Received: 13 January 2022 | Revised: 24 May 2022 | Accepted: 27 July 2022

DOI: 10.1002/asmb.2710

SPECIAL ISSUE PAPER

WILEY

Designing acceptance single sampling plans: An optimization-based approach under generalized beta distribution

Silvia Facchinetti | **Silvia Angela Osmetti** | **Umberto Magagnoli**

Department of Statistical Sciences, Università Cattolica del Sacro Cuore, Milan, Italy

Correspondence
Silvia Facchinetti and Silvia Angela Osmetti, Department of Statistical Sciences, Università Cattolica del Sacro Cuore, Milan, Italy.
Email: silvia.facchinetti@unicatt.it; silvia.osmetti@unicatt.it

Abstract

The availability of information that suppliers possess about the production process, as well as about the technical and economic consequences for customers, encourages the development and application of acceptance sampling plans that follow economic criteria, such as the Bayesian ones proposed in the literature. The combination of prior knowledge described by the prior distribution and empirical knowledge based on the sample leads to the decision to accept or reject the lot under inspection. The main purpose of this study was to derive acceptance sampling plans for attributes based on a prior generalized beta distribution following the economic criterion to minimize the expected total cost of quality. Specifically, a procedure is proposed to define the optimal sampling plan based on the technical characteristics of the production process and the costs inherent in the quality of the product. After the methodological aspects are described in detail, an extensive simulation study is reported that demonstrates how the optimal plan changes according to the main parameters, providing guidance for practitioners.

KEYWORDS

acceptance sampling plans, cost function, generalized beta distribution, statistical quality control

1 | INTRODUCTION

Decision making regarding product quality is one of the oldest aspects covered by statistical quality control.[1] The main goal is to provide the desired level of production for customers as well as suppliers. Acceptance sampling is a statistical methodology commonly used to check whether products meet the required quality level. A typical application is in the context of manufacturing, where acceptance sampling plans may be applied in many stages of manufacturing lines. Particularly, they may be applied in incoming material inspection to verify that the quality of goods satisfies certain requirements before they are used or assembled, in online production control to verify that the quality of semi-finished products is acceptable before moving to the next manufacturing step, and in finished product quality auditing to verify that the customer's specifications are satisfied before the shipment. For example, in food industries the production process involves the microbiological inspection of incoming goods, preparation, production, transformation, processing, packaging, and transport of food. In this sector, Luca et al.[2] reporting three different case studies regarding the inspection of a lot of apples by an apple juice company, the control of the eggs freshness before and during the egg processing, and the inspection of an outgoing shipment of pieces of cheese.

The sampling procedure is based on the inspection and classification of a sample of units selected at random from a lot and the final decision to accept or reject the lot. Often, sampling plans are chosen by committees composed of suppliers and customers to consider the opposite needs of the two categories to define the economic/technical effects of product quality with great reliability and validity. This leads to improved quality and reduced costs.

Different types of schema can be applied, such as a single sampling plan, double sampling plan, and multiple sampling plan. In this study, the focus is on single sampling plans by attributes. For a lot of size N submitted for inspection, a single sampling plan is based on the parameters n and c: a lot of size N is accepted if the number of nonconforming units in a sample of size n taken from the lot is less than or equal to the acceptance number c. Therefore, for a given lot, the design of the plan is based on the determination of the sample size to be inspected and the number of nonconformities tolerated for the lot to be accepted. The rejection of a lot means returning the lot to the supplier or its submission to 100% inspection.

The basic assumption underlying the theory of sampling plans by attributes is that the lot quality, in terms of a defective fraction, is a constant, that is, the production process is stable. In practical situations, however, the production process quality often is not constant, and the proportion of nonconforming units in the lot varies frequently. In such situations, the Bayesian approach is more appropriate because it allows incorporate prior process knowledge to account for the variation in the sampling scheme.

This approach has been extensively used in statistical process control (SPC). For example, some researchers have developed different types of control chart for the process mean and variance based on a Bayesian methodology,[3-6] while others have proposed Bayesian sampling plans for attributes[1,7] and variables.[8-10] Brush[11] and Sharma and Bhuttani[12] compared the risks of classical and Bayes producers and consumer, respectively. Recently, Cobb and Li,[13] among others, have suggested the use of Bayesian networks as an alternative to other approaches for Bayesian SPC with attributes, and Ouyang et al.[14] proposed a Bayesian approach for online robust process design.

During the past two decades, Bayesian acceptance sampling plans have received considerable interest in both research and applications. Of particular relevance are the works by Hald,[15-17] who gave an extensive account of sampling plans based on discrete and continuous prior distributions of product quality. Yuvaraj et al.[18] compared sampling plans in the presence of Bayesian and classical settings. GonzÁLez and Palomo[19] developed a Bayesian acceptance sampling plan for the number of defects which minimizes the expected total cost of quality, using both gamma and noninformative prior distributions. Yuvaraj and Subramaniam[20] proposed a Bayesian single sampling plan based on the beta prior distribution obtained by minimizing a cost function and minimizing the cost associated with a defective item that is accepted. Hemalatha and Iyappan[21] presented a procedure for the selection of Bayesian single sampling plans through acceptable and limiting quality levels with the application of numerical algorithms. Nezhad and Saredorahi[22] proposed a policy for designing Bayesian acceptance sampling plans based on the minimum proportion of the lot that should be inspected in the presence of inspection errors. Kaviyarasu and Sivakumar[23] developed a design for a Bayesian single sampling plan under gamma-zero inflated Poisson distributions.

The most commonly used prior distributions to describe the random fluctuations in the lot or process quality $\mathbf{p}$ are binomial, gamma, or beta. The purpose of this study was to develop a sampling schema based on a prior generalized beta distribution, which is more flexible and can adapt to multiple and different situations. In fact, for a suitable choice of the parameters that characterized the generalized beta, we are able to select the most compliant distribution for the considered productive process such as beta, binomial, uniform, triangular and so forth. Moreover, the proposed prior generalizes the beta distribution assuming a restriction in the range of $\mathbf{p}$ in the interval $[p_L, p_U] \in [0, 1]$, whose extremes can be easily identified by the quality controller on the bases of his knowledge about the considered process. This need is motivated by the fact that in many processes (e.g., manufacturing processes) there is high prior evidence that product quality is very high, and priors models on $\mathbf{p}$ with values in the interval [0,1] can be quite unattractive.

A procedure was implemented to define the optimal sampling plan for several values of lot size N on the basis of the technical characteristics of the production process and the costs inherent to the quality of the product. Specifically, minimizing the expected total cost of quality $C(N, n, c)$ is proposed subject to the condition that the consumer risk is reduced.

The remainder of this article is organized as follows. Section 2 provides a description of traditional sapling plans and introduces the Bayesian ones. The Bayesian acceptance sampling plan under a prior generalized beta distribution for the lot fraction nonconforming is given in Section 3. In Section 4, two hypotheses about lot formation are presented, along with the distribution of defective items in a lot. Section 5 focuses on the distribution law of the discrete random variable that describes the number of defective elements in a sample. In Section 6, we derive the average cost function to be minimized to find an economical balance between the cost of sampling and the expected loss resulting from an incorrect decision about the lot. In Section 7, we discuss the main results of an extensive simulation study conducted to evaluate the sensitivity of the procedure to the change of the hypotheses and the main parameters, while Section 8 offers the conclusion.

2 | TRADITIONAL AND BAYESIAN SINGLE SAMPLING PLANS, DECISION RULE AND OC CURVE

Considered a lot of size N, single sampling plans requires two numbers to be specified: the sample size n and the acceptance numbers c. Assume that the lot is submitted for inspection. Based on the outcome of the inspection process in terms of the observed number of nonconforming elements, the decision maker decides to accept the lot, reject it, or perform more inspections.

