A sequential test and a sequential sampling plan based on the process capability index Cpmk”. Computational Statistics



*Published Version:*
Michele Scagliarini (2022). A sequential test and a sequential sampling plan based on the process capability index Cpmk”. Computational Statistics. COMPUTATIONAL STATISTICS, 37(3 (July)), 1523-1550 [10.1007/s00180-021-01169-1].

*Availability:*
This version is available at: https://hdl.handle.net/11585/851733 since: 2022-06-14

*Published:*
DOI: http://doi.org/10.1007/s00180-021-01169-1

**TITLE**

# A sequential test and a sequential sampling plan based on the process capability index $C_{pmk}$

**AUTHOR**

Michele Scagliarini ORCID: 0000-0003-3209-4745
Department of Statistical Sciences, University of Bologna, Italy
Via delle Belle Arti, 41, 40126, Bologna, Italy
Email: michele.scagliarini@unibo.it

**Abstract**

In this study we propose a sequential test for hypothesis testing on the $C_{pmk}$ process capability index. Furthermore, we propose a sequential sampling plan for lot acceptance based on $C_{pmk}$. We compare the statistical properties of the sequential procedures with the performance of the corresponding non-sequential methodologies by carrying out an extensive simulation study. The results show that the proposed sequential methods make it possible to reach decisions much more quickly, on average, than the fixed sample size procedures with the same discriminating power.

**Keywords**: process capability indices; acceptance sampling plan; average sample size, power function, producer and consumer risks

Funding: There is no funding source.
Conflict of Interest: None.
Availability of data and material: Not applicabile
Code availability: (R code available as Supplementary Information)

## 1. Introduction

Process capability indices are the tools usually used for assessing processes performance in relationship to the design specifications. If we represent the process mean with $\mu$, the process standard deviation with $\sigma$, the specification limits with *LSL* and *USL*, the half-length of the specification interval as $d=(USL-LSL)/2$, the midpoint of the specification interval as $m=(USL+LSL)/2$, and the target value of the process as *T*, the capability indices most widely used are (Kotz and Lovelace 1998, Polansky and Kirmani 2003, Pearn and Kotz 2006, Montgomery 2009): the potential capability index

$$C_p = \frac{USL-LSL}{6\sigma} = \frac{d}{3\sigma}, \tag{1}$$

the actual capability index

$$C_{pk} = \frac{d-\left|\mu-\frac{1}{2}(USL+LSL)\right|}{3\sigma} = \frac{d-|\mu-m|}{3\sigma}, \tag{2}$$

the loss-based capability index

$$C_{pm} = \frac{USL-LSL}{6\sqrt{\sigma^2+(\mu-T)^2}} = \frac{d}{3\sqrt{\sigma^2+(\mu-T)^2}} \tag{3}$$

and the index

$$C_{pmk} = \frac{d-|\mu-m|}{3\sqrt{\sigma^2+(\mu-T)^2}}, \tag{4}$$

which is constructed by combining the indices $C_{pk}$ and $C_{pm}$. Note that, for processes with target value set to the mid-point of the specification limits (*T*=*m*), the index $C_{pmk}$ can be rewritten as

$$C_{pmk} = \frac{d/\sigma-|\xi|}{3\sqrt{1+\xi^2}}, \tag{5}$$

where $\xi = \dfrac{\mu-T}{\sigma}$.

In the context of contractual agreements, it is often a requirement to provide evidence that a manufacturing process satisfies a minimum level of capability. The decision-making problem of demonstrating whether the process capability exceeds a predetermined capability requirement can be approached in terms of hypothesis testing.

Literature concerning process capability hypothesis testing includes significant amounts of interesting research. To name but a few: the pioneering work by Kane (1986); the tests on $C_{pk}$ investigated by Pearn and Chen (1999), Perakis and Xekalaki (2003), Pearn and Lin (2004), Chen and Hsu (2004) and Lin (2006); the Bayesian approach proposed by Fan and Kao (2006); the hypothesis testing studies on $C_{pmk}$ by Pearn and Lin (2002) and Pearn et al. (2005); the model free approach testing procedure proposed by Vännman and Kulachi (2002); the study of the effect of pooled and un-pooled variance estimators on hypothesis testing concerning $C_{pm}$ by Hubele and Vännman (2004); the unified analysis of hypothesis testing with process capability indices by Lepore and Palumbo (2015); the sequential procedure for testing the equality of two indices $C_{pm}$ by Hussein et al. (2012), and the sequential test proposed by Scagliarini (2018).

Furthermore, due to their ability to summarize adherence to design specifications in a single number, capability indices have also proven to be useful in the framework of lot-by-lot acceptance-sampling plans.

Sampling plans provide the producer and consumer with general decision rules for product acceptance while meeting their needs for product quality and managing the risks of not adequately reflecting the quality conditions of the lot. Acceptance sampling plans basically consist of a required sample size for inspection and an acceptance criterion, so that the producer and consumer risks meet predetermined standards.

Sampling plans can be classified by: variables when the quality characteristics are measured on a numerical or continuous scale; attributes when characteristics are expressed on a "go, no-go" basis. Generally, numerical measurements of quality characteristics provide more information about the manufacturing process or lot than do attributes data. Thus, in today's manufacturing industries where the allowable proportion of nonconforming products is very low, often measured in parts per million (PPM), variables sampling plans become very attractive since they allow the sample size to be significantly reduced.

The relevance of process capability indices in the context of acceptance variables sampling plans is confirmed by numerous studies, including: Arizono et al. (1997); Pearn and Wu (2006a, and 2006b); Wu and Pearn (2008); Yen and Chang (2009); Negrin et al. (2009 and 2011); Wu et al. (2012); Lepore et al. (2018).

In this study, we propose a sequential test for the index $C_{pmk}$ and a sequential sampling plan for lot acceptance based on $C_{pmk}$.

We analytically derive the test statistic of the sequential test and describe in detail the testing procedure.

We then compare the statistical properties of the sequential test with those of the most widely used non-sequential test by performing an extensive simulation study.

In analysing the performance of the sequential procedure, several issues should be taken into consideration, such as, for example, the behaviour of the stopping sample size required by the test to decide in favour of $H_0$ and $H_1$. This type of analysis is quite complex. However, it is more intuitive and easier to develop if worked out in the framework of a variables sampling plan based on $C_{pmk}$. In such a way we can also propose a sequential sampling plan for variables which uses the index $C_{pmk}$ as the benchmark for acceptance of a batch of products.

The results show that the proposed sequential procedures make it possible to reach decisions which, on average, allow the sample size to be reduced when compared with the non-sequential procedures, with consequent financial benefits and without any loss in quality.

The paper is organized as follows. In Section 2, we review the non-sequential test, proposed by Pearn and Lin (2002), used for assessing whether a process is capable or not based on the $C_{pmk}$ process capability index. In Section 3, we present the general sequential test procedure, proposed by Hussein (2005) and Hussein et al. (2012). In Section 4, we analytically obtain the test statistic and propose the sequential procedure for hypotheses testing on $C_{pmk}$. In Section 5, using the results obtained in Section 4, we develop a sequential sampling plan based on $C_{pmk}$. In Section 6, we study the performance of the sequential test and the sequential sampling plan through a set of Monte Carlo simulations, and compare the proposed methods with the corresponding non-sequential procedures. Section 7 contains a discussion of the results. Finally, our concluding remarks are given in Section 8.

## 2. Hypothesis test for the index $C_{pmk}$

To verify whether a process conforms to specifications using the index $C_{pmk}$ the hypotheses of interest are:

$H_0 : C_{pmk} \leq C_{pmk0}$ (the process is not capable),

versus

$H_1 : C_{pmk} > C_{pmk0}$ (the process is capable).

Assuming a normally distributed quality characteristic, $X \sim N\left(\mu, \sigma^2\right)$, Pearn and Lin (2002) proposed a statistical test (PL-test) based on the distribution of the estimator:

$$\hat{C}_{pmk} = \min\left\{ \frac{USL - \bar{X}}{3\sqrt{S_n^2 + \left(\bar{X} - T\right)^2}}, \frac{\bar{X} - LSL}{3\sqrt{S_n^2 + \left(\bar{X} - T\right)^2}} \right\} = \frac{d - \left|\bar{X} - T\right|}{3\sqrt{S_n^2 + \left(\bar{X} - T\right)^2}}, \tag{6}$$

where $\bar{X} = \sum_{i=1}^{n} X_i \Big/ n$ and $S_n^2 = \sqrt{\sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2 \Big/ n}$ are the maximum likelihood estimators of $\mu$ and $\sigma^2$, respectively.

For the case $T = \left(USL + LSL\right)/2$, the cumulative distribution function of $\hat{C}_{pmk}$ can be expressed as

$$F_{\hat{C}_{pmk}}\left(x\right) = 1 - \int_0^{b\sqrt{n}/(1+3x)} G\left( \frac{\left(b\sqrt{n} - t\right)^2}{9x^2} - t^2 \right) \left[\phi\left(t + \xi\sqrt{n}\right) + \phi\left(t - \xi\sqrt{n}\right)\right] dt \tag{7}$$

for $x>0$, where $b=d/\sigma$, $G(\cdot)$ is the cumulative distribution function of the chi-square distribution $\chi^2_{n-1}$, and $\phi(\cdot)$ is the probability density function of the standard normal distribution.

The decision rule of the test is to find $C_0$ and reject $H_0 : C_{pmk} \le C_{pmk0}$ if $\hat{C}_{pmk} > C_0$ and do not reject $H_0$ otherwise.

Given the values of the type-I error probability $\alpha_{PL}$, the capability requirement $C_{pmk0}$, the sample size $n$ and the parameter $\xi$, the critical value $C_0$ can be obtained by solving the equation

$$\int_0^{b\sqrt{n}/(1+3C_0)} G\left( \frac{\left(b\sqrt{n} - t\right)^2}{9C_0^2} - t^2 \right) \left[\phi\left(t + \xi\sqrt{n}\right) + \phi\left(t - \xi\sqrt{n}\right)\right] dt = \alpha_{PL}. \tag{8}$$

It can be noted that equation (8) depends on the additional parameter $\xi$, which in real applications is unknown. The estimation of $\xi$ introduces additional sampling errors in finding the critical value $C_0$. Therefore, to eliminate the need to estimate $\xi$, Pearn and Lin (2002) studied the behaviour of $C_0$ as a function of $\xi$. They found that the condition $\xi = 0.5$ should provide conservative critical values (the condition would be $\left|\xi\right| = 0.5$, however since (8) is an even function of $\xi$, the results for positive and negative values are the same). The results for $C_0$ and the sample size *n*, for several combinations of $\alpha_{PL}$ and $C_{pmk0}$, are given by Pearn and Lin (2002) in their Tables 1-5.

Hence, for $C_{pmk1} \geq C_{pmk0}$, given $\alpha_{PL}$, *n* and $C_0$, the power of the PL-test can be written as

$$\begin{aligned}\pi_{PL}\left(C_{pmk1}\right) &= \Pr\left(\hat{C}_{pmk} \geq C_0 \middle| C_{pmk} = C_{pmk1}\right) \\ &= \int_0^{b\sqrt{n}/(1+3C_0)} G\left(\frac{\left(b\sqrt{n}-t\right)^2}{9C_0^2} - t^2\right)\left[\phi\left(t+\xi\sqrt{n}\right)+\phi\left(t-\xi\sqrt{n}\right)\right]dt,\end{aligned} \tag{9}$$

where $b = d/\sigma = 3C_{pmk1}\sqrt{1+\xi^2} + |\xi|$.

## 3. A general sequential method

Here we describe the general sequential testing procedure proposed by Hussein (2005) and Hussein et al. (2012). In Section 4, this general multivariate framework will be specified for sequential hypothesis testing on $C_{pmk}$.

