
Sampling in Manufacturing

Chapters 1 to 6 of Volume 1 described various methods for building maturity. However, from a manufacturing perspective, these methods must be cost-effective. One of the solutions that can be considered to reduce costs is to test less than 100% of the products before delivery to the system manufacturer. This is called sampling.

As expected, there are various standards dealing with this subject, such as ISO 28590. These standards clarify the sampling rules to be applied, and the interested reader is invited to read them for further details.

However, the standards do not cover several aspects that are very important for the manufacturer:

- The cost aspect, which leads to the following questions:
 - What is the benefit of applying a sampling rule?
 - Is a sampling rule adapted for my application?
 - What rule should I use to minimize costs?
- What is the impact of test coverage rates if they are not 100%?
- What is the impact of considering a distribution of potential defects?

This chapter aims to suggest a solution for each of these cases in order to formulate optimum sampling in terms of quality and cost.

Theoretically speaking, sampling techniques rely on discrete probability distributions (the random variable can only take certain values), unlike the probability distributions for estimating the reliability of a failure mechanism, which are continuous (e.g. exponential, Weibull, etc.). The Bernoulli distribution is used for the result of a test (failure or success). When this test is repeated several times, two cases are possible.

- The “draw” is **unrestricted**, and in this case the binomial distribution is applicable.

- The “draw” is **restricted**, and in this case the hypergeometric distribution is applicable.

Sampling is obviously a “restricted” draw. Therefore, the theoretical basis of the sampling norms is the hypergeometric distribution.

1.1. Cost aspects

The function of “cost associated with a size n sampling” is a random variable (the batch is accepted or rejected). Therefore, its average value (mathematical expectation) is considered here.

Given:

- C is the cost per unit of the non-compliance test;
- K is the cost per unit of accepting a non-compliant product in a batch;
- N is the number of products to be tested;
- X is the number of “non-compliant” products in the batch;
- n is the size of the sample.

The average cost is equal to the cost of compliant products plus the cost of non-compliant products. The cost of compliant products (if the batch was accepted) is given by:

$$C1(n, N, X, C, K) = (C \cdot n + X \cdot K) \cdot P_1 \quad [1.1]$$

P₁ is the probability of having “no non-compliant parts” in a size “n” sample, randomly drawn from a batch of size “N”. Therefore it is:

$$P_1 = \frac{C_{N-x}^n}{C_N^n} \quad [1.2]$$

where C_y^x is the number of combinations of x out of y elements.

It is worth considering the details of this result. The probability of an event can be estimated by the ratio of the number of possible cases to the total number of cases. It is clear that the total number of cases is equal to the number of combinations of “n” taken out of “N”, or C_N^n .

The number of possible cases is equal to the number of combinations of compliant parts, or $N-X$, and is therefore equal to C_{N-X}^n , hence the result of equation [1.2].

The cost of non-compliant products (if the batch was rejected) is given by:

$$C2(n, N, X, C, K) = C \cdot N \cdot (1 - P1) \quad [1.3]$$

Based on equations [1.1], [1.2] and [1.3], the total cost is:

$$C(n, N, X, C, K) = (C \cdot n + X \cdot K) \cdot \frac{C_{N-X}^n}{C_N^n} + C \cdot N \cdot \left(1 - \frac{C_{N-X}^n}{C_N^n}\right) \quad [1.4]$$

As an illustration, assume that:

- there are 100 products to be delivered $\rightarrow N = 100$;
- the cost of a compliant part is $C = 30\text{€}$;
- the cost of a non-compliant part is $K = 1,500\text{€}$.

According to this data, the cost of a non-compliant part is very high compared to the cost of a compliant part. Let us now consider the evolution of the total average cost depending on the size (n) and the number of defective products (X) (see Figure 1.1).

It can be noted that starting with $X > 1$, the sampling rate of 100% is optimal. In this example, the sampling is not interesting in terms of cost.

Now assume that:

- there are 100 products to be delivered $\rightarrow N = 100$;
- the cost of a compliant part is $C = 1,000\text{€}$;
- the cost of a non-compliant part is $K = 1,500\text{€}$.

