12th IMEKO TC10 Workshop on Technical Diagnostics
New Perspectives in Measurements, Tools and Techniques for Industrial Applications
June 6-7, 2013, Florence, Italy

On the design of truncated sequential sampling plans

Giuseppe Cavone[1], Nicola Giaquinto[1], Filomena Innone[1], Valeria L. Scarano[2]

[1] *Politecnico di Bari, Department of Electrical and Information Engineering (DEI), via E. Orabona 4, 70126 Bari, giaquinto@misure.poliba.it*

2 Università di Firenze, Department of Information Engineering, via S. Marta 3, 50139, Firenze, Italy, valeria.scarano@unifi.it

Abstract- The paper presents a study, via Monte Carlo simulations, of the effect of truncation in sequential sampling plans. In sequential sampling plans, truncation is needed to avoid an excessive sample size, which can occur in same cases: on the other hand, truncation affects both the Operating Characteristic (OC) and the Average Sample Number (ASN). In the paper, the approximate analytical formulas used in the literature to determine OC and ASN are compared with the actual curves obtained via Monte Carlo simulations, taking into consideration three different truncation thresholds. Besides, a fourth truncation threshold is empirically determined: it yields approximately the best correspondence between theoretical and Monte Carlo OC and ASN.

I. Introduction

Sampling plans are a key tool for quality management, sharing many theoretical issues with measurement system analysis techniques, reliability tests, and so on. In all these fields, the uncertainty and the optimization of the statistical design assumes an extremely practical and experimental meaning, as pointed out in [1], [2], [3], [4]: it results, for example, in a certain fraction of good lots mistakenly discarded in the process (the “producer’s risk”). A sampling plan is, indeed, a procedure employed in order to decide whether to "accept" or "reject" a set of items, on the basis of observations of a small number of items (sample). In particular, a Sequential Sampling Plan (SSP) is a procedure in which, in contrast with the single sampling plan, the number of items to be inspected is not fixed beforehand, but it varies as a function of the results progressively obtained. This type of procedure, introduced by Wald [5], [6], majorly reduces the Average Sample Number (ASN), which is necessary to make a decision with a given power of discrimination, i.e. with a given Operating Characteristic (OC).
Designing proper SSPs has great practical importance. They are in fact widely used nowadays not only in the industrial field (statistical process control, acceptance sampling, etc.), but also in many other situations, such as the early-life reliability assessment of electronic equipment [7], the management of intervention with pesticides in agriculture [8], or for establishing the prevalence of a disease in a population [9], etc.
In previous work, the authors have developed simplified formulas to design single sampling plans, establishing the deviation between the desired and the actual OC [10], [11]. In this paper, the design of SSPs is examined by Monte Carlo simulations, with a focus on the effect of the truncation on OC and ASN. Different thresholds to truncate the SSP are considered; besides, the focus is on the case of an isolated lot with a limited number of items, whose statistic is, therefore, hypergeometric. The study allows to choose among different ways to trim a SSP, depending on the desired goal. In particular, we show how to have a good reduction of ASN without sacrificing accuracy in the compliance with the specifications, and how to have an almost perfect correspondence between theoretical and actual ASN.

II. Equations for SSP project

In the design of a SSN, the OC must pass through two critical points: the Producer’s Point (PP) and the Customer’s Point (CP), defined by

$$PP = (p_0, 1-\alpha_0) \qquad (1)$$
$$CP = (p_1, \beta_1) \qquad (2)$$

where p_0 and p_1 are the fractions of corresponding defective parts, named Acceptable Quality Level (AQL) and Lot Tolerance Percent Defective (LTPD), while α_0 e β_1 represent, respectively, the Producer’s Risk (PR) and the Customer’s Risk (CR).
Given the specifications in terms of PP and CP, the design equations are:

$$k = \ln\left(\frac{p_1(1-p_0)}{p_0(1-p_1)}\right) \quad (3); \qquad h_0 = \frac{1}{k}\ln\left(\frac{1-\alpha_0}{\beta_1}\right) \quad (4)$$

$$h_1 = \frac{1}{k}\ln\left(\frac{1-\beta_1}{\alpha_0}\right) \quad (5); \qquad s = \frac{1}{k}\ln\left(\frac{1-p_0}{1-p_1}\right). \quad (6)$$

The design equations defines the acceptance and rejection lines:

$$a_n = -h_0 + s\cdot n \quad (7); \qquad r_n = h_1 + s\cdot n \quad (8)$$

These lines are shown in the Figure 1 for a specific case ($p_0 = 3/50$; $p_1 = 9/50$; $\alpha_0 = 0.05$; $\beta_1 = 0.10$).