Let us denote with $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k, ...,$ a sequence of independent observations of the random vector **X**, collected over time, with $\mathbf{X} \in R^l$. We assume that these data come from a common multivariate distribution with density function $f\left(\mathbf{x};\boldsymbol{\theta}\right)$, where the $d \times 1$ vector of parameters $\boldsymbol{\theta}$ is unknown.

We are interested in testing

$$H_0 : h\left(\boldsymbol{\theta}\right) = \mathbf{0} \text{ versus } H_1 : h\left(\boldsymbol{\theta}\right) \neq \mathbf{0}, \tag{10}$$

where $h\left(\boldsymbol{\theta}\right) : R^d \to R^q$, with $q \leq d$, is a function with first order derivative matrix denoted by $H\left(\boldsymbol{\theta}\right)$.

Let us assume that for $\boldsymbol{\theta} \in \Theta \subset R^d$ with $q \leq d$ the following regularity conditions hold (Hussein et al. 2012, Scholz 2006):

C1. The distribution function $F\left(\mathbf{x};\boldsymbol{\theta}\right)$ of the random vector **X** is identifiable over $\boldsymbol{\theta} \in \Theta$, *i.e.* if $\boldsymbol{\theta} \neq \boldsymbol{\theta}'$, then $F\left(\mathbf{x};\boldsymbol{\theta}\right)$ and $F\left(\mathbf{x};\boldsymbol{\theta}'\right)$ are different distributions.

C2. There exists an open subset, $\Theta_0 \subset \Theta$, containing the true value of the parameter under $H_0$, such that the partial derivatives

$$\frac{\partial}{\partial \theta_i} \ln f\left(\mathbf{x};\boldsymbol{\theta}\right),\ \frac{\partial^2}{\partial \theta_i \partial \theta_j} \ln f\left(\mathbf{x};\boldsymbol{\theta}\right),\ \frac{\partial^3}{\partial \theta_i \partial \theta_j \partial \theta_k} \ln f\left(\mathbf{x};\boldsymbol{\theta}\right)$$

exist, are exchangeable, and are continuous for all $\mathbf{x} \in R^l$, $\boldsymbol{\theta} \in \Theta_0$.

C3. For each $\boldsymbol{\theta} \in \Theta_0$ and $k = 1, 2, 3, ...,$ the score equation $\sum_{i=1}^{k} \frac{\partial \ln f(\mathbf{x};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = 0$ has a unique solution.

C4. There are functions of **x**, $M_1(\mathbf{x})$ and $M_2(\mathbf{x})$, that have finite expectations under any of the parameter values, $\boldsymbol{\theta} \in \Theta_0$, such that

$$\left|\frac{\partial}{\partial \theta_i} \ln f(\mathbf{x};\boldsymbol{\theta})\right| \le M_1(\mathbf{x}),\ \left|\frac{\partial^2}{\partial \theta_i \partial \theta_j} \ln f(\mathbf{x};\boldsymbol{\theta})\right| \le M_2(\mathbf{x}),\ \left|\frac{\partial^3}{\partial \theta_i \partial \theta_j \partial \theta_k} \ln f(\mathbf{x};\boldsymbol{\theta})\right| \le M_2(\mathbf{x}),$$

for all $\boldsymbol{\theta} \in \Theta_0$, $1 \le i \quad j, k \le d$.

C5. $E_\theta \left(\partial / \partial \theta_i\right) \ln f(\mathbf{x};\boldsymbol{\theta}) = 0$, $1 \le i \le d$, $\boldsymbol{\theta} \in \Theta_0$, where $E_\theta(\ )$ denotes that the expectation is parameterized by $\theta$.

C6. The inverse of the Fisher information matrix $I^{-1}(\theta)$ and its elements $I_{ij}(\theta) = E_\theta\left[\left(\partial^2 / \partial \theta_i \partial \theta_j\right) \ln f(\mathbf{x};\boldsymbol{\theta})\right]$ exist and are continuous for all $\boldsymbol{\theta} \in \Theta_0$, $1 \le i$, $j \le d$.

C7. $\mathrm{Var}_\theta\left[\left(\partial^2 / \partial \theta_i \partial \theta_j\right) \ln f(\mathbf{x};\boldsymbol{\theta})\right] < \infty$ for $1 \le i$, $j \le d$.

Let us further assume that:

C8. $E_\theta \left|\left(\partial / \partial \theta_i\right) \ln f(\mathbf{x};\boldsymbol{\theta})\right|^{2+\delta} < \infty$, $i = 1, 2, ..., d$, and for some $\delta > 0$.

C9. The function $h(\boldsymbol{\theta})$ is continuously differentiable over $\Theta_0$, and its first-order derivative matrix $H(\boldsymbol{\theta})$ is bounded and of rank $q$.

Let us now consider a fixed sample design with sample size equal to $k$ and let us consider the Wald's statistic

$$W_k = k h\left(\hat{\boldsymbol{\theta}}_k\right)\left[H^t(\boldsymbol{\theta}) I^{-1}(\boldsymbol{\theta}) H(\boldsymbol{\theta})\right]^{-1} h\left(\hat{\boldsymbol{\theta}}_k\right)^t, \tag{11}$$

where $\hat{\boldsymbol{\theta}}_k$ is a consistent estimator of $\boldsymbol{\theta}$.

Hussein et al. (2012) in Theorem 1 showed that under $H_0$, and if conditions C1-C9 hold, there exists an independent Wiener process, $B_j(t)$, $j = 1, 2, ..., q$, such that for $\alpha \le \frac{1}{2} - 1/(2+\delta)$ with $\delta$>0,

$$\sup_{1 \le t < \infty} \left|W_{[kt]} - U_{[kt]}\right| = O\left(k^{-\alpha} (\ln \ln k)^{1/2}\right), \tag{12}$$

where, $O(.)$ means "of the same order as",

$$U(x)=\frac{1}{x}\sum_{j=1}^{q}B_j^2(x), \tag{13}$$

$$W_{[kt]}=[kt]h\left(\hat{\boldsymbol{\theta}}_{[kt]}\right)\left[H^t(\boldsymbol{\theta})I^{-1}(\boldsymbol{\theta})H(\boldsymbol{\theta})\right]^{-1}h\left(\hat{\boldsymbol{\theta}}_{[kt]}\right)^t \tag{14}$$

and $[.]$ denotes the integer part of its argument.

The statistic $W_k$ can therefore be approximated by a functional of Brownian motions.

Furthermore, the authors derived the limiting distribution of $W_k$. In detail they showed that (Corollary 1):

- Under the conditions of Theorem 1

$$\max_{1<k\le n}\left[\frac{k}{n}W_k\right]^{1/2}\xrightarrow{D}\sup\left(\sum_{j=1}^{q}B_j^2(t)\right)^{1/2}, \tag{15}$$

  where $\xrightarrow{D}$ denotes convergence in distribution;

- When replacing the unknown $\boldsymbol{\theta}$ in the term $H^t(\boldsymbol{\theta})I^{-1}(\boldsymbol{\theta})H(\boldsymbol{\theta})$ with any almost surely convergent estimator, Corollary 1 remains valid.

Therefore, Hussein et al. (2012) defined the following as test statistic

$$W_k^*=kh\left(\hat{\boldsymbol{\theta}}_k\right)\left[H^t\left(\hat{\boldsymbol{\theta}}_k\right)I^{-1}\left(\hat{\boldsymbol{\theta}}_k\right)H\left(\hat{\boldsymbol{\theta}}_k\right)\right]^{-1}h\left(\hat{\boldsymbol{\theta}}_k\right)^t, \tag{16}$$

where $\hat{\boldsymbol{\theta}}_k$ is the maximum likelihood estimator of $\boldsymbol{\theta}$, and proposed the following $\alpha_S$-level sequential test truncated at the maximum allowable sample size $n_0$.

The sequential test procedure is performed as follows:

- For $k=2,3,...,n_0$ calculate the statistic

$$W_k^{*(1)}=\sqrt{k/n_0}\sqrt{W_k^*}\ ; \tag{17}$$

- Reject the hypothesis $H_0$ the first time that $W_k^{*(1)}$ exceeds the critical value $w_{\alpha_S}$. Given the Type I error probability $\alpha_S$, the critical value $w_{\alpha_S}$ can be obtained in accordance with Borodin and Salminen (1996);
- If $W_k^{*(1)}$ does not exceed $w_{\alpha_S}$ by $n_0$, then stop the sampling and do not reject $H_0$.

The maximum sample size $n_0$ can be decided on the basis of financial, ethical or statistical reasons, for example as the smallest one that achieves a desired power level.

## 4. A sequential test for $C_{pmk}$

Let us now consider the hypothesis

$$H_0 : C_{pmk} = C_{pmk0}, \tag{18}$$

versus

$$H_1 : C_{pmk} \neq C_{pmk0} \tag{19}$$

and assume that the quality characteristic is normally distributed $X \sim N\left(\mu, \sigma^2\right)$.

For $C_{pmk} \geq 0$, $H_0$ is equivalent to

$$H_0 : \ln\left(\left(C_{pmk}\right)^2\right) - \ln\left(\left(C_{pmk0}\right)^2\right) = 0 \tag{20}$$

and the alternative hypothesis is equivalent to

$$H_1 : \ln\left(\left(C_{pmk}\right)^2\right) - \ln\left(\left(C_{pmk0}\right)^2\right) \neq 0. \tag{21}$$

Let us define the function $h(\boldsymbol{\theta})$ as

$$h(\boldsymbol{\theta}) = \ln\left(\left(C_{pmk}\right)^2\right) - \ln\left(\left(C_{pmk0}\right)^2\right) = \ln\left[\left(d/\sigma - |\xi|\right)^2 \Big/ 9\left(1+\xi^2\right) C_{pmk0}^2\right], \tag{22}$$

where, in analogy with the PL-test, we assume $\xi$ as known. In such a way, as far as this issue is concerned, the sequential test is comparable with the PL-test. Please note that this assumption will be relaxed later in this Section and the effects $\xi$ on the sequential test will be analysed in Section 7.

With $\xi$ assumed to be known, we have $\boldsymbol{\theta} = \sigma^2$ and $h(\boldsymbol{\theta}) : R^d \to R^q$ with $d$=$q$=1 (Appendix A provides details concerning the satisfaction of conditions C1 to C9). Thus, the first order derivative matrix $H(\boldsymbol{\theta})$ can be written as

$$H(\boldsymbol{\theta}) = \left[\frac{\partial h\left(\sigma^2\right)}{\partial \sigma^2}\right] = \frac{d\left[\sqrt{\sigma^2}\,|\xi| - d\right]}{\sigma^2\left(\xi^2\sigma^2 - 2d\,|\xi|\sqrt{\sigma^2} + d^2\right)}. \tag{23}$$

In the case at hand the statistic $W_k^*$ is

$$W_k^* = kh^2\left(\hat{\boldsymbol{\theta}}_k\right)\left[H^t\left(\hat{\boldsymbol{\theta}}_k\right) I^{-1}\left(\hat{\boldsymbol{\theta}}_k\right) H\left(\hat{\boldsymbol{\theta}}_k\right)\right]^{-1}, \tag{24}$$

where $\hat{\boldsymbol{\theta}}_k = S_k^2$ with $S_k^2 = \sum_{i=1}^{k}\left(X_i - \bar{X}_k\right)^2 \Big/ k$, the function $h\left(\hat{\boldsymbol{\theta}}_k\right)$ is given by

$$h\left(\hat{\boldsymbol{\theta}}_k\right)=\ln\left[\left(d\big/\sqrt{S_k^2}-|\xi|\right)^2\Big/9\left(1+\xi^2\right)C_{pmk0}^2\right] \tag{25}$$

and the partial derivative matrix of $h(\boldsymbol{\theta})$ computed at $\hat{\boldsymbol{\theta}}_k$ can be written as

$$H\left(\hat{\boldsymbol{\theta}}_k\right)=\frac{d\left[\sqrt{S_k^2}\,|\xi|-d\right]}{S_k^2\left(\xi^2 S_k^2-2d\,|\xi|\sqrt{S_k^2}+d^2\right)}. \tag{26}$$

