0.43 0.37 0.10 0.07 0.03]	1.1978917	0.4868807	1.1676787	0.4424871
P = [0.40 0.32 0.12 0.10 0.08]	1.0025063	0.0501257	1.0015023	0.4424871

Table 8. ARL for Known Proportions with q=4 and n=500.

Baseline proportions	n = 500			
	q=4			
P = [0.60 0.25 0.10 0.05]	UCL(Z_i^2)	SDRL(Z_i^2)	UCL(Y_i^2)	SDRL(Y_i^2)
	7.8147279	19.334736	31.426705	19.453656
	ARL(Z_i^2)		ARL(Y_i^2)	
	19.8412		19.96008	
Improvement	ARL(Z_i^2)	SDRL(Z_i^2)	ARL(Y_i^2)	SDRL(Y_i^2)

Table 8. Continued…

Baseline proportions	n = 500			
	q=4			
P = [0.64 0.25 0.10 0.01]	.0017029	0.0413013	103.09278	102.59156
P = [0.65 0.25 0.05 0.05]	1.0446046	0.2158568	15.151515	14.642981
P = [0.69 0.25 0.05 0.01]	1	0	46.296296	45.793566
P = [0.70 0.15 0.10 0.05]	1.0007005	0.0264762	1.0003001	0.0173260
P = [0.75 0.15 0.05 0.05]	1	0	1.0001	0.0100005
Deterioration	ARL(Z_i^2)	SDRL(Z_i^2)	ARL(Y_i^2)	SDRL(Y_i^2)
P = [0.55 0.25 0.10 0.10]	1.021868	0.1494865	1.0060362	0.0779271
P = [0.52 0.25 0.15 0.08]	1.0092854	0.0968071	1.0131712	0.1155192
P = [0.50 0.35 0.10 0.05]	1.0034116	0.0585085	1.0083695	0.0918670
P = [0.50 0.30 0.15 0.05]	1.0928962	0.3186313	1.0084712	0.0918670

Table 9. ARL for Known Proportions with q=3 and n=500.

Baseline proportions	n = 500			
	q=3			
P = [0.7 0.2 0.1]	UCL(Z_i^2)	SDRL(Z_i^2)	UCL(Y_i^2)	SDRL(Y_i^2)
	5.9914646	20.725393	11.48985	19.334806
	ARL(Z_i^2)		ARL(Y_i^2)	
	21.231423		19.84127	
Improvement	ARL(Z_i^2)	SDRL(Z_i^2)	ARL(Y_i^2)	SDRL(Y_i^2)
P = [0.75 0.20 0.05]	1.0276436	0.1685460	63.694268	63.19229
P = [0.75 0.15 0.10]	1.3504389	0.6879290	1.3417416	0.6771476
P = [0.76 0.20 0.04]	1.0015023	0.0387886	64.102564	63.600599
P = [0.78 0.16 0.06]	1.0239607	0.1566359	1.8261505	1.2282814
P = [0.80 0.10 0.10]	1	0	1	0
Deterioration	ARL(Z_i^2)	SDRL(Z_i^2)	ARL(Y_i^2)	SDRL(Y_i^2)
P = [0.68 0.20 0.12	3.8109756	3.2730046	2.6226069	2.0628766
P = [0.65 0.23 0.12]	1.69549	1.0859081	1.4976786	0.8633438
P = [0.60 0.20 0.20	1.0001	0.0100005	1	0
P = [0.55 0.25 0.20]	1	0	1	0

Table 10. ARL for Known Proportions with q=5 and n=300.