In single sampling plans the decision about the lot is based on the information contained in only one sample. If more than one sample is required in order to reach a decision regarding the disposition of the lot, double, and multiple sampling plans are considered, as extensions of the single sampling ones.

The procedure for arriving at a decision about the lot is based on three steps:

1. extracting a random sample of size n from a lot of size N received from the producer;
2. inspecting all units in the sample and count the number of defective elements d;
3. comparing the number of defective elements d in the sample with acceptance number c: if $d \leq c$, the lot is accepted; if $d > c$, the lot is rejected.

The decision rule can be expressed considering a function ϕ_c that assumes values 0 and 1 in correspondence with the acceptance or rejection of the lot:

$$\phi_c = \begin{cases} 0 & \text{if } d \leq c, \\ 1 & \text{if } d > c. \end{cases} \tag{1}$$

This function can be seen as the probability of rejecting the lot conditioned to a number of defects in the sample equal to d:

$$P\left\{\text{Rejection}|d\right\} = \phi_c(d),$$

and therefore

$$P\left\{\text{Acceptance}|d\right\} = 1 - \phi_c(d).$$

Acceptance sampling is subject to decision-making risks that an incorrect decision might be made to accept or reject a lot by a consumer or producer. This type of risk is divided into two categories: the producer's risk (α) that a lot with a satisfactory quality level may be rejected based on the statistics of a single sample, and the consumer's risk (β) that a lot with an unsatisfactory quality level may be accepted based on the statistics of a single sample.

The risk for the consumer negatively affects customers because they will use a substandard product which could lead to premature failures and/or an increased maintenance regime. Therefore, clients have not received the product quality for which they paid. The risk for the producer negatively affects manufacturers because a conforming-quality lot will be deemed unsatisfactory, and, consequently, more expensive processes to improve its quality must be implemented. Finally, to combat the risk of rejection, suppliers increase product prices, and this leads to a value-for-money that is invalid for both consumers and producers. Sampling plans are usually designed to control one or both of these risks.[24]

In this context, an important measure of the performance of a sampling plan is the operative characteristic (OC) curve that provides the relationship between the probability of accepting a lot and its production quality.[1] Let D be the number of defective elements in the lot. Let $s_{\mathbf{d}|D}$ and $S_{\mathbf{d}|D}$ be the probability distribution function (p.d.f.) and cumulative

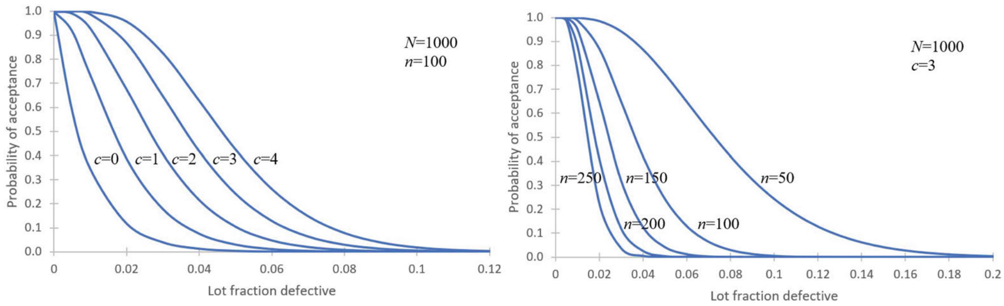

FIGURE 1 OC curves of the single sampling plan. Left: $N = 1000, n = 100$ and several values of c; right: $N = 1000, c = 3$, and several values of n

distribution function (c.d.f.) of the number of defective elements in a sample, respectively. The probability of accepting a lot is defined as

$$P_a = P\{\text{Acceptance}|D\} = P(d \le c) = \sum_{d=0}^{n}[1-\phi_c(d)]s_{\mathbf{d}|D}(d) = S_{\mathbf{d}|D}(c). \tag{2}$$

For example, if we suppose that the distribution of the number of defective elements in a sample follows a binomial distribution with parameters n and p, the P_a in Equation (2) becomes

$$P_a = \sum_{d=0}^{c} \binom{n}{d} p^d (1-p)^{n-d}, \tag{3}$$

where p is the fraction of defective elements in the lot.

The OC curve is developed by evaluating P_a for different values of p.

For $N = 1000$, Figure 1 shows how the OC curve changes as the acceptance number (left figure) and sample size (right figure) change. It is generated by evaluating Equation (3) for several values of p.

This curve plot describes the performance of the plan for different quality levels p, sample sizes n, and acceptance numbers c, and it displays the discriminatory power of the sampling plan. It helps to control the quality of products and to find adequate sample sizes, contributing to the design of better and more efficient experiments. Specifically, in Figure 1 if c increases, the OC curve shifts to the right. Thus, plans with smaller acceptance numbers provide discrimination at lower levels of lot defective fraction than those with larger values of c. Moreover, the OC curve becomes closer to the idealized OC curve shape as the sample size increases. This means that the ability of a sampling plan to differentiate between good and bad lots increases with n.

There is a widely used approach of picking n and c: require that the OC curve pass through two specific points such that $P_a(p) = 1 - \alpha$ for lots with fraction defective $p = p_1$, and $P_a(p) = \beta$ for lots with fraction defective $p = p_2$. It is usual in many industries to consider $p_1 =$ AQL (acceptable quality level), that is the quality level corresponding to the consumer's risk and $p_2 =$ LQL (limiting quality level), that is the quality level corresponding to the producer's risk.

The basic assumption underlying the theory of sampling plans by attributes is that the lot quality, in terms of a defective fraction, is a constant, that is, the production process is stable. This happens when the two types of quality variation in sampling inspection, the within-lot (sampling) and between-lot (sampling and process) variations, are equal. In practical situations, however, the production process often is not constant because the between-lot variation is greater than the within-lot variation, and the proportion of nonconforming units in the lot varies frequently. In these cases, decisions regarding a lot should be made considering the between-lot variations, and the quality of the lot should be treated as a random variable; hence, the conventional sampling schemes cannot be employed (see Suaresh and Umamaheswari[25] and Vijayaraghavan et al.[26] for details). In such situations, the Bayesian approach is more appropriate for studying the sampling plan based on attributes.

The difference between the frequentist and Bayesian approaches is related to the use of prior process history or knowledge in making decisions about a lot.[27] In fact, the idea of the Bayesian approach is to consider not only the most recent observations from sample data, but also other information about unknown parameter probability distributions available from expert judgment, past experience, observations of managers and engineers, or the literature. Quantification of this prior information would lead to better decisions.

One of the most common choices for a prior distribution of productive process quality is the beta distribution, owing to its mathematical simplicity, especially the reproducibility property.[16] We generalize this approach by considering the generalized beta distribution as prior, and we analyze its properties. Moreover, the Bayesian approach can be used to optimize some utility function in order to select the optimal sampling plan. In the following sections we will show how to determine the optimum single sampling plan (n, c) based on prior generalized beta distribution by minimizing the average acceptance cost $C(N, n, c)$ subject to the condition that the consumer risk is reduced.

3 | GENERALIZED BETA PRIOR DISTRIBUTION

The Bayesian acceptance sampling approach is based on prior process history for the selection of the appropriate distributions used to describe the random fluctuation involved in acceptance sampling. The prior distribution is the expected distribution of the productive process quality on which the sampling plan operates.