For normally distributed data with $\boldsymbol{\theta}=\sigma^2$, the Fisher information matrix is $I(\boldsymbol{\theta})=\dfrac{1}{2\sigma^4}$; consequently, $W_k^*$ can be written as

$$W_k^*=\frac{k\left(\ln\left[\left(d\big/\sqrt{S_k^2}-|\xi|\right)^2\Big/9\left(1+\xi^2\right)C_{pmk0}^2\right]\right)^2}{\dfrac{\left[d\left(\sqrt{S_k^2}\,|\xi|-d\right)\right]^2}{\left[S_k^2\left(S_k^2\xi^2-2d\,|\xi|\sqrt{S_k^2}+d^2\right)\right]^2}2S_k^4}. \tag{27}$$

Therefore, given the value of $\alpha_S$ and the maximum allowable sample size $n_0$, the test is performed by computing, for $k$=2,3,…, $n_0$, the statistic

$$\begin{aligned}W_k^{*(1)}&=\sqrt{k/n_0\,W_k^*}\\&=\sqrt{k/n_0}\sqrt{\frac{k\left(\ln\left[\left(d\big/\sqrt{S_k^2}-|\xi|\right)^2\Big/9\left(1+\xi^2\right)C_{pmk0}^2\right]\right)^2}{\dfrac{\left[d\left(\sqrt{S_k^2}\,|\xi|-d\right)\right]^2}{\left[S_k^2\left(S_k^2\xi^2-2d\,|\xi|\sqrt{S_k^2}+d^2\right)\right]^2}2S_k^4}}.\end{aligned} \tag{28}$$

Let $n_{stop}$ be the first integer $k$=2,3,…, $n_0$ for which $W_k^{*(1)}>w_{\alpha_S}$:

- We reject $H_0$ if $W_k^{*(1)}>w_{\alpha_S}$ and $n_{stop}$ is the stopping sample size of the test;
- We do not reject $H_0$ if $W_k^{*(1)}$ does not exceed $w_{\alpha_S}$ by $n_0$ and $n_0$ is the stopping sample size of the test;
- If $W_k^{*(1)}$ does not exceed $w_{\alpha_S}$ for $2\le k<n_0$, no decision can yet be made in favour of either of the two hypotheses and the decision is taken to continue observation.

In the case at hand, where $q$=1, the critical value $w_{\alpha_S}$ is obtained from the distribution of $\sup_{0<t<1}\left|B(t)\right|$. Specifically, $w_{\alpha_S}$ is such that (Feller, 1970):

$$1-\alpha_S=\frac{4}{\pi}\sum_{k=0}^{\infty}\frac{(-1)^k}{2k+1}\exp\left(\frac{-(2k+1)^2\pi^2}{8w_\alpha^2}\right). \qquad (29)$$

As an example, for $\alpha_S$ = 0.02, 0.05, 0.1 and 0.2, the values of $w_{\alpha_S}$ are 2.576, 2.241, 1.96 and 1.645 respectively.

In real applications, the parameter $\xi=(\mu-T)/\sigma$ is unknown and has to be estimated by substituting $\mu$ and $\sigma$ with their estimators, i.e. $\hat{\xi}=\left(\bar{X}_k-T\right)/S_k$, and the statistics $W_k^{*(1)}$ should be calculated by substituting $\xi$ with $\hat{\xi}$. We are aware that such an approach introduces a certain amount of variability. Therefore, for the sake of completeness, we will analyse the effects of $\xi$ on the finite sample behaviour of the sequential test statistic in Section 7.

## 5. A sequential sampling plan for variables based on $C_{pmk}$

In order to compare the statistical properties of the sequential test with those of the PL-test, a suitable approach is to study the power functions of the tests under equal conditions.

Note that the sequential test is two sided with composite alternative hypothesis $H_1: C_{pmk}\neq C_{pmk0}$, while the PL-test is unilateral. Thus, to correctly compare the two tests, cases under $H_1$ should be taken into consideration where $C_{pmk}=C_{pmk1}$ with $C_{pmk1}>C_{pmk0}$. In this manner, the sequential bilateral test with Type I error probability $\alpha_S$ can be compared with the non-sequential unilateral PL-test with Type I error probability equal to $\alpha_{PL}=\alpha_S/2$.

Furthermore, as far as the sequential test is concerned, it is also important to consider:

- The role of the maximum allowable sample size $n_0$, since it affects the statistical properties of the sequential test both under $H_0$ and $H_1$;
- The behaviour of the stopping sample size $n_{stop}$ required by the sequential test in order to decide in favour of $H_0$ and $H_1$.

This complex analysis can be developed by studying the use of the proposed sequential test in the framework of a variables sampling plan based on $C_{pmk}$. In such a way we can also propose a Sequential Sampling Plan (S-SP) based on $C_{pmk}$.

Wu and Pearn (2008) used several results from Pearn and Lin (2002) for developing a variables sampling plan (WP-SP) based on $C_{pmk}$. Their approach can be summarised as follows.

The acceptance sampling plan uses $C_{pmk}$ as benchmark for lot sentencing. The criterion used for measuring the performance of the sampling plan is the operating characteristic (OC) curve which quantifies the risks for producers and consumers. In the WP-SP framework, the OC curve plots the probability of lot acceptance versus the values of the $C_{pmk}$ index.

The acceptable quality level (AQL) and the lot tolerance percent defective (LTPD) are both defined in terms of the $C_{pmk}$ index: $C_{AQL}$ and $C_{LTPD}$ respectively.

The WP-SP is defined by the sample size *n* and the critical acceptance value $C_0$, and is such that for $C_{pmk} = C_{AQL}$ the producer's risk is $\alpha_{WP-SP}$ and for $C_{pmk} = C_{LTPD}$ the consumer's risk is $\beta_{WP-SP}$. Here, the producer's risk $\alpha_{WP-SP}$ is the probability of rejecting lots with $C_{pmk} \geq C_{AQL}$, and the consumer's risk $\beta_{WP-SP}$ is the probability of accepting lots with $C_{pmk} \leq C_{LTPD}$. In other words, the two points of interest on the OC curve are $\left(C_{AQL}, 1-\alpha_{WP-SP}\right)$ and $\left(C_{LTPD}, \beta_{WP-SP}\right)$.

Based on a sample of size *n*, a value of $\hat{C}_{pmk}$ (6) is calculated: if $\hat{C}_{pmk} > C_0$ the lot is accepted, otherwise the lot is rejected.

This sampling plan is equivalent to a PL-test with sample size *n* and critical value $C_0$ for testing $H_0 : C_{pmk} \leq C_{LTPD}$ versus $H_1 : C_{pmk} > C_{LTPD}$. With these parameters (*n* and $C_0$), when $C_{pmk} = C_{LTPD}$ the first error probability of the PL-test is $\alpha_{PL} = \beta_{WP-SP}$, and for $C_{pmk} = C_{AQL}$ the power function of the PL-test (9) is $\pi_{PL}\left(C_{AQL}\right) = \Pr\left(\hat{C}_{pmk} \geq C_0 \middle| C_{pmk} = C_{AQL}\right) = 1-\alpha_{WP-SP}$.

Wu and Pearn (2008), assuming $\xi$ as known and equal to $\xi = 0.5$, computed and tabulated the critical acceptance values $C_0$ and the sample sizes *n* of the sampling plan for several values of $\alpha_{WP-SP}$ and $\beta_{WP-SP}$, with various benchmarking quality levels $\left(C_{AQL}, C_{LTPD}\right)$=(1.33, 1.00), (1.50, 1.00), (1.50, 1.33), (1.67, 1.33), (1.67, 1.50), (2.00, 1.67).

Given the equivalence, explained above, between the PL-test and the sequential test, we are now able to propose a sequential sampling plan (S-SP) equivalent and comparable to the WP-SP.

The S-SP based on the statistic $W_k^{*(1)}$ (28) can be set up with a critical value $w_{\beta_{S-SP}}$ with $\beta_{S-SP} = \alpha_S = 2\alpha_{PL}$, and is performed as follows.

Let $n_{stop}$ be the first integer $k$=2,3,…, $n_0$ for which $W_k^{*(1)} > w_{\beta_{S-SP}}$ :

- We accept the lot (we reject $H_0 : C_{pmk} = C_{LTPD}$ ) if $W_{n_{stop}}^{*(1)} > w_{\beta_{S-SP}}$ and $n_{stop}$ is the stopping sample size of the sampling plan;
- We reject the lot (we do not reject $H_0$ ) if $W_k^{*(1)}$ does not exceed $w_{\beta_{S-SP}}$ by $n_0$ ;
- If $W_k^{*(1)}$ does not exceed $w_{\beta_{S-SP}}$ for $2 \le k < n_0$ , no decision can be made on the lot and the decision is taken to continue the inspection.

## 6. The Simulation Study

To properly compare the properties of the sequential test with those of the PL-test, the performance of the proposed sequential method should be evaluated in a worst-case scenario. However, we need an initial benchmark since, at this stage of the research, the statistical properties of the sequential procedure are unknown.

To this end, the properties of the sequential test under $H_0$ and $H_1$, or equivalently the properties of the S-SP, are studied under several scenarios by means of simulation studies (details on how the scenarios were constructed are given in Appendix B). Please note that in all the simulations that follow, we set $\xi = 0.5$ for the sequential test. In such a way, with regard to this aspect, we use the same setting as the PL-test. Hence, on the basis of the results obtained a further simulation study will be carried out to identify a value for $\xi$ that can represent the most unfavourable condition for the sequential procedure.

As a first step, we study the performance of the sequential sampling plan in connection with the consumer's risk, or equivalently the performance of the sequential test under $H_0 : C_{pmk} = C_{LTPD}$ .

With this aim, for each combination of $C_{LTPD}$ (1.00, 1.33, 1.50, 1.67), and $\beta_{S-SP}$ (0.02, 0.05, 0.1) we generated, using *R* (*R* Core Team 2020), $5 \cdot 10^4$ replicates from a normally distributed process. The aim of these simulations was to determine the smallest maximum allowable sample size, $n_{0;\hat{\beta}_{S-SP} < \beta_{S-SP}}$ , which gives an empirical consumer's risk $\hat{\beta}_{S-SP}$ smaller than the nominal $\beta_{S-SP}$ value: $\hat{\beta}_{S-SP} < \beta_{S-SP}$. Or, equivalently, to determine the smallest maximum allowable sample size, $n_{0;\hat{\alpha}_S < \alpha_S}$ , which gives an empirical type I error probability $\hat{\alpha}_S$ smaller than the nominal $\alpha_S$ value: $\hat{\alpha}_S < \alpha_S$ . The empirical consumer's risk is estimated as the proportion of accepted lots when $C_{pmk} = C_{LTPD}$ (the proportion of correctly accepted $H_0 : C_{pmk} = C_{LTPD}$ ).

In greater detail, in order to obtain $n_{0;\hat{\beta}_{S-SP}<\beta_{S-SP}}$, we implemented an iterative search algorithm with initial value for $n_{0;\hat{\beta}_{S-SP}<\beta_{S-SP}}$ denoted by $n_{start}$. The algorithm works as follows:

1. With $n_{start}$ as the maximum allowable sample size of the sequential test, the value of $\hat{\beta}_{S-SP}$ is estimated as the fraction of correctly accepted $H_0 : C_{pmk} = C_{LTPD}$ over $m = 5 \cdot 10^4$ simulations;
2. If

$$0 < \frac{\beta_{S-SP} - \hat{\beta}_{S-SP}}{\beta_{S-SP}} \leq 0.025,$$

then $n_{0;\hat{\beta}_{S-SP}<\beta_{S-SP}}$ is set equal to $n_{start}$ and the search algorithm stops;

3. Otherwise, if $\hat{\beta}_{S-SP} > \beta_{S-SP}$, then $n_{start} = n_{start} + 1$; if $\hat{\beta}_{S-SP} < \beta_{S-SP}$, then $n_{start} = n_{start} - 1$ and the algorithm starts a further *m* simulations.