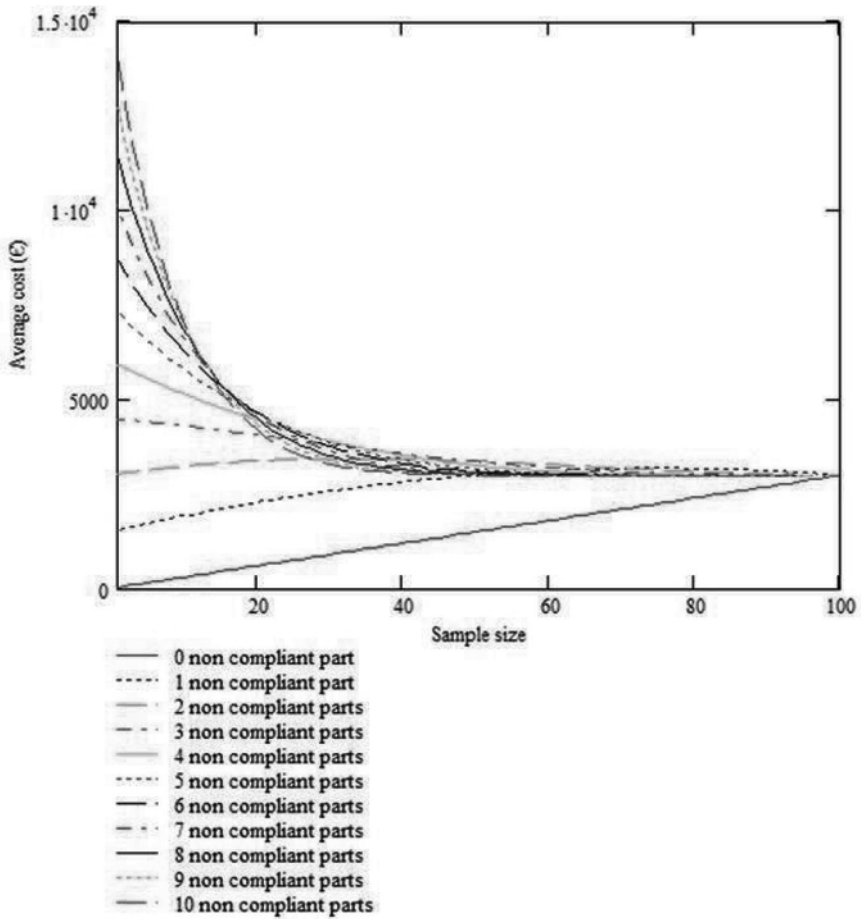


Figure 1.1. Evolution of the total average cost depending on the size of the sample and the number of defective products

Although low, the cost of a compliant part is of the order of the cost of non-compliant parts. Let us now consider the evolution of the total average cost depending on the sample size (n) and the number of defective products (X) (see Figure 1.2).