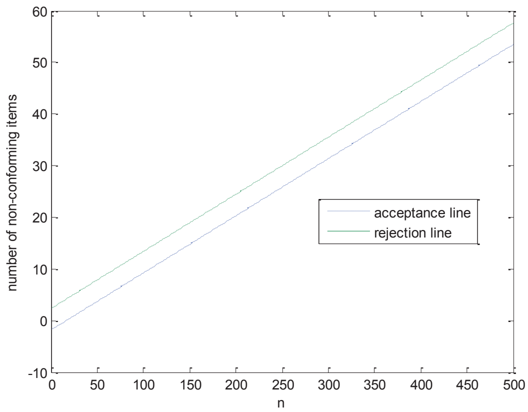


Figure 1. Graphical representation of a sequential sampling plan.

The plan consists in examining one item at a time, applying the rule:

$F_n \le a_n \Rightarrow$ accept

$F_n \ge r_n \Rightarrow$ reject

$a_n < F_n < r_n \Rightarrow$ take another sample

where F_n is the cumulative number of defective parts on n examined items. This procedure leads, after a theoretical analysis, to an OC and an ASN given by the equations:

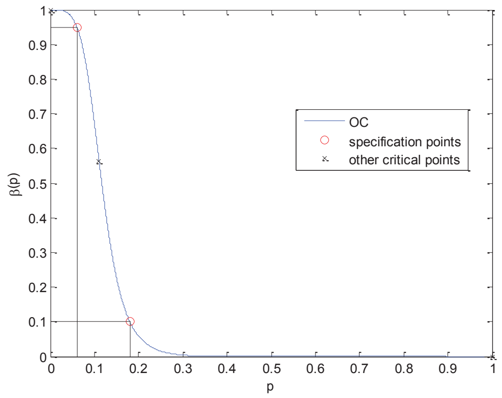
$$p = \frac{1-\left(\frac{1-p_1}{1-p_0}\right)^h}{\left(\frac{p_1}{p_0}\right)^h - \left(\frac{1-p_1}{1-p_0}\right)^h} \quad (9);$$

$$\beta(p) = \frac{\left(\frac{1-\beta_1}{\alpha_0}\right)^h - 1}{\left(\frac{1-\beta_1}{\alpha_0}\right)^h - \left(\frac{\beta_1}{1-\alpha_0}\right)^h} \quad (10)$$

$$ASN(p) = \frac{\beta(p)\ln\left(\frac{\beta_1}{1-\alpha_0}\right) + [1-\beta(p)]\ln\left[\frac{(1-\beta_1)}{\alpha_0}\right]}{p\ln\left(\frac{p_1}{p_0}\right) + (1-p)\ln\left(\frac{1-p_1}{1-p_0}\right)} \quad (11)$$

where the first pair of equations, (9) and (10), defines parametrically the OC, with $-\infty < h < +\infty$.
The OC and the ASN, with the data of the example, are shown in Figure 2.a and Figure 2.b. The figures show five critical points of the OC and the ASN, also reported in Table 1. The five points correspond, of course, to the extreme values $p = 0$ and $p = 1$, to the specification points PP e

CP, and to an intermediate point between PP e CP.

a) 

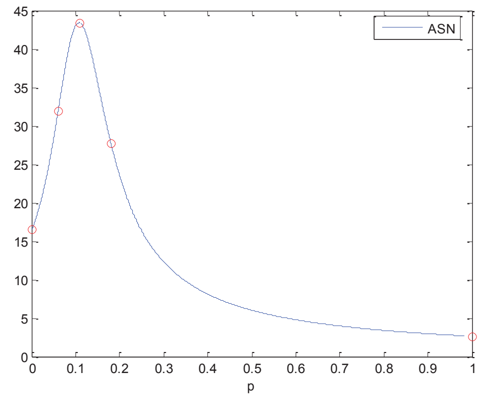
b) 

Figure 2. a) OC plot of a SSP; b) ASN plot of a SSP. Critical points are highlighted.