The simulation results are summarized in Table 1 (rows with white background), where for each combination of $\beta_{S-SP}$ and $C_{LTPD}$, the following quantities are given:

- $n_{0;\hat{\beta}_{S-SP}<\beta_{S-SP}}$ the smallest maximum allowable sample size for sequential sampling plan for achieving an empirical consumer's risk $\hat{\beta}_{S-SP}$ smaller than the nominal value $\beta_{S-SP}$ $\left(\hat{\beta}_{S-SP} < \beta_{S-SP}\right)$; or equivalently, the smallest maximum allowable sample size for the sequential test for achieving an empirical type I error probability $\hat{\alpha}_S$ smaller than the nominal $\alpha_S$ value $\left(\hat{\alpha}_S < \alpha_S\right)$;
- $\hat{\beta}_{S-SP}$ the empirical consumer's risk.

| $C_{LTPD}$ | $\beta_{S-SP}=\mathbf{0.1}$ | | $\beta_{S-SP}=\mathbf{0.05}$ | | $\beta_{S-SP}=\mathbf{0.02}$ | |
|---|---|---|---|---|---|---|
| | $n_{0;\hat{\beta}_{S-SP}<\beta_{S-SP}}$ | $\hat{\beta}_{S-SP}$ | $n_{0;\hat{\beta}_{S-SP}<\beta_{S-SP}}$ | $\hat{\beta}_{S-SP}$ | $n_{0;\hat{\beta}_{S-SP}<\beta_{S-SP}}$ | $\hat{\beta}_{S-SP}$ |
| **1.00** | 158 | 0.0995 | 200 | 0.0487 | 331 | 0.0194 |
| | 172 | 0.0998 | 242 | 0.0495 | 359 | 0.0193 |
| **1.33** | 149 | 0.0987 | 198 | 0.0497 | 296 | 0.0199 |
| | 161 | 0.0992 | 232 | 0.0499 | 324 | 0.0199 |
| **1.50** | 153 | 0.0996 | 213 | 0.0492 | 319 | 0.0198 |
| | 175 | 0.0993 | 240 | 0.0499 | 320 | 0.0199 |
| **1.67** | 147 | 0.0995 | 197 | 0.0497 | 258 | 0.0199 |
| | 157 | 0.0999 | 219 | 0.0498 | 288 | 0.0197 |

**Table 1.** Simulation results for $C_{pmk}=C_{LTPD}$: $n_{0;\hat{\beta}_{S-SP}<\beta_{S-SP}}$ and $\hat{\beta}_{S-SP}$ (white background for $\xi=0.5$; shaded background for $\xi=3$)

The R code for reproducing the simulation study for the case $C_{LTPD}=1.50$ and $\beta_{S-SP}=0.05$ is provided as Supplementary Information (SH0_150.R for $\xi=0.5$ and SH0_150N3.R for $\xi=3$).

As a second step we study the statistical properties of the S-SP in connection with the producer's risk, or equivalently the statistical properties of the sequential test under $H_1\ \left(C_{pmk}=C_{AQL}\right)$.

With this aim, for the benchmarking quality levels $\left(C_{AQL},C_{LTPD}\right)$=(1.33, 1.00), (1.50, 1.00), (1.50, 1.33), (1.67, 1.33), (1.67, 1.50), (2.00, 1.67) and for producer's risk $\alpha_{S-SP}=\alpha_{WP-SP}$=(0.01, 0.025, 0.05), we looked for the maximum allowable sample size $n_0$ required by the S-SP that for $C_{pmk}=C_{AQL}$ allows to achieving the same producer's risk as the WP-SP. More precisely for each value of $\alpha_{S-SP}$, $\beta_{S-SP}$, $C_{AQL}$, $C_{LTPD}$ we used a simulation experiment to determine the smallest maximum allowable sample size, $n_{0;\hat{\pi}_S>1-\alpha_{S-SP}}$, which gives an empirical power of the sequential test $\hat{\pi}_S$ greater than $1-\alpha_{S-SP}$, or equivalently to obtain an empirical producer's risk $\leq\alpha_{S-SP}$. The empirical power $\hat{\pi}_S$ of the sequential test is estimated as the proportion of correctly rejected $H_0: C_{pmk} = C_{LTPD}$ when $C_{pmk} = C_{AQL}$.

In order to obtain $n_{0;\hat{\pi}_S > 1-\alpha_{S-SP}}$ we implemented an iterative search algorithm with an initial value for $n_{0;\hat{\pi}_S > 1-\alpha_{S-SP}}$ denoted by $n_{start}$. The algorithm works as follows:

1. With $n_{start}$ as the maximum allowable sample size of the sequential test, the empirical power of the test $\hat{\pi}_S$ is estimated as the proportion of correctly rejected $H_0$ over $m = 5 \cdot 10^4$ simulations;
2. If

$$0 < \frac{\hat{\pi}_S - (1-\alpha_{S-SP})}{1-\alpha_{S-SP}} \leq 0.025,$$

then $n_{0;\hat{\pi}_S > 1-\alpha_{S-SP}}$ is set equal to $n_{start}$ and the search algorithm stops. At the same time the average stopping sample size $n_{avg}$ was empirically assessed as the average of the stopping sample sizes $n_{stop}$ required by the sequential test to correctly reject $H_0$ (to accept the lot when $C_{pmk} = C_{AQL}$) when the maximum allowable sample size is equal to $n_{0;\hat{\pi}_S > 1-\alpha_{S-SP}}$;
3. Otherwise, if $\hat{\pi}_S \leq 1-\alpha_{S-SP}$, then $n_{start} = n_{start} + 1$; if $\hat{\pi}_S > 1-\alpha_{S-SP}$, then $n_{start} = n_{start} - 1$ and the algorithm starts other *m* simulations.

The simulation results are summarized in Tables 2-7 (rows with white background). For each combination of $\alpha_{S-SP}$, $\beta_{S-SP}$ $C_{AQL}$ and $C_{LTPD}$, the following quantities are given: $n_{0;\hat{\pi}_S > 1-\alpha_{S-SP}}$ the smallest maximum allowable sample size for the sequential test for achieving an empirical power $\hat{\pi}_S > 1-\alpha_{S-SP}$ (or equivalently an empirical producer's risk $\hat{\alpha}_{S-SP}$ smaller than or equal to the nominal value $\alpha_{S-SP}$); $n_{avg}$ the average of the stopping sample sizes $n_{stop}$ required for the sequential test with maximum allowable sample size $n_{0;\hat{\pi}_S > 1-\alpha_{S-SP}}$ for concluding in favour of $H_1$; S.D.($n_{stop}$) the standard deviation of the final sample sizes $n_{stop}$; $\hat{\alpha}_{S-SP} = 1-\hat{\pi}_S$ the estimated producer's risk. In order to allow comparisons, the parameters of the corresponding WP-SP (Wu and Pearn 2008) are also given in Tables 2-7.

The R code for reproducing the simulation study for the case $C_{AQL} = 1.50$, $C_{LTPD} = 1.33$, $\alpha_{WP-SP} = 0.01$ and $\beta_{WP-SP} = 0.01$ is provided as Supplementary Information (PS133_150_002.R for $\xi = 0.5$, and PS133_150_002N3.R for $\xi = 0.3$).

| $C_{AQL} = 1.33$ $C_{LTPD} = 1.00$ | | | | | | |
|---|---|---|---|---|---|---|
| $\alpha_{WP-SP}$ | $\beta_{WP-SP}$ | $n$ | $C_0$ | $n_{0;\hat{\pi}_S > 1-\alpha_{S-SP}}$ | $n_{avg}$(SD) | $\hat{\alpha}_{S-SP}$ |
| 0.010 | 0.010 | 202 | 1.1634 | 165 | 87.31(25.52) | 0.0097 |
| | | | | 209 | 109.28(32.24) | 0.0093 |
| | 0.025 | 170 | 1.1497 | 139 | 68.72(22.44) | 0.0093 |
| | | | | 176 | 86.56(28.69) | 0.0094 |
| | 0.050 | 144 | 1.1360 | 120 | 55.10(20.19) | 0.0097 |
| | | | | 153 | 69.36(25.60) | 0.0094 |
| 0.025 | 0.010 | 174 | 1.1779 | 141 | 79.38(23.08) | 0.0243 |
| | | | | 176 | 98.73(29.24) | 0.0248 |
| | 0.025 | 144 | 1.1642 | 117 | 62.12(24.43) | 0.0238 |
| | | | | 148 | 77.56(25.99) | 0.0237 |
| | 0.050 | 120 | 1.1504 | 98 | 48.49(18.05) | 0.0243 |
| | | | | 125 | 61.22(23.08) | 0.0239 |
| 0.050 | 0.010 | 151 | 1.1925 | 119 | 71.15(20.52) | 0.0498 |
| | | | | 152 | 89.89(26.43) | 0.0496 |
| | 0.025 | 123 | 1.1792 | 98 | 55.06(18.13) | 0.0492 |
| | | | | 125 | 69.35(23.04) | 0.0479 |
| | 0.050 | 102 | 1.1654 | 81 | 42.40(16.06) | 0.0496 |
| | | | | 104 | 53.79(20.34) | 0.0478 |

Table 2. WP-SP ($n$ and $C_0$) and S-SP simulation results ($n_{0;\hat{\pi}_S > 1-\alpha_{S-SP}}$, $n_{avg}$ and SD) for $C_{AQL} = 1.33$ and $C_{LTPD} = 1.00$. (white background for $\xi = 0.5$; shaded background for $\xi = 3$)

| | | | $C_{AQL}=1.50$ | $C_{LTPD}=1.00$ | | |
|---|---|---|---|---|---|---|
| $\alpha_{WP-SP}$ | $\beta_{WP-SP}$ | $n$ | $C_0$ | $n_{0;\hat{\pi}_S>1-\alpha_{S-SP}}$ | $n_{avg}$(SD) | $\hat{\alpha}_{S-SP}$ |
| 0.010 | 0.010 | 98 | 1.2466 | 78 | 41.04(12.41) | 0.0098 |
| | | | | 98 | 50.54(15.50) | 0.0091 |
| | 0.025 | 82 | 1.2261 | 66 | 32.15(11.17) | 0.0089 |
| | | | | 83 | 39.88(13.98) | 0.0096 |
| | 0.050 | 70 | 1.2057 | 56 | 25.12(10.22) | 0.0092 |
| | | | | 70 | 31.17(12.63) | 0.0099 |
| 0.025 | 0.010 | 85 | 1.2688 | 66 | 37.01(11.33) | 0.0246 |
| | | | | 82 | 45.54(14.23) | 0.0248 |
| | 0.025 | 70 | 1.2484 | 55 | 28.58(10.25) | 0.0243 |
| | | | | 68 | 35.22(12.56) | 0.0247 |
| | 0.050 | 58 | 1.2279 | 46 | 21.78(9.27) | 0.0242 |
| | | | | 58 | 27.44(11.50) | 0.0242 |
| 0.050 | 0.010 | 74 | 1.2913 | 56 | 33.07(10.20) | 0.0499 |
| | | | | 70 | 40.90(12.80) | 0.0491 |
| | 0.025 | 60 | 1.2714 | 46 | 25.15(9.28) | 0.0464 |
| | | | | 57 | 30.93(11.32) | 0.0494 |
| | 0.050 | 50 | 1.2511 | 38 | 18.89(8.37) | 0.0463 |
| | | | | 47 | 23.34(10.14) | 0.0499 |