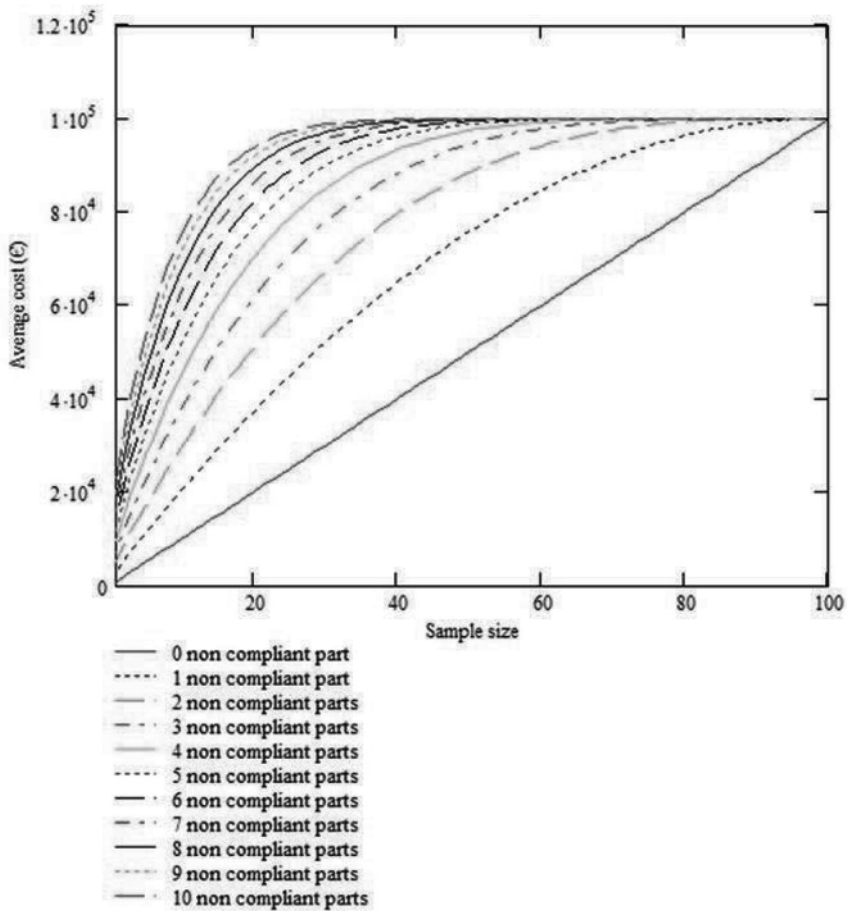


Figure 1.2. Evolution of the total average cost depending on the size of the sample and the number of defective products

In contrast to the previous example, in this case the sampling is better than the test at 100%.

Now assume that:

- there are 100 products to be delivered $\rightarrow N = 100$;
- the cost of a compliant part is $C = 100\text{€}$;
- the cost of a non-compliant part is $K = 1,500\text{€}$.

Let us now consider the evolution of the total average cost depending on the sample size (n) and the number of defective products (X) (see Figure 1.3).

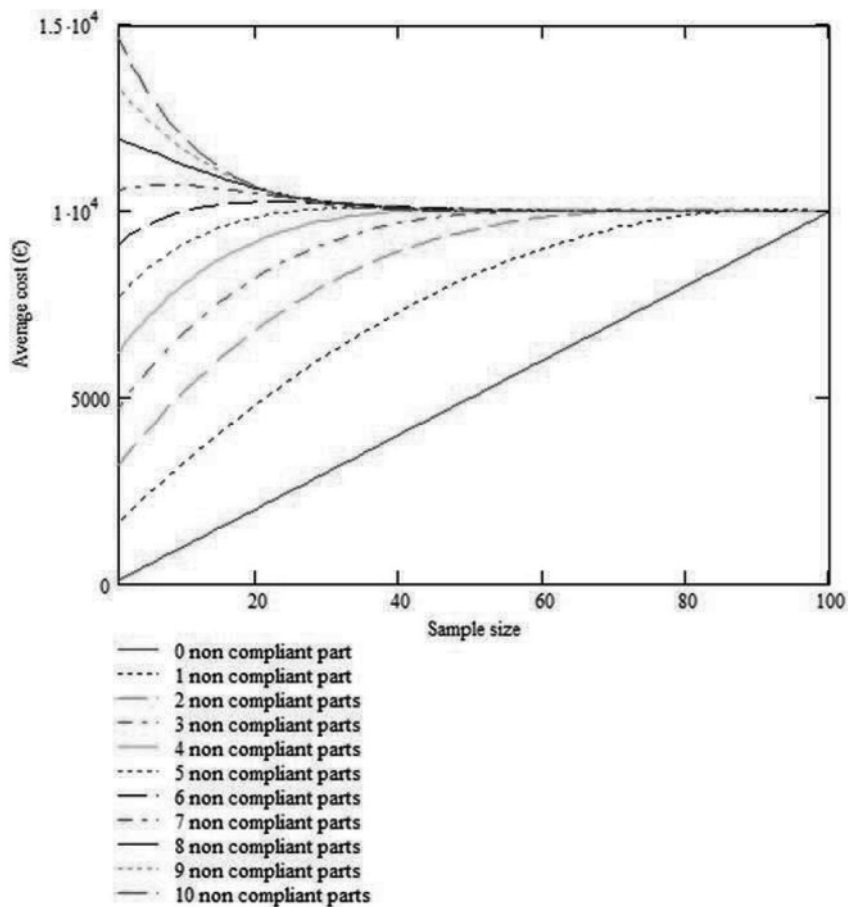


Figure 1.3. Evolution of the total average cost depending on the size of the sample and the number of defective products

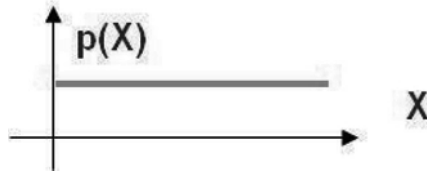
This situation is an intermediate one between the two previously mentioned examples, as the sampling is optimal when the number of defects is equal to or greater than 7.