The quality $p(t)$ is defined as the probability of obtaining a nonconforming element in the lot at production time t. It is assumed that the quality of the productive process $\mathbf{p}$ is a random variable (r.v.) that varies over time following a statistical law that depends on the nature of the productive process and satisfies the stationarity property. In this study, it was assumed that $\mathbf{p}$ follows a generalized beta distribution (GB), which is a linear transformation of a beta r.v. ($Beta$). The major details about continuous distributions can be found elsewhere.[28]

Specifically, let $X \sim Beta(a, b)$ be a r.v. with density function (d.f.) and c.d.f. as follows:

$$\varphi(x) = \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)} \quad \text{for} \quad 0 \le x \le 1,$$

and

$$\phi(x) = \int_0^x \varphi(t)\, dt = I_x(a,b) \quad \text{for} \quad 0 \le x \le 1,$$

where $B(a, b)$ is the beta function

$$B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1} dt,$$

and $I_x(a, b)$ is the incomplete beta function

$$I_x(a,b) = \int_0^x \frac{t^{a-1}(1-t)^{b-1}}{B(a,b)} dt.$$

The expected value and variance of X are, respectively:

$$\begin{cases} \mu_X = \frac{a}{a+b}, \\ \sigma_X^2 = \frac{ab}{(a+b)^2(a+b+1)}. \end{cases} \tag{4}$$

The quality of the productive process $\mathbf{p}$ is defined as

$$\mathbf{p} = p_L + (p_U - p_L)X, \tag{5}$$

where p_L and $(p_U - p_L)$ are a location and a scale parameter, respectively, such that $0 \le p_L \le p_U \le 1$. Therefore, the quality of the productive process $\mathbf{p}$ assumes values in the interval $[p_L, p_U]$ whose extremes can be easily identified by the quality controller on the bases of his knowledge about the considered process. In particular, if the quality controller selects $p_L = 0$ and $p_U = 1$, $\mathbf{p}$ follows a Beta distribution with parameters a and b.

It follows that $\mathbf{p} \sim GB(p_L, p_U, a, b)$, with d.f.

$$f_{\mathbf{p}}(p) = \frac{(p_U - p_L)^{1-a-b}}{B(a,b)}(p - p_L)^{a-1}(p_U - p)^{b-1} \quad \text{for} \quad p_L \leq p \leq p_U, \tag{6}$$

and c.d.f.

$$F_{\mathbf{p}}(p) = \int_{p_L}^{p} f_{\mathbf{p}}(t)dt = I_{\frac{p-p_L}{p_U-p_L}}(a,b) \quad \text{for} \quad p_L \leq p \leq p_U. \tag{7}$$

The expected value and variance of $\mathbf{p}$ can be derived from the expected value and variance of X as follows:

$$\begin{cases} \mu_{\mathbf{p}} = p_L + (p_U - p_L)\mu_X, \\ \sigma^2_{\mathbf{p}} = (p_U - p_L)^2\sigma^2_X. \end{cases} \tag{8}$$

The generalized beta can adapt to multiple and different situations and hypotheses by a suitable selection of the parameters p_L, p_U, a, b that characterize the distribution. As said above, the proposed prior generalizes the beta distribution assuming a restriction in the range of $\mathbf{p}$ in the interval $[p_L, p_U] \in [0, 1]$. The values of p_L an p_U can be selected by the quality controller on the bases of his knowledge about the process. This restriction is very suitable for such processes characterized by high prior evidence that product quality is very high, and for which the choice of $\mathbf{p} \in [0, 1]$ is quite unattractive. Moreover, the choice of a and b allow us to select the most compliant trend of the density function in (6). In fact, if $a = b, X$ follows a symmetric distribution; if $a = b = 1, X$ follows a uniform distribution in $(0, 1)$; if $a = 1, b = 2, X$ follows a triangular increasing distribution in $(0, 1)$, if $a = 2, b = 1, X$ follows a triangular decreasing distribution in $(0, 1)$.

In the next sections we will consider several *GB* distributions with different combinations of parameters.

Parameters a and b are linked to the expected value and variance of the r.v. $\mathbf{p}$, and they can be estimated using Equations (4) and (8). From (8), one can obtain

$$\mu_X = \frac{\mu_{\mathbf{p}} - p_L}{p_U - p_L} \quad \sigma^2_X = \frac{\sigma^2_{\mathbf{p}}}{(p_U - p_L)^2}.$$

Solving Equation (4) with respect to a and b results in

$$\begin{cases} a = \dfrac{(\mu_X - \mu^2_X - \sigma^2_X)\mu_X}{\sigma^2_X}, \\ b = \dfrac{(\mu_X - \mu^2_X - \sigma^2_X)(1 - \mu_X)}{\sigma^2_X}. \end{cases} \tag{9}$$

Because $a, b > 0$, the values of μ_X and σ^2_X must be such that*

$$0 < \mu_X < 1 \quad 0 < \sigma^2_X < \mu_X(1 - \mu_X). \tag{10}$$

Therefore, the quality of the considered productive process can be summarized with a prior generalized beta distribution: $\mathbf{p} \sim GB(p_L, p_U, \mu_X, \sigma^2_X)$.

4 | LOT FORMATION AND DISTRIBUTION OF DEFECTIVE ITEMS

In the previous section, we defined the quality of the productive process $\mathbf{p}$ and the associated prior distribution $F_{\mathbf{p}}(p)$ reported in Equation (7). This distribution does not highlight the sequential structure that could exist in the production cycle, and therefore, the time series nature of the quality $\mathbf{p}$ which may affects the defects of the lot elements.

In this regard, it is difficult to put forward concrete hypotheses because the type of product, the nature of the defect, the technological methods of production, and the collection and construction of lots, can differ greatly.

In this study, two hypotheses about lot formation were considered.

*If μ_X assumes an extreme value (0 or 1), the r.v. $\mathbf{p}$ degenerates into a discrete r.v. with a single value (p_L or p_U).

4.1 Hypothesis 1

The lot is composed of elements coming from a phase of the production cycle characterized by a constant quality value, which may differ from one lot to another. In this situation we suppose that there is high dependence among the quality of the lot elements. This can reasonably be assumed when the lot size is limited.

For a given lot of size N, the number of defects D follows a discrete distribution identified in relation to the prior distribution of the process quality $\mathbf{p}$. Under the assumption of a constant probability for each element of the lot to be defective, the p.d.f. of the conditioned random variable $\mathbf{D}|p$ results in

$$r_{\mathbf{D}|p}(D) = \binom{N}{D} p^D (1-p)^{N-D} \qquad D = 0, 1, \ldots, N. \tag{11}$$

In other words, the conditioned r.v. $\mathbf{D}|p$ follows a binomial distribution with parameters N and p: $\mathbf{D}|p \sim Bin(N, p)$.

To obtain the distribution law of the number of defects D in a lot considering the prior distribution of $\mathbf{p}$, it is necessary to consider the distribution function of the marginal r.v. $\mathbf{D}$

$$r_{\mathbf{D}}(D) = \int_0^1 r_{\mathbf{D}|p}(D) dF_{\mathbf{p}}(p) = \binom{N}{D} \int_0^1 p^D (1-p)^{N-D} dF_{\mathbf{p}}(p) \qquad D = 0, 1, \ldots, N. \tag{12}$$

Assuming as the prior distribution of $\mathbf{p}$ the one reported in Equation (6), one obtains

$$r_{\mathbf{D}}(D) = \binom{N}{D} \frac{(p_U - p_L)^{1-a-b}}{B(a,b)} \int_{p_L}^{p_U} p^D (1-p)^{N-D} (p - p_L)^{a-1} (p_U - p)^{b-1} dp. \tag{13}$$

This relationship is complex and does not lead to synthetic analytical expressions. From this equation, the associative structure between the binomial and the generalized beta distributions is evident.

It is of interest to consider some particular situations.