Table 3. WP-SP ($n$ and $C_0$) and S-SP simulation results ($n_{0;\hat{\pi}_S>1-\alpha_{S-SP}}$, $n_{avg}$ and SD) for $C_{AQL}=1.50$ and $C_{LTPD}=1.00$ (white background for $\xi=0.5$; shaded background for $\xi=3$).

| $C_{AQL}=1.50$ $C_{LTPD}=1.33$ | | | | | | |
|---|---|---|---|---|---|---|
| $\alpha_{WP-SP}$ | $\beta_{WP-SP}$ | $n$ | $C_0$ | $n_{0;\hat{\pi}_S>1-\alpha_{S-SP}}$ | $n_{avg}$(SD) | $\hat{\alpha}_{S-SP}$ |
| 0.010 | 0.010 | 1039 | 1.4147 | 925 | 493.71(136.34) | 0.0098 |
| | | | | 1116 | 597.40(167.28) | 0.0098 |
| | 0.025 | 877 | 1.4075 | 790 | 394.41(122.54) | 0.0099 |
| | | | | 963 | 480.48(149.80) | 0.0098 |
| | 0.050 | 749 | 1.4003 | 678 | 318.17(109.64) | 0.0096 |
| | | | | 819 | 386.17(133.97) | 0.0099 |
| 0.025 | 0.010 | 887 | 1.4220 | 782 | 447.67(123.85) | 0.0235 |
| | | | | 957 | 544.40(152.62) | 0.0249 |
| | 0.025 | 738 | 1.4149 | 658 | 354.10(109.53) | 0.0249 |
| | | | | 805 | 431.19(134.42) | 0.0248 |
| | 0.050 | 621 | 1.4076 | 561 | 283.34(97.74) | 0.0248 |
| | | | | 688 | 346.79(120.44) | 0.0245 |
| 0.050 | 0.010 | 765 | 1.4295 | 675 | 408.00(111.02) | 0.0494 |
| | | | | 819 | 493.67(135.81) | 0.0498 |
| | 0.025 | 627 | 1.4224 | 563 | 319.47(98.01) | 0.0469 |
| | | | | 678 | 386.57(118.41) | 0.0498 |
| | 0.050 | 520 | 1.4152 | 472 | 253.30(86.21) | 0.0485 |
| | | | | 571 | 306.21(105.58) | 0.0494 |

Table 4. WP-SP ($n$ and $C_0$) and S-SP simulation results ($n_{0;\hat{\pi}_S>1-\alpha_{S-SP}}$, $n_{avg}$ and SD) for $C_{AQL}=1.50$ and $C_{LTPD}=1.33$ (white background for $\xi=0.5$; shaded background for $\xi=3$).

| $C_{AQL} = 1.67$ $C_{LTPD} = 1.33$ | | | | | | |
|---|---|---|---|---|---|---|
| $\alpha_{WP-SP}$ | $\beta_{WP-SP}$ | $n$ | $C_0$ | $n_{0;\hat{\pi}_S > 1-\alpha_{S-SP}}$ | $n_{avg}$(SD) | $\hat{\alpha}_{S-SP}$ |
| 0.010 | 0.010 | 286 | 1.4988 | 250 | 133.04(37.91) | 0.0092 |
| | | | | 301 | 159.36(46.00) | 0.0097 |
| | 0.025 | 240 | 1.4846 | 212 | 105.33(33.76) | 0.0095 |
| | | | | 254 | 125.99(40.76) | 0.0097 |
| | 0.050 | 204 | 1.4704 | 182 | 84.51(30.15) | 0.0092 |
| | | | | 221 | 101.70(36.42) | 0.0099 |
| 0.025 | 0.010 | 245 | 1.5137 | 212 | 120.18(34.38) | 0.0240 |
| | | | | 253 | 143.61(41.25) | 0.0249 |
| | 0.025 | 203 | 1.4995 | 177 | 94.40(30.32) | 0.0241 |
| | | | | 214 | 113.43(36.79) | 0.0244 |
| | 0.050 | 170 | 1.4852 | 151 | 75.25(27.11) | 0.0235 |
| | | | | 180 | 89.77(32.33) | 0.0244 |
| 0.050 | 0.010 | 213 | 1.5287 | 183 | 109.31(30.92) | 0.0493 |
| | | | | 218 | 130.51(37.20) | 0.0495 |
| | 0.025 | 174 | 1.5149 | 149 | 84.32(26.91) | 0.0486 |
| | | | | 179 | 101.01(32.35) | 0.0499 |
| | 0.050 | 143 | 1.5006 | 122 | 64.95(23.23) | 0.0498 |
| | | | | 150 | 79.22(28.63) | 0.0494 |

Table 5. WP-SP ($n$ and $C_0$) and S-SP simulation results ($n_{0;\hat{\pi}_S > 1-\alpha_{S-SP}}$, $n_{avg}$ and SD) for $C_{AQL} = 1.67$ and $C_{LTPD} = 1.33$ (white background for $\xi = 0.5$; shaded background for $\xi = 3$).

| | | | $C_{AQL} = 1.67$ $\quad C_{LTPD} = 1.50$ | | | |
|---|---|---|---|---|---|---|
| $\alpha_{WP-SP}$ | $\beta_{WP-SP}$ | $n$ | $C_0$ | $n_{0;\hat{\pi}_S > 1-\alpha_{S-SP}}$ | $n_{avg}$(SD) | $\hat{\alpha}_{S-SP}$ |
| 0.010 | 0.010 | 1253 | 1.5847 | 1160 | 615.52(170.65) | 0.0080 |
| | | | | 1359 | 726.96(203.51) | 0.0090 |
| | 0.025 | 1058 | 1.5775 | 975 | 489.08(151.33) | 0.0094 |
| | | | | 1158 | 581.11(180.28) | 0.0094 |
| | 0.050 | 904 | 1.5703 | 842 | 394.67(135.05) | 0.0097 |
| | | | | 1004 | 471.38(161.65) | 0.0097 |
| 0.025 | 0.010 | 1068 | 1.5921 | 966 | 552.49(152.79) | 0.0231 |
| | | | | 1161 | 662.10(184.15) | 0.0234 |
| | 0.025 | 890 | 1.5849 | 813 | 438.64(135.03) | 0.0249 |
| | | | | 970 | 523.44(162.42) | 0.0248 |
| | 0.050 | 749 | 1.5776 | 690 | 349.80(119.41) | 0.0247 |
| | | | | 827 | 417.75(143.83) | 0.0249 |
| 0.050 | 0.010 | 922 | 1.5995 | 840 | 507.27(137.42) | 0.0471 |
| | | | | 993 | 598.93(164.08) | 0.0491 |
| | 0.025 | 756 | 15924 | 686 | 393.65(119.22) | 0.0498 |
| | | | | 820 | 467.16(143.15) | 0.0493 |
| | 0.050 | 627 | 1.5852 | 574 | 310.70(105.14) | 0.0489 |
| | | | | 690 | 370.31(125.80) | 0.0497 |

Table 6. WP-SP ($n$ and $C_0$) and S-SP simulation results ($n_{0;\hat{\pi}_S > 1-\alpha_{S-SP}}$, $n_{avg}$ and SD) for $C_{AQL} = 1.67$ and $C_{LTPD} = 1.50$ (white background for $\xi = 0.5$; shaded background for $\xi = 3$).

| $C_{AQL} = 2.00 \quad C_{LTPD} = 1.67$ | | | | | | |
|---|---|---|---|---|---|---|
| $\alpha_{WP-SP}$ | $\beta_{WP-SP}$ | $n$ | $C_0$ | $n_{0;\hat{\pi}_S > 1-\alpha_{S-SP}}$ | $n_{avg}$(SD) | $\hat{\alpha}_{S-SP}$ |
| 0.010 | 0.010 | 426 | 1.8341 | 388 | 206.69(58.18) | 0.0097 |
| | | | | 453 | 241.81(68.06) | 0.0084 |
| | 0.025 | 359 | 1.8203 | 331 | 165.06(51.71) | 0.0097 |
| | | | | 388 | 192.88(61.02) | 0.0095 |
| | 0.050 | 305 | 1.8065 | 287 | 133.58(46.66) | 0.0099 |
| | | | | 333 | 154.68(54.50) | 0.0096 |
| 0.025 | 0.010 | 365 | 1.8485 | 331 | 188.48(52.83) | 0.0247 |
| | | | | 383 | 218.14(61.57) | 0.0239 |
| | 0.025 | 303 | 1.8347 | 276 | 148.07(46.71) | 0.0243 |
| | | | | 323 | 172.43(64.72) | 0.0236 |
| | 0.050 | 254 | 1.8207 | 233 | 117.29(41.25) | 0.0246 |
| | | | | 275 | 137.70(48.47) | 0.0247 |
| 0.050 | 0.010 | 316 | 1.8630 | 281 | 169.82(46.92) | 0.0493 |
| | | | | 331 | 198.42(55.42) | 0.0489 |
| | 0.025 | 258 | 1.8495 | 233 | 148.07(41.33) | 0.0497 |
| | | | | 274 | 155.14(48.74) | 0.0491 |
| | 0.050 | 213 | 1.8356 | 195 | 104.19(36.37) | 0.0497 |
| | | | | 228 | 121.36(42.71) | 0.0497 |

Table 7: WP-SP ($n$ and $C_0$) and S-SP simulation results ($n_{0;\hat{\pi}_S > 1-\alpha_{S-SP}}$, $n_{avg}$ and SD) for $C_{AQL} = 2.00$ and $C_{LTPD} = 1.67$ (white background for $\xi = 0.5$; shaded background for $\xi = 3$).

## 7 Discussion

Before drawing any conclusions, it is important to deal with the relevant issue introduced in the previous Section. For the purposes of comparison with the PL-test, the results for the sequential test (Tables 1-7 rows with white background) were obtained by assuming $\xi$ as being known and equal to 0.5.

However, in real-world applications, $\xi$ is unknown. Furthermore, for the PL-test and the WP-SP by setting $\xi = 0.5$ conservative sample size and critical values were obtained, whereas the effect of $\xi$ on the statistical properties of the sequential test is unknown.

Therefore, in order to evaluate the impact of $\xi$ on the sequential test, or equivalently on the sequential sampling plan, and identify a value of $\xi$ that can reasonably represent the worst-case scenario, we conducted further simulation studies.

For each combination of $\alpha_S$, $C_{pmk0}\left(C_{LTPD}\right)$, we considered values of $\xi$ ranging from 0 to 5 $\left(\xi = 0, 0.5, 1.0, 2.0, 3.0, 5.0\right)$, and we simulated $10^4$ observations from processes with $C_{pmk} \geq C_{pmk0}$ (for example if $C_{pmk0} = 1.67$ we considered $C_{pmk} = 1.67(0.05)2.12$.

The aim of these simulations was to evaluate the empirical power $\hat{\pi}_S$ of the sequential test estimated as the proportion of rejected $H_0$. As an example, Figure 1 plots $\hat{\pi}_S$ for the case $C_{pmk0} = 1.67$ and $\alpha_S = 0.02$ versus the values of $C_{pmk} = 1.67(0.05)2.12$.

The curves are plotted for each value of $\xi$, with the addition of the bold curve for $\xi = 0.5$ and a dashed red curve for $\xi = 3$.