Baseline proportions	n = 300			
	q=5			
P = [0.50 0.30 0.10 0.07 0.03]	UCL(Z_i^2)	SDRL(Z_i^2)	UCL(Y_i^2)	SDRL(Y_i^2)
	9.487729	19.453656	86.818157	19.453656
	ARL(Z_i^2)		ARL(Y_i^2)	
	20.576132		19.96008	
Improvement	ARL(Z_i^2)	SDRL(Z_i^2)	ARL(Y_i^2)	SDRL(Y_i^2)
P = [0.52 0.30 0.10 0.07 0.01}	3.7664783	19.453655	19.96008	19.453656
P = [0.55 0.30 0.10 0.02 0.03]	1.0954102	19.453655	34.965035	34.461408
P = [0.55 0.30 0.05 0.07 0.03}	1.4361626	19.453655	7.3855244	6.8673464
P = [0.58 0.22 0.10 0.07 0.03}	1.3640704	19.453655	1.3390466	0.6737946
P = [0.60 0.27 0.09 0.03 0.01]	1.0302905	19.453655	11.600928	11.089662
Deterioration	ARL(Z_i^2)	SDRL(Z_i^2)	ARL(Y_i^2)	SDRL(Y_i^2)
P = [0.48 0.30 0.10 0.07 0.05]	2.7847396	19.453655	1.9219681	1.3311624
P = [0.43 0.30 0.17 0.07 0.03]	1.1093854	19.453655	1.4755792	0.8377080
P = [0.43 0.37 0.10 0.07 0.03]	1.6943409	19.453655	1.5239256	0.8935455
P = [0.40 0.32 0.12 0.10 0.08]	1.0493179	19.453655	1.0171905	0.1322347

Table 11. ARL for Known Proportions with q=4 and n=300.

Baseline proportions	n = 300			
	q=4			
P = [0.60 0.25 0.10 0.05]	UCL(Z_i^2)	SDRL(Z_i^2)	UCL(Y_i^2)	SDRL(Y_i^2)
	7.8147279	19.178522	34.163768	19.453656
	ARL(Z_i^2)		ARL(Y_i^2)	
	19.685039		19.96008	
Improvement	ARL(Z_i^2)	SDRL(Z_i^2)	ARL(Y_i^2)	SDRL(Y_i^2)
P = [0.64 0.25 0.10 0.01]	1.0980564	0.3281333	133.33333	132.83239
P = [0.65 0.25 0.05 0.05	1.3299641	0.6624503	22.421525	21.915822
P = [0.69 0.25 0.05 0.01]	1.0012014	0.0346820	80	79.498428
P = [0.70 0.15 0.10 0.05]	1.0303967	0.1769764	1.0234367	0.1548741
P = [0.75 0.15 0.05 0.05]	1.0002	0.0141435	1.0188487	0.1385784
Deterioration	ARL(Z_i^2)	SDRL(Z_i^2)	ARL(Y_i^2)	SDRL(Y_i^2)
P = [0.55 0.25 0.10 0.10]	1.1405109	0.4003176	1.0622477	0.2571429
P = [0.52 0.25 0.15 0.08	1.0976948	0.3274738	1.1201972	0.3669395
P = [0.50 0.35 0.10 0.05]	1.0678057	0.2690786	1.0992635	0.3303282
P = [0.50 0.30 0.15 0.05]	1.0928962	0.3186313	1.0949305	0.3224008

Table 12. ARL for Known Proportions with q=3 and n=300.

Baseline proportions	n = 300			
	q=4			
P = [0.7 0.2 0.1]	UCL(Z_i^2)	SDRL(Z_i^2)	UCL(Y_i^2)	SDRL(Y_i^2)
	5.9914646	18.577242	12.064423	
	ARL(Z_i^2)		ARL(Y_i^2)	
	19.083969		19.96008	
Improvement	ARL(Z_i^2)	SDRL(Z_i^2)	ARL(Y_i^2)	SDRL(Y_i^2)
P = [0.75 0.20 0.05]	1.2189176	0.5165680	70.921986	70.420211
P = [0.75 0.15 0.10]	2.0661157	1.4841558	2.0682523	1.4864102
P = [0.76 0.20 0.04]	1.039177	0.2017717	74.074074	73.572375
P = [0.78 0.16 0.06]	1.1944577	0.4819455	3.3557047	2.811592
P = [0.80 0.10 0.10]	1.0049241	0.0703445	1.0042177	0.0650806
Deterioration	ARL(Z_i^2)	SDRL(Z_i^2)	ARL(Y_i^2)	SDRL(Y_i^2)
P = [0.68 0.20 0.12]	5.4112554	4.8857374	3.8774719	3.3402569
P = [0.65 0.23 0.12]	2.463661	1.8989377	2.1710812	1.5945258
P = [0.60 0.20 0.20]	1.002004	0.0448109	1.0006004	0.0245104
P = [0.55 0.25 0.20]	1.0001	0.0100005	1	0