Table 1. Values of five critical points

h	p	$\beta(p)$	$ASN(p)$
$+\infty$	0	1	h_0 / s
1	p_0	$1-\alpha_0$	$[(1-\alpha_0)h_0 - \alpha_0 h_1]/(s-p_0)$
0	s	$h_1/(h_0+h_1)$	$h_1 h_0 / [s(1-s)]$
-1	p_1	β_1	$[(1-\beta_1)h_1 - \beta_1 h_0]/(p_1 - s)$
$-\infty$	1	0	$h_1/(1-s)$

III. The problem of SSP truncation

The described SSP has the advantage of reducing the ASN, but the defect of leading, occasionally, to excessively high sample size, because it is theoretically possible to remain indefinitely in the region of non-decision, even if with a probability vanishing for increasing sample size. For this reason it is necessary to truncate the SSN, or to establish a maximum sample size, $n_{\max}$, at which a decision is nevertheless taken. The truncation may have an effect on the actual OC, but especially on the ASN, in the neighborhood of the maximum of the curve (i.e. for intermediate values of the fraction of defective p, leading more easily to situations of ambiguity). Another factor which can possibly contribute to have real curves different from theoretical ones is the limited size of the lot to be sampled. As is well-known, in this case the equations, derived under the assumption of binomial distribution, have higher approximations, since the sampling is governed by the hypergeometric distribution. In order to verify the equations of OC and ASN in case of truncated SSN on a lot of finite dimension, Monte Carlo simulation is used. In particular, the following rules of truncation are considered:

$$n_{\max} = 3 \cdot \max\{ASN(p_0), ASN(s), ASN(p_1)\} \qquad (12)$$

$$n_{\max} = n' \cdot \frac{N}{N-1+n'}\text{; with } n' = \left(\frac{z_{\beta_0}\sqrt{p_0(1-p_0)} + z_{\alpha_1}\sqrt{p_1(1-p_1)}}{p_1 - p_0} \right)^2 \qquad (13)$$

$$n_{\max} = \frac{\ln\left[\dfrac{(1-\alpha_0)}{\beta_1}\right] \cdot \ln\left[\dfrac{(1-\beta_1)}{\alpha_0}\right]}{\ln\left(\dfrac{p_1}{p_0}\right) \cdot \ln\left(\dfrac{1-p_1}{1-p_0}\right)} \qquad (14)$$

The first rule, (12), is quite common and is the one followed in [12]; the second, (13), is the length of a single-sampling plan with OC through the same points [10], [11]; the third, (14), is a rule proposed in [9]. In all cases, the decision rule for $n = n_{\max}$ consists in the acceptance for $F_{n_{\max}} \le s \cdot n_{\max}$, rejection otherwise.

The results of the Monte Carlo simulations are shown in the figures below, Figures 4-7. The sampling of an isolated lot of $N = 500$ items was simulated. The AQL corresponds to $K_0 = p_o \cdot N = 30$ defective parts, and the LTPD corresponds to $K_1 = p_1 \cdot N = 90$ defective parts. Moreover, in the execution of simulations, the acceptance number a_n has been rounded up, while the rejection number r_n has been rounded down. This choice, which is also usual, with the imposition of a limit $n_{\max}$ the sample size, leads to represent the SSP as in Figure 3.