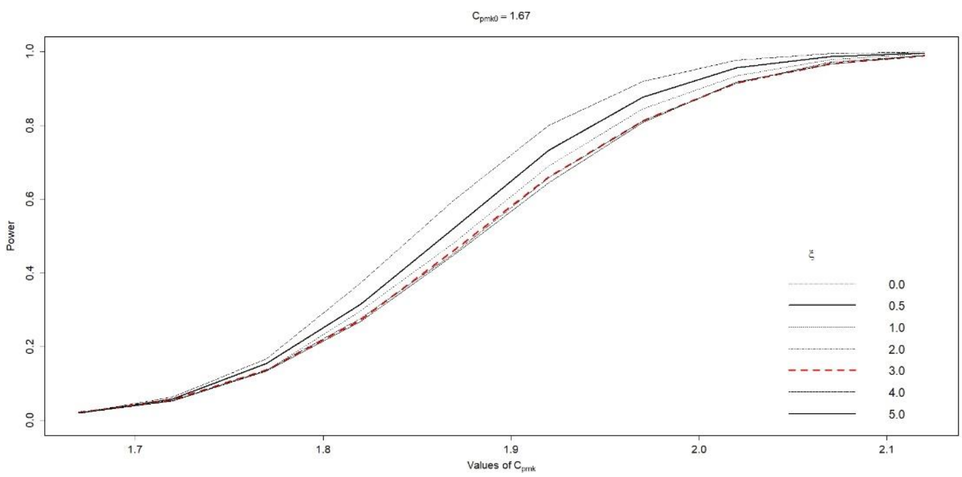


Figure 1. Empirical power functions for the case $\alpha_S = 0.02$ and $C_{pmk0} = 1.67$ for $\xi$ from 0 to 5

For completeness, in Figures 2-4 we reported the results for all the cases examined where for clearness purpose, we limit the curves plotted to $\xi = (0.5, 3.0, 4.0, 5.0)$.

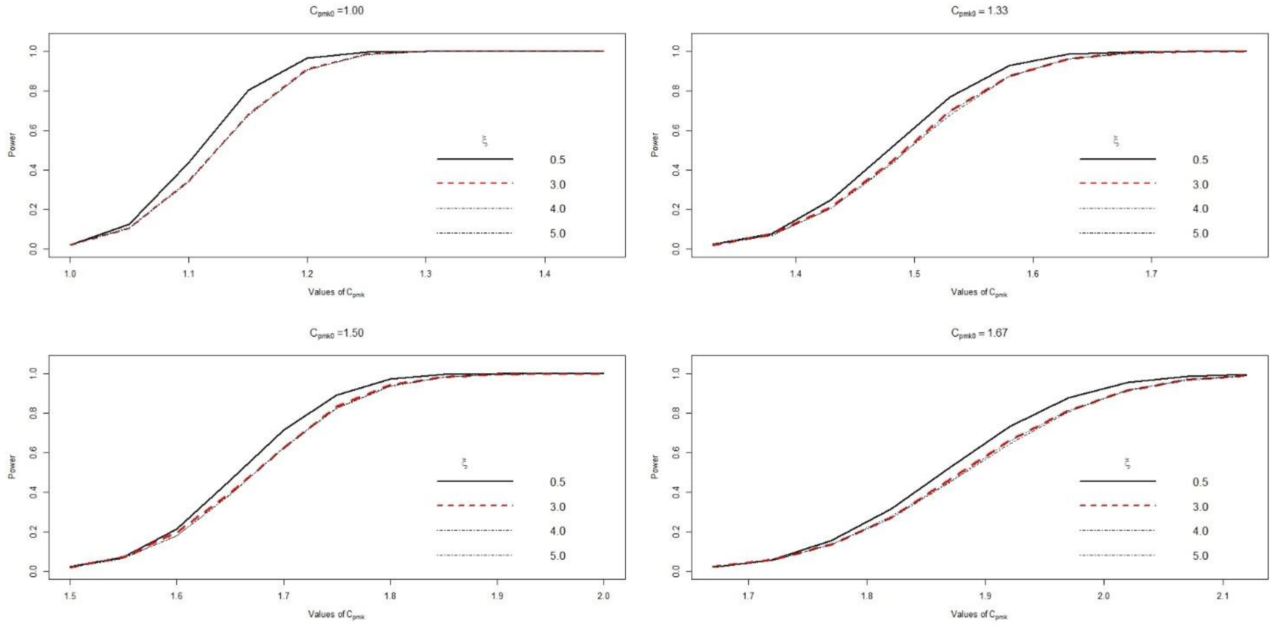


Figure 2. Empirical power functions for the case $\alpha_S = 0.02$ and $C_{pmk0} = (1.00, 1.33, 1.50, 1.67)$ for $\xi = 0.5, 3.0, 4.0, 5.0$

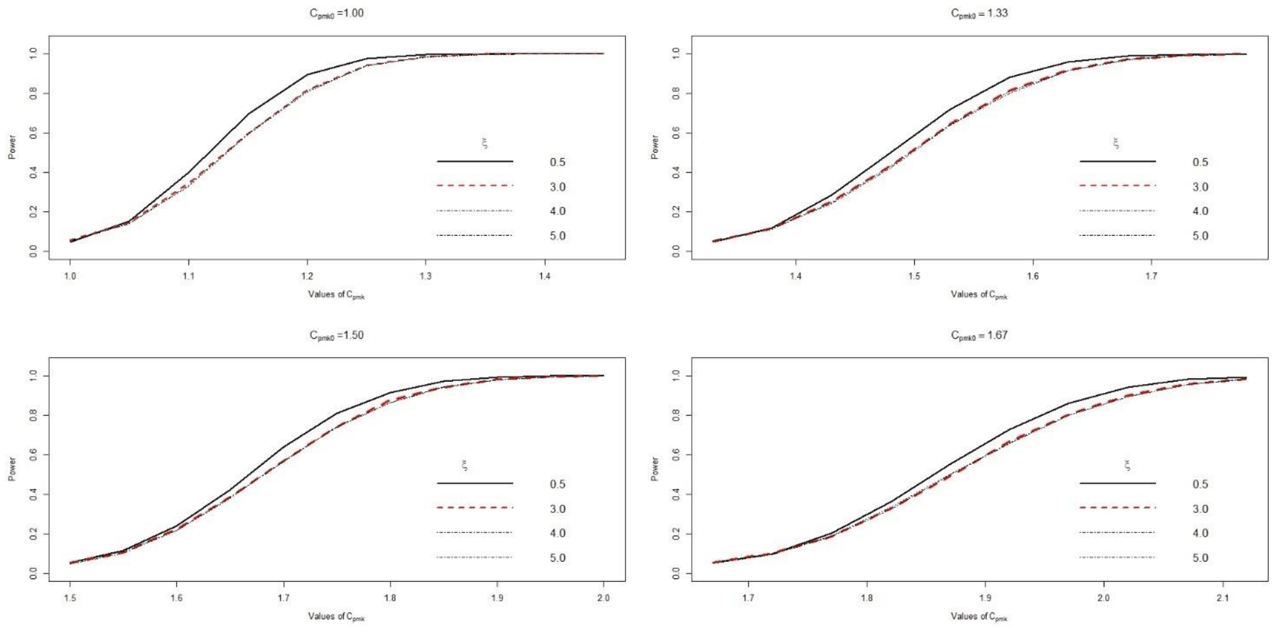


Figure 3. Empirical power functions for the case $\alpha_S = 0.05$ and $C_{pmk0} = (1.00, 1.33, 1.50, 1.67)$ for $\xi = (0.5, 3.0, 4.0, 5.0)$

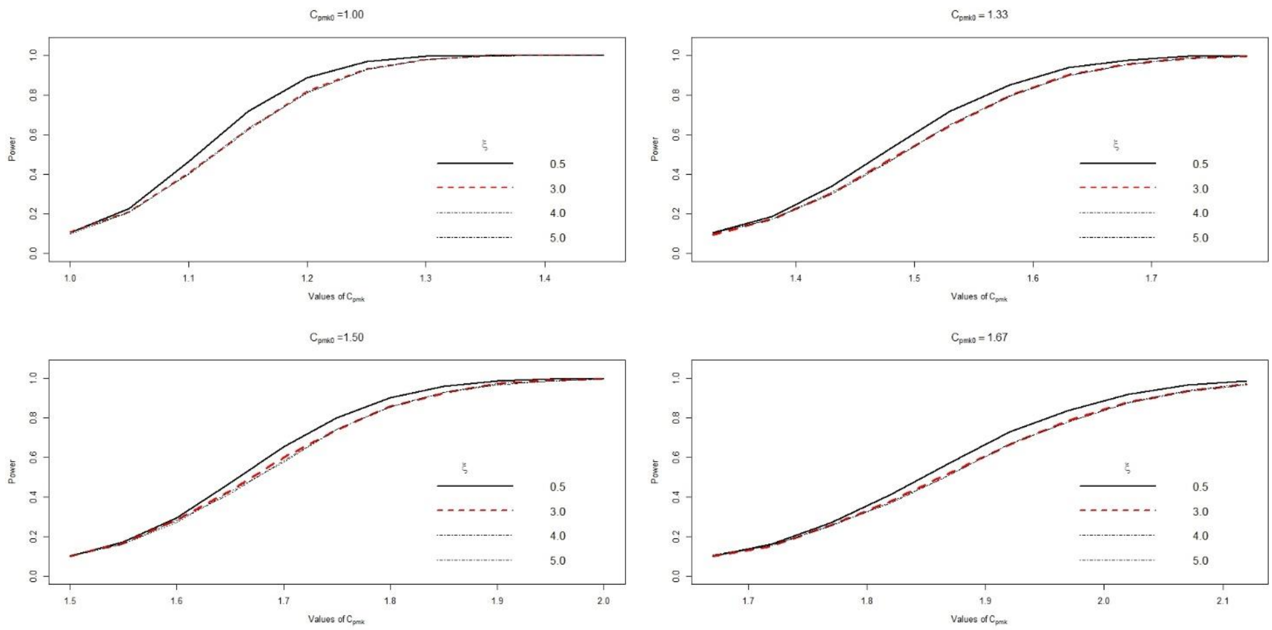


Figure 4. Empirical power functions for the case $\alpha_S = 0.1$ and $C_{pmk0} = (1.00, 1.33, 1.50, 1.67)$ for $\xi = (0.5, 3.0, 4.0, 5.0)$

By examining the results, it can be noted that the empirical power functions $\hat{\pi}_S$ decrease as $\xi$ increases and that the changes for $\xi \geq 3$ are negligible. Hence, considering also that in literature (Wu and Pearn, 2008, Lepore et al. 2017) the range usually considered to study the effects of $\xi$ is $0 \leq \xi \leq 3$ and that these values cover a wide range of real situations, the value $\xi = 3$ can reasonably represent a pessimistic situation for the sequential procedures.

Thus, to study the performance of the sequential sampling plan in connection with the consumer's risk, or equivalently the performance of the sequential test under $H_0 : C_{pmk} = C_{LTPD}$ in the worst-case scenario, we repeated with $\xi = 3$ the simulation experiment described in Section 6. In such a way, for each combination of $\beta_{S\text{-}SP}$ and $C_{LTPD}$, we obtain the values of $n_{0;\hat{\beta}_{S-SP} < \beta_{S-SP}}$ and $\hat{\beta}_{S-SP}$. The simulation results are summarized in Table 1 (rows with shaded background). Similarly, to study the statistical properties of the S-SP in connection with the producer's risk, or equivalently the statistical properties of the sequential test under $H_1$ $\left(C_{pmk} = C_{AQL}\right)$, in the worst-case scenario, we repeated with $\xi = 3$ the simulation experiment described di Section 6. The results concerning the values $n_{0;\hat{\pi}_S > 1-\alpha_{S-SP}}$, $n_{avg}$ and $\hat{\alpha}_{S-SP}$ are reported in Tables 2-7 (rows with shaded background).

Therefore, by examining Tables 2-7 it is possible to quantify the changes occurred in the performance of the sequential test, moving from a situation with $\xi = 0.5$ to a situation with $\xi = 3$ and the following discussion will be made by comparing the outcomes in the worst-case scenario of the sequential procedures with the performance of the PL test.

The results indicate that, when the capability requirements $C_{AQL}$ and $C_{LTPD}$ are high ($C_{LTPD} \geq 1.33$), the S-SP allows, on average, a considerable reduction in sample size compared to the WP-SP.