Table 13. ARL for Known Proportions with q=5 and n=100.

Baseline proportions	n = 100			
	q=5			
P = [0.50 0.30 0.10 0.07 0.03]	UCL(Z_i^2)	SDRL(Z_i^2)	UCL(Y_i^2)	SDRL(Y_i^2)
	9.487729	19.062916	122.05381	19.453656
	ARL(Z_i^2)		ARL(Y_i^2)	
	19.569472		19.96008	
Improvement	ARL(Z_i^2)	SDRL(Z_i^2)	ARL(Y_i^2)	SDRL(Y_i^2)
P = [0.52 0.30 0.10 0.07 0.01]	16.528926	16.021126	99.009901	98.508632

Table 13. Continued…

Baseline proportions	n = 100			
	q=5			
P = [0.55 0.30 0.10 0.02 0.03]	4.199916	3.6659758	26.525199	26.020396
P = [0.55 0.30 0.05 0.07 0.03]	5.3763441	4.8506424	26.315789	25.810947
P = [0.58 0.22 0.10 0.07 0.03]	3.9184953	3.3817318	4.3994721	3.8672836
P = [0.60 0.27 0.09 0.03 0.01]	2.6226069	2.0628766	41.322314	40.819252
Deterioration	ARL(Z_i^2)	SDRL(Z_i^2)	ARL(Y_i^2)	SDRL(Y_i^2)
P = [0.48 0.30 0.10 0.07 0.05]	5.6338028	5.1093964	3.875969	3.3387373
P = [0.43 0.30 0.17 0.07 0.03]	2.18436	1.6084367	4.2753313	3.7420751
P = [0.43 0.37 0.10 0.07 0.03]	4.4802867	3.948757	4.456328	3.9246059
P = [0.40 0.32 0.12 0.10 0.08]	1.7476407	3.948757	1.5807777	0.9581651

Table 14. ARL for Known Proportions with q=4 and n=100.

Baseline proportions	n = 100			
	q=4			
P = [0.60 0.25 0.10 0.05]	UCL(Z_i^2)	SDRL(Z_i^2)	UCL(Y_i^2)	SDRL(Y_i^2)
	7.8147279	21.56926	32.038648	19.453656
	ARL(Z_i^2)		ARL(Y_i^2)	
	22.075055		19.96008	
Improvement	ARL(Z_i^2)	SDRL(Z_i^2)	ARL(Y_i^2)	SDRL(Y_i^2)
P = [0.64 0.25 0.10 0.01]	6.0569352	5.5343951	151.51515	151.01432
P = [0.65 0.25 0.05 0.05]	4.7641734	4.2347579	26.385224	25.88039
P = [0.69 0.25 0.05 0.01]	1.699813	1.0906655	133.33333	132.83239
P = [0.70 0.15 0.10 0.05]	1.9912386	1.4049173	1.9853087	1.3986214
P = [0.75 0.15 0.05 0.05]	1.2674271	0.5821893	1.7979144	1.1977403
Deterioration	ARL(Z_i^2)	SDRL(Z_i^2)	ARL(Y_i^2)	SDRL(Y_i^2)
P = [0.55 0.25 0.10 0.10]	2.1181953	1.5390114	1.6594756	1.0461279
P = [0.52 0.25 0.15 0.08]	1.9868865	1.4002969	1.9508389	1.3619594
P = [0.50 0.35 0.10 0.05]	2.0395676	1.4561141	2.1168501	1.5375969
P = [0.50 0.30 0.15 0.05]	2.1168501	1.5375969	2.093364	1.512881

Table 15. ARL for Known Proportions with q=3 and n=100.