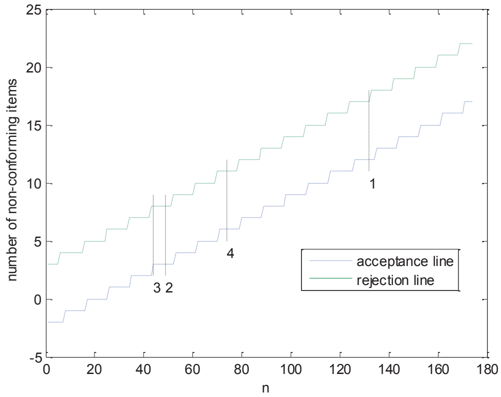


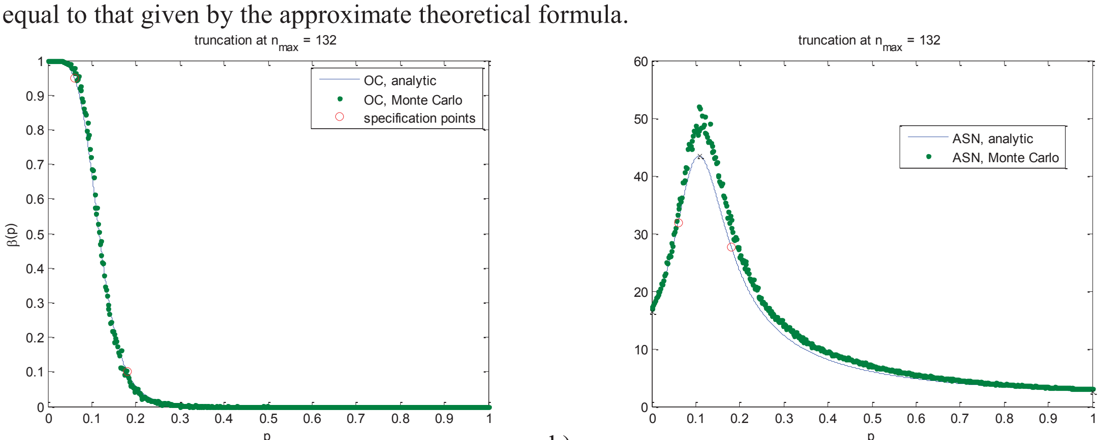
Figure 3. Graphical representation of a sequential sampling plan with integer acceptance and rejection numbers, and truncation of the plan (for four different sample sizes, considered in the paper).

IV. Results from the Monte Carlo simulation

The application of truncation according to (12) (leading to the value $n_{\max} = 132$ in the example) has led to a Monte Carlo OC substantially identical to the theoretical one. The same thing cannot be said, instead, for the ASN, as shown in Figure 4. Actual ASN is greater than the theoretical one, in particular for the critical point $p = s$. Therefore, with this common rule of truncation the sampling has a higher cost than expected.

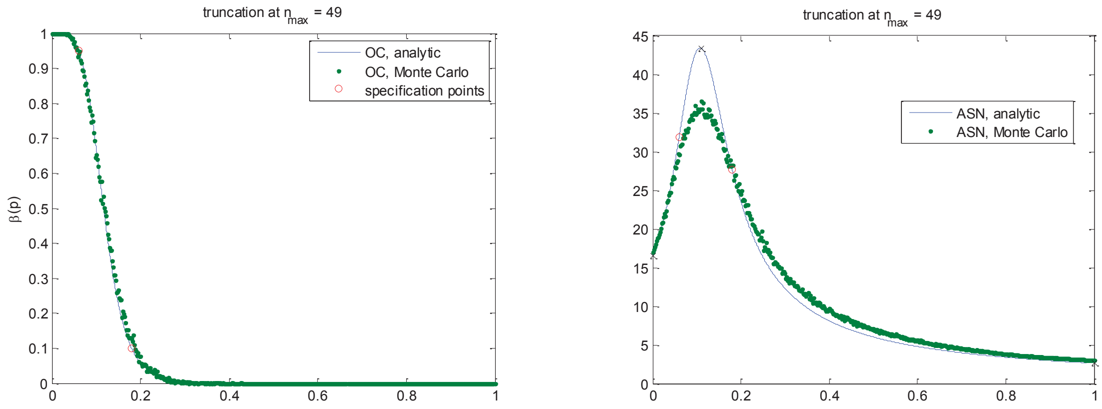
With (13) the truncation occurs much earlier; in the example considered, $n_{\max} = 49$ is obtained. The Monte Carlo simulation shows (Figure 5) that the OC is slightly shifted with respect to the specification points and the theoretical curve. In particular, the actual OC has a greater Producer's Risk (PR). Also the ASN is very different from the theoretical one and, in this case, it has a lower number of samples around the critical point $p = s$.