As an example, for $C_{LTPD}$=1.33 and $C_{AQL}$=1.50 with producer's risk $\alpha_{WP-SP}$=0.01 and consumer's risk $\beta_{WP-SP}$=0.01, the WP sampling plan requires a sample size of $n$=1039 with a critical value $C_0$=1.4147 (Table 4). The S-SP:

- When $C_{pmk} = C_{AQL}$=1.50, with a maximum allowable sample size equal to $n_{0;\hat{\pi}_S > 1-\alpha_{S-SP}}$=1116, the estimated producer's risk is equal to $\hat{\alpha}_{S-SP}$=0.0098 with $n_{avg} \cong$597 (Table 4);
- When $C_{pmk} = C_{LTPD}$=1.33 (Table 1), the smallest maximum allowable sample size required by the sequential test to ensure an estimated consumer's risk $\hat{\beta}_{S-SP}$ smaller than the nominal value $\beta_{S-SP}$ is $n_{0;\hat{\beta}_{S-SP} < \beta_{S-SP}}$=324.

Therefore, in this case the S-SP: for high-quality lots ($C_{pmk} = C_{AQL}$) allows, on average, a 42.54% reduction in sample size compared to the WP-SP ($n_{avg} \cong$597); for low-quality lots ($C_{pmk} = C_{LTPD}$) allows a 68.82% reduction in sample size compared to the WP-SP, since it rejects the lot using a sample size not greater than $n_{0;\hat{\beta}_{S-SP} < \beta_{S-SP}}$=324.

For $C_{LTPD}$=1.67 and $C_{AQL}$=2.00 with $\alpha_{WP-SP}$=0.025 and $\beta_{WP-SP}$=0.05, the WP sampling plan has a sample size $n$=254 with a critical value $C_0$=1.8207 (Table 7). The S-SP:

- When $C_{pmk} = C_{AQL}$=2.00, with a maximum allowable sample size equal to $n_{0;\hat{\pi}_S > 1-\alpha_{S-SP}}$=275, has an estimated producer's risk equal to $\hat{\alpha}_{S-SP}$=0.0247 with $n_{avg} \cong$137.70 (Table 7);
- When $C_{pmk} = C_{LTPD}$=1.67 (Table 1), the smallest maximum allowable sample size required by the sequential test to ensure an estimated consumer's risk $\hat{\beta}_{S-SP}$ smaller than the nominal value $\beta_{S-SP}$ is $n_{0;\hat{\beta}_{S-SP} < \beta_{S-SP}}$=157 (Table 1).

In this case, the S-SP: for high-quality lots ($C_{pmk} = C_{AQL}$) allows, on average, a 45.7% reduction in sample size compared to the WP-SP ($n_{avg} \cong$138); for low-quality lots ($C_{pmk} = C_{LTPD}$) allows a 38.2% reduction in sample size compared to the WP-SP, since it rejects the lot using a sample size not greater than $n_{0;\hat{\beta}_{S-SP} < \beta_{S-SP}}$=157.

However, for the sake of completeness, it is important to note that the sequential sampling plan, in order to ensure that the empirical consumer's risk does not exceed the nominal level $\beta_{S-SP}$, requires at a minimum the sequential inspection of a number of units not less than the values given in Table 1 (rows with shaded background).

By comparing the values of $n_{0;\hat{\beta}_{S-SP}<\beta_{S-SP}}$, given in Table 1, with the sample sizes of an equivalent WP-SP, we find that the sequential sampling plan can have lower performance than the WP-SP when the capability requirements are low ($C_{LTPD} < 1.33$).

As an example, for $C_{LTPD}$=1.00 and $C_{AQL}$=1.33 with $\alpha_{WP-SP}$=0.025 and $\beta_{WP-SP}$=0.025, the WP sampling plan has a sample size $n$=144 with a critical value $C_0$=1.1642 (Table 2).

The SSP:

- When $C_{pmk} = C_{AQL}$=1.33, with a maximum allowable sample size equal to $n_{0;\hat{\pi}_S>1-\alpha_{S-SP}}$ =148, has an estimated producer's risk equal to $\hat{\alpha}_{S-SP}$=0.0237 with $n_{avg} \cong$77.56 (Table 2);
- When $C_{pmk} = C_{LTPD}$=1.00, the smallest maximum allowable sample size required by the sequential test to ensure an estimated consumer's risk $\hat{\beta}_{S-SP}$ smaller than the nominal value $\beta_{S-SP}$ is $n_{0;\hat{\beta}_{S-SP}<\beta_{S-SP}}$ =242 (Table 1).

In this case $n_{0;\hat{\beta}_{S-SP}<\beta_{S-SP}}$ =242 is larger than $n_{0;\hat{\pi}_S>1-\alpha_{S-SP}}$ =148; consequently, to ensure that the both the risks $\hat{\beta}_{S-SP}$ and $\hat{\alpha}_{S-SP}$ do not exceed their nominal values, $\beta_{S-SP}$ and $\alpha_{S-SP}$ respectively, it is necessary to set up the sequential sampling plan using, as maximum sample size, the largest between $n_{0;\hat{\beta}_{S-SP}<\beta_{S-SP}}$ and $n_{0;\hat{\pi}_S>1-\alpha_{S-SP}}$. Therefore, this is a scenario in which the S-SP requires a maximum sample size (242) larger than the sample size (144) required by WP-SP. In particular, the S-SP provides lower performance than the WP-SP for low-quality lots $C_{pmk} = C_{LTPD}$.

In examining the results, we can state that the sequential sampling plan provides lower performance than the WP-SP only for low capability requirements ($C_{LTPD} < 1.33$) and limited to the case $C_{LTPD}$ =1.33 and $C_{AQL}$=1.67 with:

- $\beta_{WP-SP}$=0.01 for which the S-SP requires a sample size at least equal to $n_{0;\hat{\beta}_{S-SP}<\beta_{S-SP}}$ =324 (Table 1), while the parameters of the WP-SP are $n$=286 and $C_0$=1.4988 for $\alpha_{WP-SP}$=0.01, $n$=245 and $C_0$=1.5137 for $\alpha_{WP-SP}$=0.025 and $n$=213 and $C_0$=1.587 for $\alpha_{WP-SP}$=0.05, respectively (Table 5);

- $\beta_{WP-SP}$=0.025 for which the S-SP requires a sample size at least equal to $n_{0;\hat{\beta}_{S-SP}<\beta_{S-SP}}$ =232 (Table 1) while the parameters of the WP-SP are n=203 and $C_0$=1.4995 $\alpha_{WP-SP}$=0.025 (Table 5) and $n$=174 and $C_0$=1.5149 for $\alpha_{WP-SP}$=0.050 (Table 5);
- $\beta_{WP-SP}$=0.05 for which the S-SP requires a sample size at least equal to $n_{0;\hat{\beta}_{S-SP}<\beta_{S-SP}}$ =161 (Table 1) while the parameters of the WP-SP are $n$=143 and $C_0$=1.5006 for $\alpha_{WP-SP}$=0.050 (Table 5).

## 8. Concluding Remarks

In this article, we propose a sequential test for the process capability index $C_{pmk}$. The study of the statistical properties of the proposed test is complex task as it is necessary to consider several aspects jointly. For this reason, a sequential sampling plan is also proposed, which makes it possible to study the performance of the sequential procedures and compare them with the corresponding non-sequential methodologies in a more intuitive manner.

More precisely, we studied the statistical properties of the sequential test and the sequential sampling plan with an extensive simulation study with regard to the type I error (the consumer's risk of the sampling plan), the average of the sample sizes for correctly deciding, for $H_0$ and $H_1$, and the maximum allowable sample size required to achieve a predetermined power level (the producer's risk of the sampling plan).

We compared the performance of the sequential procedures with that of the PL-test and WP-SP.

The results showed that, for high quality requirements such as $C_{LTPD} > 1.33$, the S-SP allows, on average, a considerable reduction in sampling size compared to the WP-SP (or equivalently the sequential test for $C_{pmk0} > 1.33$ allows, on average, smaller stopping sample sizes compared to the fixed sample size tests, while maintaining the desired $\alpha$ -level and power).

In summary, the proposed sequential procedures have several interesting features: they may offer a substantial decrease in sample size compared to the non-sequential methods, while type I and II error probabilities are correctly maintained at their desired values.

We consider these results to be valuable, because in a highly competitive context where both cost and quality are relevant, the availability of statistical methods which make it possible to reduce sampling size can be directly translated into resource and cost savings.

## Appendix A: Conditions C1-C9

Conditions C1-C7 are the usual requirements (Serfling 1980) for the existence of maximum likelihood estimators. Bhattacharya et al. (2016) showed that these condition are met for *X* normally distributed.

**Condition C8 holds**.

*Proof*.

Let us consider the general case where we have

$$f\left(x,\boldsymbol{\theta}\right)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}\left(x-\mu\right)^2\right),$$

with $\boldsymbol{\theta}=\left(\mu,\sigma^2\right)$ and therefore

$$\ln f\left(x,\boldsymbol{\theta}\right)=-\frac{1}{2}\ln 2\pi-\frac{1}{2}\ln\sigma^2-\frac{1}{2\sigma^2}\left(x-\mu\right)^2.$$

The partial derivatives are of $\ln f\left(x,\boldsymbol{\theta}\right)$ are

$$\frac{\partial\ln f\left(x,\boldsymbol{\theta}\right)}{\partial\mu}=\frac{x-\mu}{\sigma^2}$$

and

$$\frac{\partial\ln f\left(x,\boldsymbol{\theta}\right)}{\partial\sigma^2}=-\frac{1}{2\sigma^2}+\frac{1}{2\sigma^4}\left(x-\mu\right)^2,$$

respectively.

We have to demonstrate that

$$E_\theta\left|\frac{\partial\ln f\left(x,\boldsymbol{\theta}\right)}{\partial\mu}\right|^{2+\delta}<\infty \tag{30}$$

and

$$E_\theta\left|\frac{\partial\ln f\left(x,\boldsymbol{\theta}\right)}{\partial\sigma^2}\right|^{2+\delta}<\infty, \tag{31}$$

for some $\delta>0$.

Without loss of any generality let $\delta=2$ and examine firstly $E_\theta\left|\frac{\partial\ln f\left(x,\boldsymbol{\theta}\right)}{\partial\mu}\right|^{2+\delta}$.

Then,

$$E_\theta\left|\frac{\partial \ln f\left(x,\mathbf{\theta}\right)}{\partial \mu}\right|^{2+\delta}=E_\theta\left[\left(\frac{x-\mu}{\sigma^2}\right)^4\right]=\frac{3\sigma^4}{\sigma^8}=\frac{3}{\sigma^4}, \tag{32}$$

that for $\sigma^2 \neq 0$ is $<\infty$. This proves (30).