Baseline proportions	n = 100			
	q=3			
P = [0.7 0.2 0.1]	UCL(Z_i^2)	SDRL(Z_i^2)	UCL(Y_i^2)	SDRL(Y_i^2)
	5.9914646	21.13911	11.79648	19.334806
	ARL(Z_i^2)		ARL(Y_i^2)	
	21.645022		19.84127	
Improvement	ARL(Z_i^2)	SDRL(Z_i^2)	ARL(Y_i^2)	SDRL(Y_i^2)
P = [0.75 0.20 0.05]	4.0866367	3.5516141	81.967213	81.465679
P = [0.75 0.15 0.10]	6.0096154	5.4868809	6.4267352	5.9056067
P = [0.76 0.20 0.04]	2.5826446	2.021734	75.757576	75.255915
P = [0.78 0.16 0.06]	3.3590863	2.8150265	10.438413	9.9258276
P = [0.80 0.10 0.10]	1.512173	0.8800535	1.5304561	0.9010215
Deterioration	ARL(Z_i^2)	SDRL(Z_i^2)	ARL(Y_i^2)	SDRL(Y_i^2)
P = [0.68 0.20 0.12]	9.5785441	9.0647649	6.7204301	6.2003025
P = [0.65 0.23 0.12]	5.7208238	5.1968261	4.4622936	3.9306196
P = [0.60 0.20 0.20]	1.2669454	0.5815542	1.1500863	0.4154662
P = [0.55 0.25 0.20]	1.1082788	0.3464146	1.0625863	0.2578824

The ARL and SDRL performance comparison for the case where n=100, 300, 500, 1000 and q=3, 4, 5 is shown in the previous tables. They reveal in case of process improvement, it would be better to use the WSCC. In fact, this chart has better performance compared to CSCC. Whereas, in the case of a process deterioration, the WSCC does not suggest better outcomes as to the CSCC.

Finally, The main strengths of the WSCC lies in:

1. In a case of quality deterioration, this chart allows a fast detection of an out-of-control situation;
2. On the contrary of Marccucci chart, an out-of-control situation is not quickly detected in the case of a quality improvement.

6 Conclusion

Processes with multiple categories can be modeled as multinomial processes. Marcucci (1985) proposed a method to monitor such processes. However, this method does not allow a differentiation between a process improvement or process deterioration, it indicates only if the process is out-of-control or not. Hence in this study we propose a Weighted sum of the chi-square control chart to monitor any multinomial processes. The basic concept of this chart is to enhance each category by weights.

The corresponding UCL is calculated using simulation. According to an ARL comparison, we succeed to demonstrate that the proposed method outperforms the Chi-square chart. In fact, the WSCC control chart does not only allow a faster detection of an out-of control situation in a case of quality deterioration but also outperforms the Marcucci method in the case of a quality improvement as it does not quickly detect an out-of control situation.

Future researches will focus on expanding the study on other approaches such as implementing a Weighted sum of the chi-square control chart with unknown proportions, i.e proportions are not specified and should be estimated. It is well known that when in-control parameters are estimated, the performance of control charts differs from the known parameters case due to the variability of the estimators used during the Phase I.

References

Duncan, A. J. (1950). A chi-square chart for controlling a set of percentages. *Industrial Quality Control*, 7, 11-15.

Duncan, A. J. (1974). *Quality control and industrial statistics* (4th ed.). Homewood, IL: Richard D. Irwin.