1. If the prior distribution of $\mathbf{p}$ degenerates into a discrete r.v. with a single value $p = p_0$ and unit probability, the p.d.f. of $\mathbf{D}$ in Equation (13) becomes

$$r_{\mathbf{D}}(D) = \binom{N}{D} p_0^D (1-p_0)^{N-D} \qquad D = 0, 1, \ldots, N. \tag{14}$$

   Therefore, $\mathbf{D}$ follows a binomial distribution with parameters N and p_0.
2. If the prior distribution of $\mathbf{p}$ is uniform in the interval $(0, 1)$, that is $\mathbf{p} \sim GB(0, 1, 1/2, 1/12)$, the p.d.f. of $\mathbf{D}$ in Equation (13) becomes

$$r_{\mathbf{D}}(D) = \binom{N}{D} B(D+1, N-D+1) = \frac{1}{N+1} \qquad D = 0, 1, \ldots, N. \tag{15}$$

   Therefore, $\mathbf{D}$ follows a discrete rectangular distribution.
3. If the prior distribution of $\mathbf{p}$ is uniform in the interval $(0, p_0)$, with $p_0 < 1$, that is $\mathbf{p} \sim GB(0, p_0, 1/2, 1/12)$, the p.d.f. of $\mathbf{D}$ in Equation (13) becomes

$$r_{\mathbf{D}}(D) = \frac{I_{p_0}(D+1, N-D+1)}{p_0(N+1)} \qquad D = 0, 1, \ldots, N. \tag{16}$$

4. If $\mathbf{p}$ follows a beta distribution with $p_L = 0$ and $p_U = 1$, and the parameters a and b are obtained using Equation (9), the p.d.f. of $\mathbf{D}$ in Equation (13) becomes

$$r_{\mathbf{D}}(D) = \frac{B(D+a, N-D+b)}{(N+1)B(D+1, N-D+1)} \qquad D = 0, 1, \ldots, N. \tag{17}$$

Even in these particular situations, one obtains expressions of considerable complexity for the p.d.f. of $\mathbf{D}$.

4.2 | Hypothesis 2

The lot is composed of elements from different stages of the production cycle, with quality values different from element to element. In this situation we suppose that there is independence among the quality of the lot elements.

The condition of independence implies that each element i of the lot ($i = 1, 2, \dots, N$) could be defective. Therefore, r.v. $\mathbf{D}_i$ which takes values 0 and 1 depending on whether the element of the lot is nondefective or defective, is introduced.

The p.d.f. of the r.v. $\mathbf{D}_i$ expressed in terms of the prior distribution of the quality $\mathbf{p}$, results in

$$r_{\mathbf{D}_i}(D_i) = \int_0^1 p^{D_i}(1-p)^{1-D_i} dF_{\mathbf{p}}(p) \qquad D_i = 0, 1.$$

More precisely,

$$r_{\mathbf{D}_i}(0) = 1 - \int_0^1 p dF_{\mathbf{p}}(p) = 1 - \mu_{\mathbf{p}},$$
$$r_{\mathbf{D}_i}(1) = \int_0^1 p dF_{\mathbf{p}}(p) = \mu_{\mathbf{p}}.$$

Therefore, every r.v. $\mathbf{D}_i$ follows a binomial distribution $\mathbf{D}_i \sim Bin(1, \mu_{\mathbf{p}})$ for $i = 1, 2, \dots, N$, and the total number of defects $\mathbf{D}$ in a lot follows a binomial distribution given by the sum of the N r.v.s $\mathbf{D}_i$ as follows

$$\mathbf{D} = \sum_{i=1}^{N} \mathbf{D}_i \sim Bin(N, \mu_{\mathbf{p}}).$$

The p.d.f. of $\mathbf{D}$ is therefore:

$$r_{\mathbf{D}}(D) = \binom{N}{D} \mu_{\mathbf{p}}^{D}(1-\mu_{\mathbf{p}})^{N-D} \qquad D = 0, 1, \dots, N \tag{18}$$

with mean and variance, respectively, of

$$E(\mathbf{D}) = N\mu_{\mathbf{p}} \quad Var(\mathbf{D}) = N\mu_{\mathbf{p}}(1-\mu_{\mathbf{p}}).$$

In this case, the prior distribution of $\mathbf{p}$ is related to the distribution of the lot only through its expected value $\mu_{\mathbf{p}}$.

Following Hald,[15,16] the above distributions of $\mathbf{D}$ defined under Hypothesis 1 or 2 are the prior distributions on which to base the optimal choice of the sampling plan.

Once the p.d.f. of defects in a lot is defined, one can define the quality of the lot $\hat{\mathbf{p}}$ as the discrete r.v.

$$\hat{\mathbf{p}} = \frac{\mathbf{D}}{N}, \tag{19}$$

whose p.d.f. is provided by the one of $\mathbf{D}$ as follows

$$g_{\hat{\mathbf{p}}}(\hat{p}) = r_{\mathbf{D}}(N\hat{p}) \qquad \hat{p} = 0, \frac{1}{N}, \frac{2}{N}, \dots, 1. \tag{20}$$

Both the distributions of $\mathbf{D}$ and $\hat{\mathbf{p}}$, depend on the lot size N.

5 | DISTRIBUTION OF DEFECTIVE ELEMENTS IN THE SAMPLE

In this section, a simple random sample of size n extracted from a lot of size N is considered. The focus is on the distribution law of the discrete r.v. which describes the number of defects $\mathbf{d}$ in the sample.

If D is the number of defective elements in the lot, the conditioned r.v. $\mathbf{d}|D$ follows a hypergeometric distribution $\mathbf{d}|D \sim Hypergeometric(N, D, n)$ with p.d.f.

$$s_{\mathbf{d}|D}(d) = \frac{\binom{D}{d}\binom{N-D}{n-d}}{\binom{N}{n}} \quad d = d_0, d_0+1, \ldots, d_f, \tag{21}$$

where

$$d_0 = \max[0, n-(N-D)] \quad \text{and} \quad d_f = \min[n, D].$$

For the sake of simplicity and speed of the calculation, the hypergeometric distribution can be approximated by considering the binomial distribution

$$\mathbf{d}|D \sim Bin\left(n, \hat{p} = \frac{D}{N}\right), \tag{22}$$

or the Poisson one,

$$\mathbf{d}|D \sim Poi(\lambda = n\hat{p}). \tag{23}$$

However, attention must be paid to the accuracy of the final result because, if the value of N is rather small, the considered approximation may not be satisfactory. For the motivation to the use of discrete distributions for the number of defective elements in a sample from a production line see for example, Irony.[29]

Because the distribution of the number of defects in the lot is known (see Section 4), the joint p.d.f. of the defects in the sample and in the lot is

$$s_{\mathbf{dD}}(d, D) = s_{\mathbf{d}|D} r_{\mathbf{D}}(D) = \frac{\binom{D}{d}\binom{N-D}{n-d}}{\binom{N}{n}} r_{\mathbf{D}}(D), \tag{24}$$

for $D = 0, 1, \ldots, N$ and $d = d_0, d_0+1, \ldots, d_f$.

Therefore, the marginal distribution of the number of defects in the sample is

$$s_{\mathbf{d}}(d) = \sum_{D=0}^{N} \frac{\binom{D}{d}\binom{N-D}{n-d}}{\binom{N}{n}} r_{\mathbf{D}}(D), \tag{25}$$

assuming that the binomial coefficients $\binom{x}{y}$ are null for $x < y$.

Remembering the identity

$$\frac{\binom{D}{d}\binom{N-D}{n-d}}{\binom{N}{n}} = \frac{\binom{n}{d}\binom{N-n}{D-d}}{\binom{N}{D}}, \tag{26}$$

and indicating with $M = N - n$ the number of uninspected elements in the lot, and with $E = D - d$ the number of defective elements in the uninspected part of the lot, one obtains

$$s_{\mathbf{d}}(d) = \binom{n}{d} \sum_{E=0}^{M} \frac{\binom{M}{E}}{\binom{N}{d+E}} r_{\mathbf{D}}(d+E), \tag{27}$$

which is a compound hypergeometric distribution.