Let us consider now $E_\theta\left|\frac{\partial \ln f\left(x,\mathbf{\theta}\right)}{\partial \sigma^2}\right|^{2+\delta}$ with $\delta = 2$. Then,

$$E_\theta\left|\frac{\partial \ln f\left(x,\mathbf{\theta}\right)}{\partial \sigma^2}\right|^{2+\delta}=E_\theta\left[\left(-\frac{1}{2\sigma^2}+\frac{1}{2\sigma^4}\left(x-\mu\right)^2\right)^4\right]. \tag{33}$$

Developing the fourth power in (33) we obtain

$$\begin{aligned}
&E_\theta\left[\begin{array}{l}\left(-\frac{1}{2\sigma^2}\right)^4+4\left(-\frac{1}{2\sigma^2}\right)^3\left(\frac{1}{2\sigma^4}\left(x-\mu\right)^2\right)+6\left(-\frac{1}{2\sigma^2}\right)^2\left(\frac{1}{2\sigma^4}\left(x-\mu\right)^2\right)^2\\ +4\left(-\frac{1}{2\sigma^2}\right)\left(\frac{1}{2\sigma^4}\left(x-\mu\right)^2\right)^3+\left(\frac{1}{2\sigma^4}\left(x-\mu\right)^2\right)^4\end{array}\right]=\\
&=E_\theta\left[\begin{array}{l}\frac{1}{16\sigma^8}-4\frac{1}{8\sigma^6}\frac{1}{2\sigma^4}\left(x-\mu\right)^2+6\frac{1}{4\sigma^4}\frac{1}{4\sigma^8}\left(x-\mu\right)^4\\ -\frac{4}{2\sigma^2}\frac{1}{8\sigma^{12}}\left(x-\mu\right)^6+\frac{1}{16\sigma^{16}}\left(x-\mu\right)^8\end{array}\right]=\\
&=E_\theta\left[\frac{1}{16\sigma^8}-\frac{1}{4\sigma^{10}}\left(x-\mu\right)^2+3\frac{1}{8\sigma^{12}}\left(x-\mu\right)^4-\frac{1}{4\sigma^{14}}\left(x-\mu\right)^6+\frac{1}{16\sigma^{16}}\left(x-\mu\right)^8\right].
\end{aligned} \tag{34}$$

Since, for symmetric distributions

$$E_\theta\left[\left(x-\mu\right)^{2k}\right]=\frac{\left(2k!\right)}{2^k k!}\left(\sigma^2\right)^k, \tag{35}$$

for any integer *k*, the (32) becomes

$$E_\theta\left|\frac{\partial \ln f\left(x,\mathbf{\theta}\right)}{\partial \sigma^2}\right|^{2+2}=\left[\begin{array}{l}\frac{1}{16\sigma^8}-\frac{1}{4\sigma^{10}}\sigma^2+3\frac{1}{8\sigma^{12}}\frac{4!}{4\cdot 2!}\sigma^4\\ -\frac{1}{4\sigma^{14}}\frac{6!}{8\cdot 3!}\sigma^6+\frac{1}{16\sigma^{16}}\frac{8!}{16\cdot 4!}\sigma^8\end{array}\right]=\frac{15}{4\sigma^8}, \tag{36}$$

that for $\sigma^2 \neq 0$ is $<\infty$. This proves (31) and consequently condition C8 holds.

**Condition C9 holds**

*Proof.*

Let us consider the function $h\left(\mathbf{\theta}\right)$ which in our case is defined as

$$h(\boldsymbol{\theta}) = \ln\left(\left(C_{pmk}\right)^2\right) - \ln\left(C_{pmk0}^2\right) = \ln\left[\frac{\left(d/\sigma - |\xi|\right)^2}{9\left(1+\xi^2\right)C_{pmk0}^2}\right]$$
$$= \ln\left[\frac{\left(d/\sqrt{\sigma^2} - |\xi|\right)^2}{9\left(1+\xi^2\right)C_{pmk0}^2}\right], \tag{37}$$

with $\boldsymbol{\theta} = \sigma^2$ scalar.

The (partial) derivative of $h(\boldsymbol{\theta})$ is

$$\frac{\partial h(\boldsymbol{\theta})}{\partial \sigma^2} = \frac{d\left[\sqrt{\sigma^2}\,|\xi| - d\right]}{\sigma^2\left(\xi^2\sigma^2 - 2d\,|\xi|\sqrt{\sigma^2} + d^2\right)} \tag{38}$$

and the subset $\Theta_0$ is $\boldsymbol{\theta}_0 = \sigma_0^2$. Then, we have

$$\lim_{\sigma^2 \to \sigma_0^2} \frac{\partial h(\boldsymbol{\theta})}{\partial \sigma^2} = \frac{d\left[\sqrt{\sigma_0^2}\,|\xi| - d\right]}{\sigma_0^2\left(\xi^2\sigma_0^2 - 2d\,|\xi|\sqrt{\sigma_0^2} + d^2\right)}. \tag{39}$$

Therefore, the (partial) derivative is a continuous function as requested.

Furthermore,

$$H(\boldsymbol{\theta}) = \frac{d\left[\sqrt{\sigma^2}\,|\xi| - d\right]}{\sigma^2\left(\xi^2\sigma^2 - 2d\,|\xi|\sqrt{\sigma^2} + d^2\right)}$$

is of rank q=1. These results satisfy condition C9.

## Appendix B: Scenario Details

In the sequential test, or equivalently in the S-SP, for each value of $C_{pmk0}$ and $C_{pmk1}$, in order to obtain scenarios perfectly comparable with those of the PL-test (or the WP-SP), we set $\xi = 0.5$.

Therefore, the values of the process standard deviations, under $H_0$ and $H_1$, are obtained as

$$\sigma_0 = \frac{d}{3\left(C_{pmk0}\right)\sqrt{1+\xi^2} + |\xi|}$$

and

$$\sigma_1 = \frac{d}{3\left(C_{pmk1}\right)\sqrt{1+\xi^2} + |\xi|},$$

respectively.

Without loss of any generality, we considered processes with $d$=1 and $\mu = 0$. Therefore, we generated observations from normal distributions with $\mu = 0$ and standard deviations $\sigma_0$ and $\sigma_1$ under $H_0$ and $H_1$ respectively.

## References


Arizono I, Kanagawa A, Ohta H, Watakabe K, Tateishi K (1997). Variable sampling plans for normal distribution indexed by Taguchi's loss function. Nav. Res. Logist. 44 (6): 591-603. https://doi.org/10.1002/(SICI)1520-6750(199709)44:6<591::AID-NAV5>3.0.CO;2-Z

Bhattacharya R., Lin L., Patrangenaru V. A Course in Mathematical Statistics and Large Sample Theory. Springer-Verlag: New York 2016.

Borodin AB, Salminen P (1996) Handbook of Brownian Motion-Facts and Formulae. Birkhäuser Verlag, Basel.

Chen SM, Hsu YS (2004) Uniformly most powerful test for process capability index $C_{pk}$. Qual Technol Quant Manag 1(2): 257-269. https://doi.org/10.1080/16843703.2004.11673077

Fan S, Kao C (2006) Development of confidence interval and hypothesis testing for Taguchi capability index using a Bayesian approach. Int J Oper Res, 3(1): 56-75.

Feller W (1970) An Introduction to probability theory and its Applications Vol II. John Wiley & Son, New York.

Kane VE (1986) Process capability indices. J Qual Technol 18(1): 41-52. https://doi.org/10.1080/00224065.1986.11978984

Kotz S, Lovelace CR (1998) Introduction to Process Capability Indices: Theory and Practice. Arnold, London.

Hubele NF, Vännman K (2004) The Effect of Pooled and Un-pooled Variance Estimators on $C_{pm}$ When Using Subsamples. J Qual Technol 36(2): 207-222. https://doi.org/10.1080/00224065.2004.11980266

Hussein A (2005) Sequential Comparison of Two Treatments Using Weighted Wald-Type Statistics. Commun Stat Theory Methods 34(7): 1631-1641.

Hussein A, Ahmed SE, Bhatti S (2012) Sequential testing of process capability indices, J Stat Comput Simul 82(2): 279-282. https://doi.org/10.1080/00949655.2011.638925

Lepore A, Palumbo B, (2015) New Insights into the Decisional Use of Process Capability Indices via Hypothesis Testing. Qual Reliab Eng Int 31(8): 1725-1741. https://doi.org/10.1002/qre.1713

Lepore A, Palumbo B, Castagliola P (2018) A note on decision making method for product acceptance based on process capability indices $C_{pk}$ and $C_{pmk}$. Eur. J. Oper. Res. 267(1): 393-398. https://doi.org/10.1016/j.ejor.2017.12.032

Lin GH (2006) A hypothesis testing procedure on assessing process performance. J Stat Manag Syst 9(2): 319-339. https://doi.org/10.1080/09720510.2006.10701209

Montgomery D (2009) Introduction to Statistical Quality Control (6th ed). John Wiley & Sons, Hoboken.

Negrin I, Parmet Y, Schechtman E (2009). Developing a sampling plan based on $C_{pk}$. Qual Eng 21(3): 306–318. https://doi.org/10.1080/08982110902873597

Negrin, I., Parmet, Y., Schechtman, E., 2011. Developing a sampling plan based on $C_{pk}$-unknown variance. Qual Reliab Eng Int 27 (1): 3-14. https://doi.org/10.1002/qre.1094

Pearn WL, Chen KS (1999) Making decision in assessing process capability index Cpk. Qual Reliab Eng Int 15(4): 321-326. https://doi.org/10.1002/(SICI)1099-1638(199907/08)15:4<321::AID-QRE258>3.0.CO;2-5

Pearn WL, Kotz S (2006) Encyclopedia and Handbook of Process Capability Indices: A Comprehensive Exposition of Quality Control Measures. World Scientific, Singapore.

Pearn WL, Lin PC (2002) Computer program for calculating the p-value in testing process capability index $C_{pmk}$. Qual Reliab Eng Int 18(4): 333-342. https://doi.org/10.1002/qre.465

Pearn WL, Lin PC (2004) Testing process performance based on capability index $C_{pk}$ with critical values. Comput Ind Eng 47(4): 351-369. https://doi.org/10.1016/j.cie.2003.03.001

Pearn WL, Shu MH, Hsu BM (2005) Monitoring manufacturing quality for multiple Li-BPIC processes based on capability index $C_{pmk}$. Int J Prod Res. 43(12): 2493-2512. https://doi.org/10.1080/00207540500045741

Pearn WL, Wu CW (2006a). Critical acceptance values and sample sizes of a variables sampling plan for very low fraction of defectives. Omega 34 (1): 90-101. https://doi.org/10.1016/j.omega.2004.08.001

Pearn WL, Wu CW (2006b). Variables sampling plans with PPM fraction of defectives and process loss consideration. J Oper Res Soc. 57 (4): 450-459. https://doi.org/10.1057/palgrave.jors.2602013

Perakis M, Xekalaki E (2003) On a comparison of the efficacy of various approximations of the critical values for tests on the process capability indices $C_{PL}$, $C_{PU}$, $C_{pk}$. Commun Stat Simul C 32(4): 1249-1264. https://doi.org/10.1081/SAC-120023888

Polansky AM, Kirmani SNUA (2003) Quantifying the capability of industrial processes. In Khattree B, Rao CR (eds.) Handbook of Statistics 22. Elsevier Science, Amsterdam, chapter 17.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Scagliarini M (2018) A sequential hypothesis testing procedure for the process capability index $C_{pk}$ Qual Reliab Eng Int 34(5): 701-806. https://doi.org/10.1002/qre.2290

Scholz FW (2006) Maximum Likelihood Estimation. In Kotz S, Read CB, Balakrishnan N, Vidakovic B, Johnson LN (eds) Encyclopedia of Statistical Sciences, 2nd edn. Wiley, Hoboken, pp.4629 4639.

Serfling RJ. Approximation Theorems of Mathematical Statistics. John Wiley & Sons: New York, 1980.

Vännman K, Kulahci M (2008) A model-free approach to eliminate autocorrelation when testing for process capability. Qual Reliab Eng Int 24(2): 213-228. https://doi.org/10.1002/qre.887

Wu CW, Aslam M, Jon CH (2012) Variables sampling inspection scheme for resubmitted lots based on the process capability index $C_{pk}$. Eur. J. Oper 217(3): 560-566. https://doi.org/10.1016/j.ejor.2011.09.042

Wu CW, Pearn WL (2008) A variables sampling plan based on $C_{pmk}$ for product acceptance determination. Eur. J. Oper. 184 (2): 549–560. https://doi.org/10.1016/j.ejor.2006.11.032

Yen CH, Chang CH (2009). Designing variables sampling plans with process loss consideration. Commun. Stat. Simul. Comput. 38(8): 1579–1591. https://doi.org/10.1080/03610910903046809