With rule (14), the truncation happens for an even lower sample size: $n_{\max} = 44$. As shown in Figure 6.a, and Figure 6.b, this case is worse than the previous one. The Monte Carlo OC has a visible difference with respect to the theoretical one, and the ASN is even smaller, presenting deviations from theory similar to those of the previous case. It is important to consider that having a smaller ASN than expected is a practical advantage, but in this case it is achieved at the expanse of the OC. Besides, it makes sense to ask which truncation makes the real ASN equal to that given by the approximate theoretical formula.



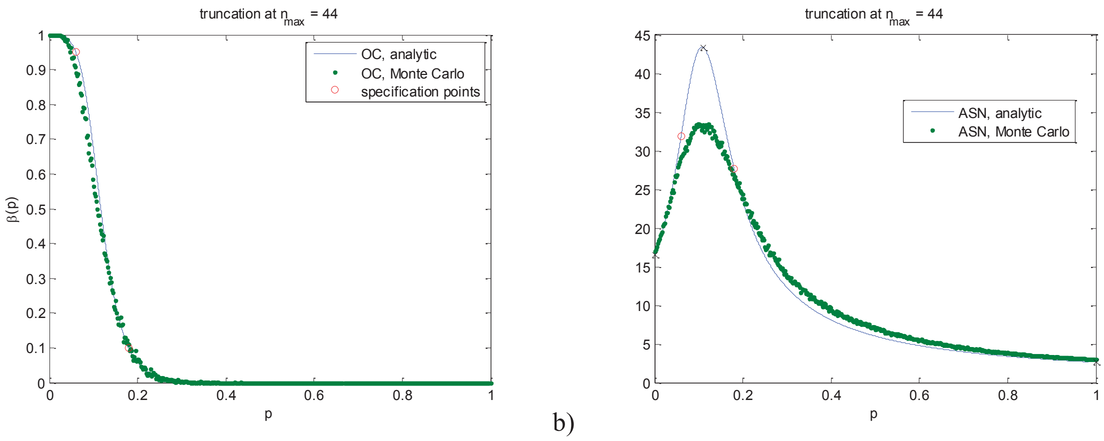
a) b)

Figure 4. a) OC plots and b) ASN plots, comparison between analytical and Monte Carlo method considering rule (12)



a) b)

Figure 5. a) OC plots and b) ASN plots, comparison between analytical and Monte Carlo method considering rule (13)



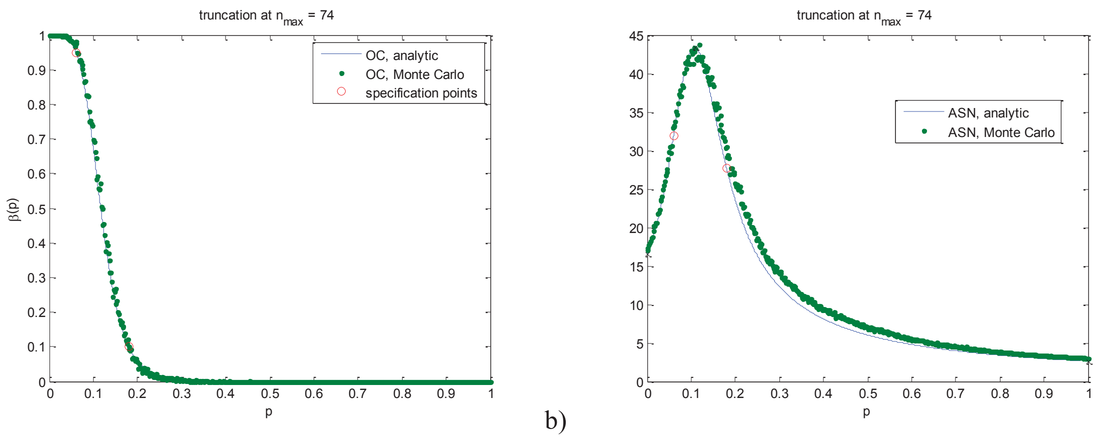
a) b)

Figure 6. a) OC plots and b) ASN plots, comparison between analytical and Monte Carlo method considering rule (14)