Considering the compound hypergeometric, Hald[15(sec. 8)] introduced the distribution reproducibility condition for the prior distributions, that is: if the compound hypergeometric distribution defined in (27) (the marginal distribution of the defective elements in the sample extracted from a lot of size N) is equal to the distribution $r_D(D)$ of the defective elements of a lot of size $N = n$, then the prior distributions are said to be reproducible or invariant to hypergeometric sampling.

This definition could be interpreted as follows: consider a process able to generate lots, corresponding to finite populations of size N, draw a random sample of size n without replacement from every lot of size $N > n$. The sample distributions will coincide with the lot distributions of size $N = n$ generated by the process. Therefore, the information obtained from the sample may be considered as drawn directly from the productive process.

The Hald's condition of distribution reproducibility is verified only under *Hypothesis 2* (see Section 4.2). Only in this case we can consider the information obtained from the sample as taken directly from the productive process. Otherwise, under *Hypothesis 1* (see Section 4.1), the information obtained from the sample depends on the selected lot which corresponds to a specific quality value. The only exception is the case of a uniform prior distribution of $\mathbf{p}$ in the interval $(0, 1)$ (case 2. in Section 4.1) which implies a rectangular distribution of $\mathbf{D}$.

6 | COST FUNCTION AND OPTIMAL SAMPLING PLAN

In this article, the choice of a sampling plan is delegated to an economic optimization that considers the costs of quality and sampling that fall on the supplier and buyer.

The expected total cost for acceptance sampling depends on the acceptance or rejection of the lot. If the lot is accepted, the remainder of the lot is not inspected and, for each defect, a cost is incurred. If the lot is rejected, there is a cost related to the lot size. In addition, the sample inspection cost is a function of its size.[19]

The following linear cost functions are assumed in the case of acceptance or rejection of the lot:

$$\begin{aligned} C_A &= K_1 d + K_2 n + K_3(D - d) && \text{if lot is accepted,} \\ C_R &= K_1 D + K_2 N && \text{if lot is rejected.} \end{aligned} \tag{28}$$

The above costs are total for the N elements of the lot and are supplementary to the manufacturing and general costs. Specifically:

1. $K_1 d$: replacing cost for defective elements found in the sample;
2. $K_2 n$: cost of sample survey;
3. $K_3(D - d)$: cost associated with defective elements in the accepted lot;
4. $K_1 D$: replacing cost for defective elements found in the lot (assuming the hypothesis of 100% inspection in case of rejection);
5. $K_2 N$: cost of sample survey and of 100% inspection carried out in case of rejection.

The coefficients K_1, K_2, and K_3 are obtained based on the cost assessments made by the supplier (K_1, K_2) and customer (K_3).

Given the random nature of some quantities present in the cost functions in Equation (28), as well as the decision rule in Equation (1), one can define the r.v. cost $\mathbf{C}$ as follows:

$$\mathbf{C} = \mathbf{C}_A(1 - \phi_c) + \mathbf{C}_R \phi_c = C(\mathbf{d}, \mathbf{D}). \tag{29}$$

The aim is to provide a procedure for striking an economical balance between the cost of sampling and the expectation of loss resulting from the acceptance of a lot with an unsatisfactory quality level or the rejection of a lot with a satisfactory quality level.

The optimization procedure is carried out by minimizing the average cost function $E(\mathbf{C})$, which, because both d and D assume discrete values, is given by

$$E(\mathbf{C}) = \sum_{D=0}^{N} E(\mathbf{C}|D) r_{\mathbf{D}}(D), \tag{30}$$

where

$$E(\mathbf{C}|D) = \sum_{d=0}^{n} C(d, D) s_{\mathbf{d}|D}(d),$$

is the average cost for a given number of defective elements D in the lot.

Thus, the expected value of the cost function $E(\mathbf{C})$ is a linear combination of the conditioned averages $E(\mathbf{C}|D)$, whose weights are the probability function of defects D in the lot.

Assuming $K_3 = 1$ for simplicity, from Equations (28) to (30), one obtains

$$E(\mathbf{C}|D) = (K_1 D + K_2 N) + [K_2 n - (K_1 - 1)D - K_2 N] S_{\mathbf{d}|D}(c) + (K_1 - 1) E_{\mathbf{d}|D}(c), \tag{31}$$

where

$$S_{\mathbf{d}|D}(x) = \sum_{d=0}^{x} s_{\mathbf{d}|D}(d) \quad \text{and} \quad E_{\mathbf{d}|D}(x) = \sum_{d=0}^{x} d s_{\mathbf{d}|D}(d), \tag{32}$$

are functions of x, with parameters N, D, and n.

Together with the objective function $E(\mathbf{C})$ for minimizing, two supplementary conditions are considered. The first is related to the lot size N, which, instead of being specified at the beginning, can be determined as a result of the optimization procedure, requiring only that it is included in an interval ($N_L \leq N < N_U$). In this case, for values of N included in the above interval, one consider, as an objective function for minimizing, the average cost per element of the lot $E(\mathbf{C})/N = C(N, n, c)$.

The second condition is related to the risk for the consumer to ensure a high rejection probability for lots whose quality $\hat{p} = D/N$ is to be considered unsatisfactory. In this case, when a small value for the probability β (5%–10%), and a limit value $\hat{p}^*$ for the quality of the lot are fixed, the optimal sample plan must be such that $P_a(\hat{p}^*) < \beta$. The presence of this constraint makes the choice of the sampling plan approach that of the traditional type or that proposed by Dodge and Roming,[30] while not neglecting the characteristic of an economic optimization based on an assumption of prior distribution for the productive process quality. In this way, the need of the consumer to be protected sufficiently from highly defective lots whose effects on the economy and efficiency of the buyer could be greater than can be assessed on the basis of the normal quality level of the product can be satisfied. For example, in the case in which the buyer is an entity that manages delicate and complex services, such as public transport, electricity, and hospitals, the direct economic burden caused by the defective elements can be added to the one related to the safety and continuity of the service. Obviously, compliance with this need of the consumer should not conflict with the definition of the economic expression, in which unit costs should be considered within the normal values and quality parameters in relation to the average production cycle, as the supplier expects. An excessively high defect cost evaluation and a excessive control by the consumer would lead the producer to increase the quality, but with increases in production costs that would fall on the customer.

In summary, a design model is proposed that minimizes the cost function $E(\mathbf{C})/N = C(N, n, c)$ subject to the condition that the lot size is bounded in an interval and the consumer risk is reduced:

$$\begin{aligned}\min_{N,n,c} C(N, n, c) = \min_{N,n,c} \Bigg\{ K_2 + \frac{1}{N} \Bigg[& K_1 \, \mu_{\mathbf{D}} - K_2 (N - n) \sum_{D=0}^{N} S_{\mathbf{d}|D}(c) \, r_{\mathbf{D}}(D) + \\ & -(K_1 - 1) \sum_{D=0}^{N} \left(D \, S_{\mathbf{d}|D}(c) - E_{\mathbf{d}|D}(c) \right) r_{\mathbf{D}}(D) \Bigg] \Bigg\},\end{aligned} \tag{33}$$

where $\mu_{\mathbf{D}} = \sum_{D=0}^{N} D \, r_{\mathbf{D}}(D)$, and for $N_L \leq N < N_U$, $S_{\mathbf{d}|D^*}(c) < \beta$, with $D^* = \hat{p}^* N$.