Feiveson, A. H., & Delaney, F. C. (1968). *A chi-square chart for controlling a set of percentages. industrial quality control* (Vol. 38). Washington: National Aeronautics and Space Administration.

Kanagawa, A., Tamaki, F., & Ohta, H. (1993). Control charts for process average and variability based on linguistic data. *International Journal of Production Research*, 31(4), 913-922. http://dx.doi.org/10.1080/00207549308956765.

Marcucci, M. (1985). Monitoring multinomial processes. *Journal of Quality Technology*, 17(2), 86-91. http://dx.doi.org/10.1080/00224065.1985.11978941.

Nelson, L. S. (1987). A chi-square control chart for several proportions. *Journal of Quality Technology*, 19(4), 229-231. http://dx.doi.org/10.1080/00224065.1987.11979069.

Raz, T., & Wang, J. (1990). On the construction of control charts using linguistic variables. *International Journal of Production Research*, 28(3), 477-487. http://dx.doi.org/10.1080/00207549008942731.

Shewhart, W. A. (1925). The application of statistics as an aid in maintaining quality of a manufactured product. *Journal of the American Statistical Association*, 20(152), 546-548. http://dx.doi.org/10.1080/01621459.1925.10502930.

Shewhart, W. A., & Deming, W. E. (1939). *Statistical method from the viewpoint of quality control*. Washington: The Graduate School, The Department of Agriculture.

Steiner, S. H., Geyer, P. L., & Wesolowsky, G. O. (1996). Shewhart control charts to detect mean and standard deviation shifts based on grouped data. *Quality and Reliability Engineering International*, 12(5), 345-353. http://dx.doi.org/10.1002/(SICI)1099-1638(199609)12:5<345::AID-QRE11>3.0.CO;2-M.

Taleb, H., Limam, M., & Hirota, H. (2006). Multivariate fuzzy multinomial control charts. *Quality Technology & Quantitative Management*, 3(4), 437-453. http://dx.doi.org/10.1080/16843703.2006.11673125.

Taleb, M., & Limam, M. (2002). H.and Limam. On fuzzy and probabilistic control charts. *International Journal of Production Research*, 40(12), 2849-2863. http://dx.doi.org/10.1080/00207540210137602.

Topalidou, E., & Psarakis’, S. (2009). Review of multinomial and multi attribute quality control charts. *Quality and Reliability Engineering International*, 25(7), 773-804. http://dx.doi.org/10.1002/qre.999.

Tucker, G. R., Woodall, W. H., & Tsui, K. L. (2002). control chart method for ordinal data. *American Journal of Mathematical and Management Sciences*, 22(1-2), 31-48. http://dx.doi.org/10.1080/01966324.2002.10737574.

Woodall, W. H. (1997). Control charts based on attribute data: bibliography and review. *Journal of Quality Technology*, 29(2), 172-183. http://dx.doi.org/10.1080/00224065.1997.11979748.

GESTÃO & PRODUÇÃO
ISSN 0104-530X (Print)
ISSN 1806-9649 (Online)

ERRATUM

ERRATUM: Monitoring multinomial processes based on a weighted chi-square control chart

Achouri Ali[1], Emira Khedhiri[2], Ramzi Talmoudi[3], Hassen Taleb[1]

[1]University of Carthage, ARBRE Laboratory, Tunis, Tunisia.
[2]University of Tunis, ARBRE Laboratory, Tunis, Tunisia.
[3]University of Carthage, ENVIE Laboratory LR18ES48, Nabeul, Tunisia.

Due to desktop publishing error the article “Monitoring multinomial processes based on a weighted chi-square control chart” (DOI https://doi.org/10.1590/1806-9649-2021v28e43), published in *Gestão & Produção*, 28(3), e43, 2021, was published with an error.

On page 1, where the text reads:
ORIGINAL ARTICLE

It should read:
THEMATIC SECTION: STATISTICAL PROCESS MONITORING AND CONTROL
The publisher apologizes for the errors.