Empirically, the best match between the theoretical ASN and the Monte Carlo ASN has been obtained for:

$$n_{\max} = 1.7 \cdot \max\{ASN(p_0), ASN(s), ASN(p_1)\} = 1.7 \cdot \max\left\{\frac{(1-\alpha_0)h_0 - \alpha_0 h_1}{s - p_0}, \frac{h_0 h_1}{s(1-s)}, \frac{(1-\beta_1)h_1 - \beta_1 h_0}{p_1 - s}\right\} \quad (15)$$

In Figure 7 this case, which gives $n_{\max} = 74$, is plotted.



a) b)

Figure 7. a) OC plots and b) ASN plots, comparison between analytical and Monte Carlo method considering the rule (15)

V. Conclusions

Several ways to truncate a sequential sampling plan were examined, calculating both the resulting OC and ASN, with Monte Carlo simulations of isolated lots with a finite number of items. The results were compared to the analytical formulas obtained by approximate analysis of the sampling plan. Based on the analysis, the commonly used rule of truncating the sample at 3 times the theoretical maximum value of the ASN involves an excessive cost of the plan, without any apparent benefit. The truncation used in [9] involves, on the contrary, some saving in the number of items to be controlled, but at the cost of not adhering exactly to the specifications. Truncating the number of items relative to a single sampling plan, calculated according to the formulas in [10], [11], involves a good savings in the number of inspections, without sacrificing accuracy in the compliance with specifications. Finally, in order to have a good match between theoretical expression of the ASN and its actual value, the truncation can be performed at 1.7 times the theoretical maximum value of the ASN.

References

[1] F. Attivissimo, A. Cataldo, L. Fabbiano, and N. Giaquinto, "Systematic errors and measurement uncertainty: An experimental approach," *Measurement*, vol. 44, no. 9, pp. 1781–1789, 2011.
[2] F. Attivissimo, A. Cataldo, N. Giaquinto, and M. Savino, "Assessment of the uncertainty associated with systematic errors in digital instruments: An experimental study on offset errors," *Meas. Sci. Technol.*, vol. 23, no. 3, 2012.
[3] M. Catelani, V. L. Scarano, F. Bertocci, and R. Berni, "Optimization of the Soldering Process With ECAs in Electronic Equipment: Characterization Measurement and Experimental Design," *IEEE Trans. Components Packag. Manuf. Technol.*, vol. 1, no. 10, pp. 1616–1626, 2011.
[4] M. Catelani, V. Scarano, and R. Berni, "Optimization of accelerated testing through Design of Experiment for ageing of Lead-free electronic interconnection material," *Int. J. Metrol. Qual. Eng.*, 2013.
[5] A. Wald, "Sequential Tests of Statistical Hypotheses," *Ann. Math. Stat.*, vol. 16, no. 2, pp. 117–186, Jun. 1945.
[6] A. Wald, *Sequential Analysis*. Courier Dover Publications, 1947.
[7] C. K. Chan, C. I. Seraidaridia, and M. Tortorella, "Sequential sampling plans for early-life reliability assessment," in *Proc. of 1997 Annual Reliability and Maintainability Symposium*, 1997, pp. 131–135.
[8] S. M. Prager, C. D. Butler, and J. T. Trumble, "A sequential binomial sampling plan for potato psyllid (Hemiptera: Triozidae) on bell pepper (Capsicumannum)," *Pest Manag. Sci.*, 2013.
[9] M. Myatt and D. E. Bennett, "A novel sequential sampling technique for the surveillance of transmitted HIV drug resistance by cross-sectional survey for use in low resource settings," *Antivir. Ther.*, vol. 13, no. Supplement 2, pp. 37–48, Apr. 2008.
[10] G. Cavone, L. Fabbiano, and N. Giaquinto, "Closed-form equations to design single sampling plans for isolated lots," in *Proc. of IMEKO XIX World Congress*, Lisbona, Portugal, 2009.
[11] G. Cavone, N. Giaquinto, L. Fabbiano, and G. Vacca, "Design of Single Sampling Plans by Closed-Form Equations," in *Proc. of IEEE I2MTC 2013*, Minneapolis, 2013.
[12] A. K. Hardeo Sahai, "A Visual BASIC Program to Determine Wald's Sequential Sampling Plan," *Rev. Investig. Oper.*, vol. 25, pp. 84–96, 2004.