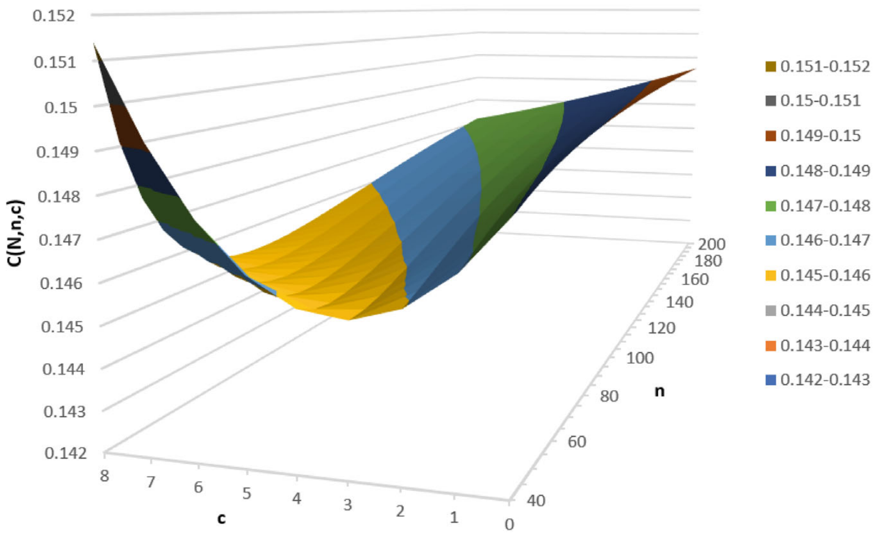

FIGURE 2 Trend of the cost function $C(N, n, c)$ for $N = 1000$

In Figure 2, an example of the behavior of the cost function $C(N, n, c)$ to minimize for $N = 1000$ and for different values of $n \in [40{,}200]$ and $c \in [0, 8]$ is provided. We observe that the average cost function tend to be U-shaped, decreasing as c decrease and n increase, and then increasing for high values of n and low values of c. The aim is to find the best combination (n, c) that minimizes the curve subject to the above constraint conditions.

7 | SIMULATION STUDY

In this section, simulation study was performed to select the optimal sampling plans such that the function in Equation (33) is minimized with respect to the condition $P_a(\hat{p}^*) < 10\%$.

It was assumed that K_1 and K_2 are cost assessments made by the supplier.

There may exist multiple solutions depending on the considered hypothesis and on a number of parameters. In this study, the quality of the considered productive process is a r.v. that follows a prior generalized beta distribution. In line with what was described in Section 4, two hypotheses about lot formation were considered: *Hypothesis 1*, for which the lot is composed of elements coming from a phase of the production cycle characterized by a constant quality value, and *Hypothesis 2*, for which the lot is composed of elements from different stages of the production cycle, with quality values different from element to element related to the prior distribution.

The results of some of the simulation study for both the hypothesis follow.

7.1 | Hypothesis 1

The distribution function $r_{\mathbf{D}}(D)$ of the number of defects in the lot is defined by Equation (13). It is supposed that $\mathbf{p}$ follows a generalized beta distribution $\mathbf{p} \sim GB(p_L, p_U, \mu_X = 0.05, \sigma_X^2 = 0.00226)$, where the values μ_X and σ_X^2 are derived from Equation (4) by choosing the *Beta* parameters $a = 1$ and $b = 19$.

Initially, the costs are fixed as $K_1 = K_2 = 0.1$. For $p_L = 0, p_U = 1$ and $c \in [0, 8]$, the tables in Figures 3–5 report the parameters of the sampling plans (N, n, c) and the values of the cost function that satisfy the condition $P_a(\hat{p}) < 10\%$ for $N = 100{,}500, 1000$, respectively. For each N, the optimal triad that minimizes the cost function is identified and is indicated in bold. Moreover, the probabilities of acceptance $P_a(\hat{p})$ and the corresponding values of $\hat{p}$ are calculated. Finally, the OC curves of the optimal sampling plans are presented on the right side of the figures. The lot fraction of defective items in correspondence to the red point represents the limiting quality levels which represents the average poorest level of quality that the consumer is willing to accept in an individual lot for $\beta = 10\%$.

N	n	c	$C(N,n,c)$	$P_a(\hat{p})$	$\hat{p}$
100	11	0	0.06500	0.09887	0.180
100	30	1	0.07189	0.09931	0.110
100	21	2	0.06075	0.09996	0.220
100	**11**	**3**	**0.05546**	**0.09986**	**0.500**
100	21	4	0.06035	0.09983	0.330
100	20	5	0.06030	0.09946	0.400
100	27	6	0.06397	0.09963	0.340
100	27	7	0.06429	0.09888	0.380
100	21	8	0.06144	0.09787	0.530

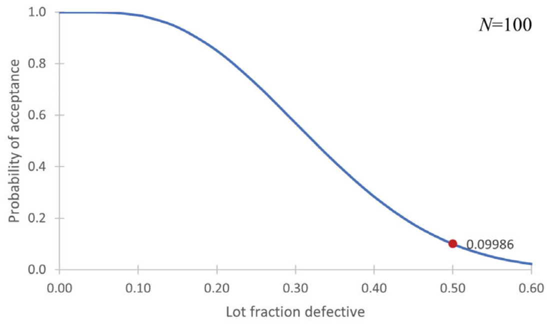

FIGURE 3 Optimal sampling plan (left) and behavior of CO (right) for $N = 100$

N	n	c	$C(N,n,c)$	$P_a(\hat{p})$	$\hat{p}$
500	112	0	0.09453	0.09991	0.018
500	52	1	0.07069	0.09995	0.070
500	80	2	0.07174	0.09943	0.062
500	56	3	0.05887	0.09998	0.112
500	44	4	0.05320	0.09997	0.170
500	**30**	**5**	**0.05143**	**0.09997**	**0.284**
500	66	6	0.05529	0.09992	0.150
500	104	7	0.06143	0.09992	0.106
500	36	8	0.05285	0.09995	0.332

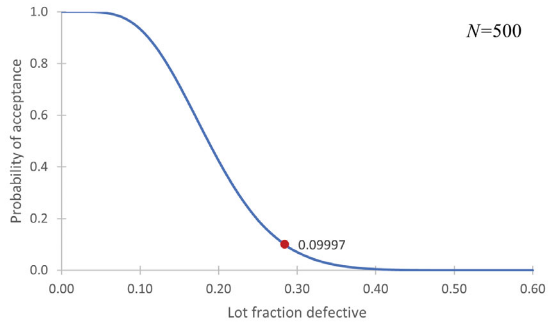

FIGURE 4 Optimal sampling plan (left) and behavior of CO (right) for $N = 500$

N	n	c	$C(N,n,c)$	$P_a(\hat{p})$	$\hat{p}$
1000	50	0	0.08220	0.09944	0.044
1000	115	1	0.08388	0.09967	0.032
1000	115	2	0.07638	0.09983	0.044
1000	135	3	0.07415	0.09997	0.047
1000	125	4	0.06776	0.09984	0.061
1000	**40**	**5**	**0.04970**	**0.09943**	**0.218**
1000	105	6	0.05770	0.09997	0.096
1000	175	7	0.06671	0.09986	0.064
1000	50	8	0.05044	0.09944	0.245

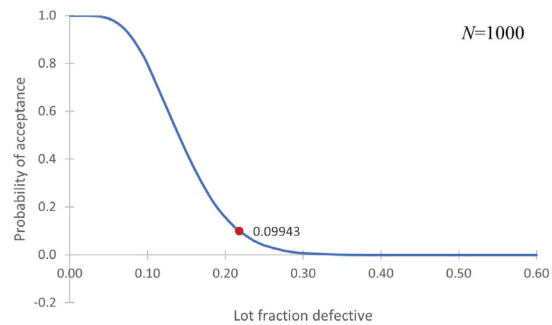

FIGURE 5 Optimal sampling plan (left) and behavior of CO (right) for $N = 1000$

To evaluate the sensitivity of the proposed procedure to the change in the main parameters, other optimal sampling plans were determined for several values of the costs regarding quality K_1 and sampling K_2, and some values of the parameters p_L and p_U characterizing the *GB* distribution. The results are listed in Table 1.