1.2. Considering the distribution of defects

Given $p(X)$, the probability density of the defects observed during this phase of the test, the mathematical expectation of the cost is:

$$C(n, N, C, K) = \sum_{X=0}^N p(x) \cdot \left[(C \cdot n + X \cdot K) \cdot \frac{C_N^R - X}{C_N^R} + C \cdot N \cdot \left(1 - \frac{C_N^R - X}{C_N^R} \right) \right] \quad [1.5]$$

If this distribution is not known, the uniform probability distribution can be used:



Given that the number of non-compliant products ranges between 0 and N , the probability density is then:

$$p(x) = \frac{1}{N+1} \quad [1.6]$$

Using equations [10.5] and [10.6], the mathematical expectation of the cost is given by:

$$C(n, N, C, K) = \sum_{X=0}^N \left(\frac{1}{N+1} \right) \cdot \left[(C \cdot n + X \cdot K) \cdot \frac{C_N^R - X}{C_N^R} + C \cdot N \cdot \left(1 - \frac{C_N^R - X}{C_N^R} \right) \right] \quad [1.7]$$

Once again, as an illustration, let us resume the following example and assume that:

- there are 100 products to be delivered $\rightarrow N = 100$;
- the cost of a compliant part is $C = 30\text{€}$;
- the cost of a non-compliant part is $K = 1,500\text{€}$.

Let us now consider the evolution of the total average cost depending on the sample size (n) (see Figure 1.4).

It can be noted that for a small sample size, the “average” cost can be very high.

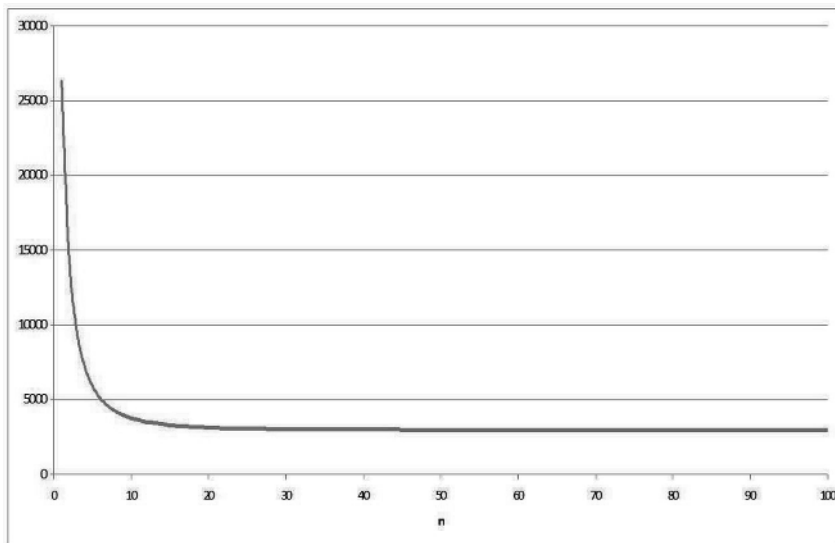


Figure 1.4. Evolution of the total average cost depending on the sample size – Example 1

Assume that:

- there are 100 products to be delivered $\rightarrow N = 100$;
- the cost of a compliant part is $C = 100\text{€}$;
- the cost of a non-compliant part is $K = 1,500\text{€}$.

Let us now examine the evolution of the total average cost depending on the sample size (n) (see Figure 1.5).

It can be noted that for a small sample size, the “average” cost can be very high. However, a slight optimum is obtained for $n = 20$, or a sampling rate of 20%, in this example.

Assume that:

- there are 100 products to be delivered $\rightarrow N = 10$;
- the cost of a compliant part is $C = 1,000\text{€}$;
- the cost of a non-compliant part is $K = 1,500\text{€}$.

Let us now examine the evolution of the total average cost depending on the sample size (n) (see Figure 1.6).