Table 1 shows that the values of the cost function depend strongly on the values of K_1 and K_2. For increasing values of $K_1 = K_2$, $C(N, n, c)$ increases. For example, the minimum cost is 0.01818 for $K_1 = K_2 = 0.01$, and it is 3.88470 for $K_1 = K_2 = 30$. Considering the values of $K_1 \neq K_2$, it seems that the sampling cost K_2 has more influence on the trend of the minimum of $C(N, n, c)$ than the quality-related cost K_1. Moreover, if the parameters of the *GB* distribution (p_L and p_U) are changed, the minimum of the cost function changes together with the optimal triad.

TABLE 1 Optimal single sampling plans based on prior *GB* distribution (*Hypothesis 1*)

N	n	c	K_1	K_2	p_L	p_U	$C(N,n,c)$	$P_a(\hat{p})$	$\hat{p}$
100	30	1	0.01	0.01	0	1	0.01818	0.09931	0.110
100	21	2	0.07	0.07	0	1	0.05056	0.09996	0.220
100	11	3	1	1	0	1	0.17072	0.09986	0.500
100	11	3	30	30	0	1	3.88470	0.09986	0.500
100	11	0	0.01	0.05	0	1	0.04032	0.09887	0.180
100	21	2	0.1	0.07	0	1	0.05126	0.09996	0.220
100	11	3	30	0.07	0	1	0.27158	0.09986	0.500
100	30	1	0.01	0.01	0.2	0.7	0.01306	0.09931	0.110
100	21	8	1	1	0.2	0.7	0.46290	0.09787	0.530
500	52	1	0.01	0.01	0	1	0.01452	0.09995	0.070
500	56	3	0.05	0.05	0	1	0.03822	0.09998	0.112
500	36	8	1	1	0	1	0.13600	0.09995	0.332
500	36	8	30	30	0	1	2.81506	0.09995	0.332
500	56	3	0.01	0.05	0	1	0.03692	0.09998	0.112
500	44	4	0.01	0.07	0	1	0.04436	0.09997	0.170
500	112	0	0.07	0.01	0	1	0.01317	0.09991	0.018
500	112	0	0.01	0.01	0.2	0.7	0.01225	0.09991	0.018
500	21	8	1	1	0.2	0.7	0.46290	0.09787	0.530
1000	115	1	0.01	0.01	0	1	0.01067	0.09967	0.032
1000	135	3	0.03	0.03	0	1	0.02602	0.09997	0.047
1000	105	6	0.07	0.07	0	1	0.04543	0.09997	0.096
1000	50	8	1	1	0	1	0.14194	0.09944	0.245
1000	50	8	30	30	0	1	3.09027	0.09944	0.245
1000	40	5	0.03	0.1	0	1	0.04873	0.09943	0.218
1000	125	4	0.1	0.03	0	1	0.02915	0.09984	0.061
1000	135	3	0.01	0.01	0.4	0.9	0.01425	0.09997	0.047
1000	50	8	1	1	0.2	0.7	1.03967	0.09944	0.245

Table 2 displays the optimal single sampling plan for the special case of *Hypothesis 1* in which $a = b = 1$ that is, $\mathbf{p}$ follows a uniform distribution in $(0, 1)$; consequently, the number of defects in the lot follows a discrete rectangular distribution, see Equation (15).

As expected, different parameters associated with the prior *GB* and, consequently, different distributions of the number of defects in the lot strongly affect the minimum cost function and thus the choice of the optimal sampling plan.

7.2 | Hypothesis 2

The number of defects in the lot follows a binomial distribution with parameters N and $\mu_{\mathbf{p}}$ as in Equation (18). The optimal sampling plans were derived for several values of the quality K_1 and sampling K_2 costs and several values of the binomial parameters $N = 100{,}500, 1000$ and $\mu_{\mathbf{p}} = 0.01, 0.1, 0.25$ (Table 3).

TABLE 2 Optimal single sampling plans based on prior rectangular distribution

N	n	c	K_1	K_2	p_L	p_U	$C(N, n, c)$	$P_a(\hat{p})$	$\hat{p}$
100	30	1	0.01	0.01	0	1	0.01664	0.09931	0.110
100	30	1	0.07	0.07	0	1	0.10381	0.09931	0.110
100	21	8	1	1	0	1	1.17682	0.09787	0.530
100	11	3	30	30	0	1	33.26709	0.09986	0.500
100	30	1	0.01	0.05	0	1	0.05484	0.09931	0.110
100	30	1	0.1	0.07	0	1	0.11874	0.09931	0.110
100	21	8	30	0.07	0	1	13.00993	0.09787	0.530
100	30	1	0.01	0.01	0.2	0.7	0.01459	0.09931	0.110
100	21	8	1	1	0.2	0.7	1.11915	0.09787	0.530
500	112	0	0.01	0.01	0	1	0.01499	0.09991	0.018
500	80	2	0.05	0.05	0	1	0.07417	0.09943	0.062
500	36	8	1	1	0	1	1.27427	0.09995	0.332
500	36	8	30	30	0	1	37.36674	0.09995	0.332
500	80	2	0.01	0.05	0	1	0.05420	0.09943	0.062
500	56	3	0.01	0.07	0	1	0.07330	0.09998	0.112
500	112	0	0.07	0.01	0	1	0.04499	0.09991	0.018
500	112	0	0.01	0.01	0.2	0.7	0.01450	0.09991	0.018
500	36	8	1	1	0.2	0.7	1.35045	0.09995	0.332
1000	115	1	0.01	0.01	0	1	0.01503	0.09967	0.032
1000	40	3	0.03	0.03	0	1	0.04757	0.09950	0.158
1000	85	4	0.07	0.07	0	1	0.10293	0.09976	0.090
1000	40	3	1	1	0	1	1.40569	0.09950	0.158
1000	40	3	30	30	0	1	42.00908	0.09950	0.158
1000	40	3	0.03	0.1	0	1	0.11102	0.09950	0.158
1000	135	3	0.1	0.03	0	1	0.07959	0.09997	0.047
1000	135	3	0.01	0.01	0.4	0.9	0.01650	0.09997	0.047
1000	40	3	1	1	0.2	0.7	1.44849	0.09950	0.158

8 | CONCLUSION

An approach to find optimal Bayesian acceptance sampling plans for attributes under the hypothesis that the process quality follows a prior generalized beta distribution was proposed. After a detailed description of the generalized beta distribution of the process quality, two hypotheses about lot formation was presented: the lot is composed of elements coming from a phase of the production cycle characterized by a constant quality value (*Hypothesis 1*), and the lot is composed of elements from different stages of the production cycle (*Hypothesis 2*). These two hypothesis lead to different distributions of defective items in a lot. The combination of the prior distribution (and consequently the distribution of defects in a lot) and the hypergeometric distribution, has lead to the compound hypergeometric distribution that describes the number of defective elements in a sample. Under these conditions, optimal sampling plans are obtained by minimizing the expected total cost of quality $C(N, n, c)$ subject to the condition that consumer risk is reduced.

TABLE 3 Optimal single sampling plans based on prior binomial distribution

N	n	c	μ_p	K_1	K_2	$C(N, n, c)$	$P_a(\hat{p})$	$\hat{p}$
100	11	3	0.01	0.01	0.01	0.01001	0.09986	0.500
100	30	1	0.1	0.01	0.01	0.02244	0.09931	0.110
100	30	1	0.25	0.01	0.01	0.01283	0.09931	0.110
100	11	3	0.01	0.1	0.07	0.01671	0.09986	0.500
100	30	1	0.1	0.1	0.07	0.08257	0.09931	0.110
100	30	1	0.25	0.1	0.07	0.09521	0.09931	0.110
500	30	5	0.01	0.01	0.01	0.01001	0.09997	0.284
500	112	0	0.1	0.01	0.01	0.01100	0.09991	0.018
500	112	0	0.25	0.01	0.01	0.01250	0.09991	0.018
500	30	5	0.01	0.01	0.07	0.01361	0.09997	0.284
500	112	0	0.1	0.01	0.07	0.07100	0.09991	0.018
500	112	0	0.25	0.01	0.07	0.07250	0.09991	0.018
1000	40	5	0.01	0.01	0.01	0.01000	0.09943	0.218
1000	115	1	0.1	0.01	0.01	0.01101	0.09967	0.032
1000	175	7	0.25	0.01	0.01	0.01250	0.09986	0.064
1000	40	5	0.01	0.1	0.03	0.01084	0.09943	0.218
1000	115	1	0.1	0.1	0.03	0.04000	0.09967	0.032
1000	175	7	0.25	0.1	0.03	0.05500	0.09986	0.064

Given the many parameters involved, an extensive simulation study was conducted to show how the optimal design of the experiment varies according to the change of the hypotheses about lot formation and the main parameters. This study confirms that our proposal turns out to be an automatic procedure for determining the optimal sampling plan following an economical criteria. In fact, it provides a mechanism so that the sampling plans would be automatically adjusted to changes in the prior generalized beta distribution related to the characteristics of the considered productive process and/or the information from previous inspection results.

In summary, for quality controllers, to select the optimal sampling plan, we suggest the following steps:

1. choosing a prior distribution of the productive process quality $f_{\mathbf{p}}(p)$ (in this case generalized beta distribution),
2. choosing between *Hypotheses 1* and *2* about lot formation,
3. setting the lot size N and the parameters of the distribution function $r_{\mathbf{D}}(D)$ of defective items in the lot,
4. deriving the distribution law of defective elements in the sample,
5. defining the cost function and setting the cost parameters (K_1, K_2),
6. deriving the optimal sampling plan (N, n, c) as a result of the conditioned optimization of the cost function for a fixed consumer risk β.

The further research will be devoted to the design of sampling plans and to evaluate their effects on the sampling and inspection cost levels considering more than one stage (double and multiple plans). These sampling methods are based on the assumption that on the first inspection the sample did not produce the necessary support to take a decision about the lot. The principal advantage of double and multiple sampling plans with respect to single ones is that they may reduce the total amount of required inspection nevertheless the high complexity of these plans.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

ACKNOWLEDGMENT

Open Access Funding provided by Universita Cattolica del Sacro Cuore within the CRUI-CARE Agreement.

ORCID

Silvia Facchinetti https://orcid.org/0000-0002-3753-0406
Silvia Angela Osmetti https://orcid.org/0000-0002-0321-5731

REFERENCES

1. Montgomery DC. *Introduction to Statistical Quality Control*. 8th ed. John Wiley & Sons; 2019.
2. Luca S, Vandercappellen J, Claes J. A web-based tool to design and analyze single- and double-stage acceptance sampling plans. *Qual Eng*. 2020;32:58-74.
3. Ali S, Riaz M. On designing a new Bayesian dispersion chart for process monitoring. *Arab J Sci Eng*. 2019;45:2093-2111.
4. Aslam M, Anwar SM. An improved Bayesian modified-EWMA location chart and its applications in mechanical and sport industry. *PLoS One*. 2020;15:e0229422.
5. Menzefricke U. Combined exponentially weighted moving average charts for the mean and variance based on the predictive distribution. *Commun Stat-Theory Methods*. 2013;42:4003-4016.
6. Saghir A. Phase-I design scheme for -chart based on posterior distribution. *Commun Stat—Theory Methods*. 2015;44:644-655.
7. Biswas S. *Statistics of Quality Control*. New International Pvt, Ltd; 1975.
8. Balamurali S, Park H, Jun C-H, Kim K-J, Lee J. Designing of variables repetitive group sampling plan involving minimum average sample number. *Commun Stat-Simul Comput*. 2005;34:799-809.
9. Moskowitz H, Tang K. Bayesian variables acceptance-sampling plans: quadratic loss function and step-loss function. *Technometrics*. 1992;34:340-347.
10. Sankle R, Singh JR. Single sampling plans for variables indexed by AQL and AOQL with measurement error. *J Modern Appl Stat Methods*. 2012;11:396-406.
11. Brush GG. A comparison of classical and Bayes producer's risk. *Technometrics*. 1986;28:69-72.
12. Sharma KK, Bhuttani RK. A comparison of classical and Bayes risks when quality varies randomly. *Microelectron Reliab*. 1992;32:493-495.
13. Cobb B, Li L. Bayesian networks for statistical process control with attribute data. *Int J Qual Reliab Manag*. 2018;36:232-256.
14. Ouyang L, Chen J, Ma Y, Park C, Jin J. Bayesian closed-loop robust process design considering model uncertainty and data quality. *IISE Trans*. 2020;52:288-230.
15. Hald A. The compound hypergeometric distribution and a system of single sampling inspection plans based on prior distributions and costs. *Technometrics*. 1960;2:275-340.
16. Hald A. Statistical Concepts and models in the theory of sampling inspection by attributes. *Nordisk Tidsskrift for Industriel Statistik*. 1960;5:144-168.
17. Hald A. Bayesian single sampling attribute plans for continuous prior distributions. *Technometrics*. 1968;10:667-683.
18. Yuvaraj V, Sakthimurugan L, Sivaranjani P. Comparative study of traditional and bayesian single sampling plan for attributes. *Int J Inf Comput Sci*. 2019;6:522-529.
19. GonzÁLez C, Palomo G. Bayesian acceptance sampling plans following economic criteria: an application to paper pulp manufacturing. *J Appl Stat*. 2003;30:319-333.
20. Yuvaraj V, Subramaniam R. Bayesian single sampling plan for attributes based on beta prior distribution. *Int J Stat Anal*. 2011;1:119-126.
21. Hemalatha TR, Iyappan M. Bayesian acceptance sampling plans through numerical methods. *Int J Math Stat Invent*. 2015;3:11-19.
22. Nezhad MSF, Saredorahi FZ. A new policy for designing acceptance sampling plan based on Bayesian inference in the presence of inspection errors. *Commun Stat Theory Methods*. 2018;47:5307-5318.
23. Kaviyarasu V, Sivakumar P. Implication of production and monitoring techniques in Bayesian single sampling plan using gamma-zero inflated Poisson distribution. *Int J Recent Technol Eng*. 2019;8:10110-10119.
24. Wadsworth HM, Stephens KS, Godfrey AB. *Modern Methods for Quality Control and Improvement*. 2nd ed. John Wiley & Sons; 2002.
25. Suresh KK, Umamaheswari S. Designing of Bayesian skip lot sampling plan under destructive testing. *J Modern Appl Stat Methods*. 2016;15:703-722.
26. Vijayaraghavan R, Loganathan A, Rajagopal K. Selection of Bayesian single sampling plans by attributes with desired discrimination. *Econ Qual Control*. 2013;28:167-177.
27. Calvin TW. The ASQC basic references in quality control: statistical techniques. *How and When to Perform Bayesian Acceptance Sampling*. Vol 7. American Society for Quality Control; 1983.
28. Johnson NL, Kotz S, Balakrishnan N. *Continuous Univariate Distributions*. Vol 2. 2nd ed. John Wiley & Sons; 1995.
29. Irony TZ, Pereira CAB. Motivation for the use of discrete distributions in quality assurance. *Test*. 1994;3:181-193.
30. Dodge HF, Romig HG. *Sampling Inspection Tables: Single and Double Sampling*. John Wiley & Sons; 1949.

How to cite this article: Facchinetti S, Osmetti SA, Magagnoli U. Designing acceptance single sampling plans: An optimization-based approach under generalized beta distribution. *Appl Stochastic Models Bus Ind*. 2022;1-17. doi: 10.1002/asmb.2710