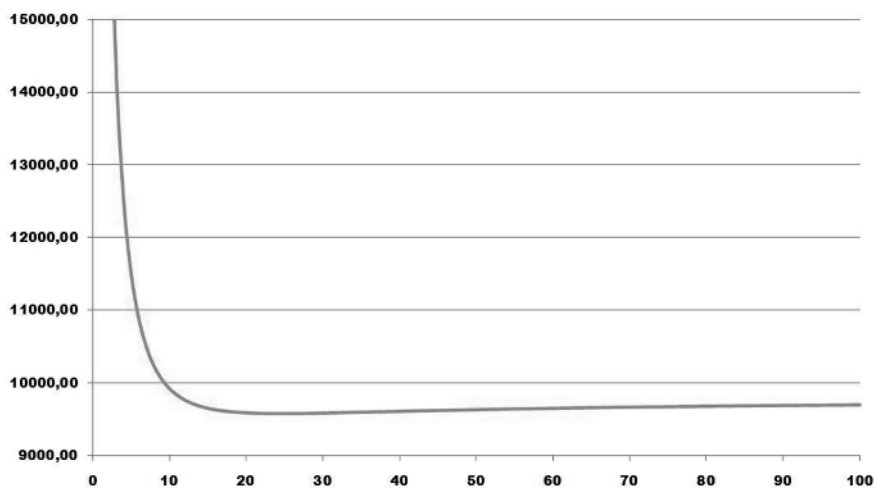


Figure 1.5. Evolution of the total average cost depending on the sample size – Example 2

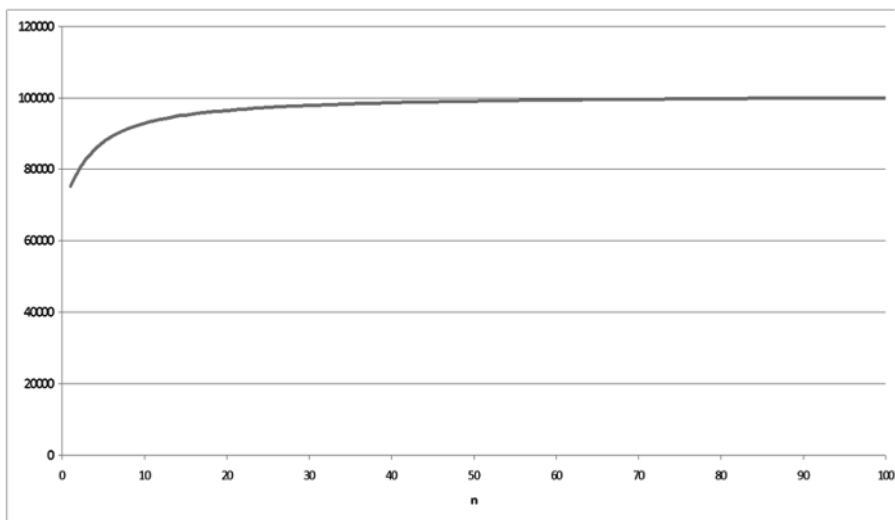


Figure 1.6. Evolution of the total average cost depending on the sample size – Example 3

It can be noted that for a small sample size, the “average” cost is the lowest.

1.3. Considering the test coverage

Given P_t , the probability of detecting a non-compliant product for a test, knowing that it will be detected by the client, let us denote by p_{nMNt} the probability of having M “non-compliant” parts in a sample of size n , drawn from a batch of size N , in which X parts are non-compliant.

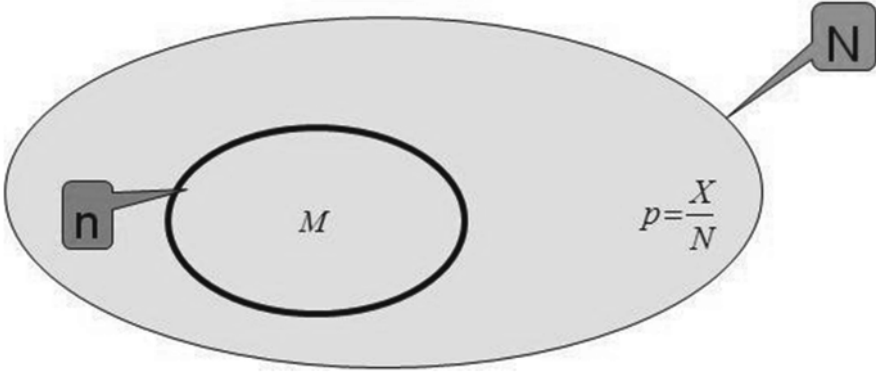


Figure 1.7. Illustration

Here, the number of possible cases is the number of possibilities of drawing M defective parts from X and drawing “ $n-M$ ” compliant parts from the “ $N-X$ ” compliant parts of the batch. The probability p_{nMNt} is then written as (hypergeometric distribution):

$$P_{nMNt} = \frac{C_X^M \cdot C_{N-X}^{n-M}}{C_N^n} \quad [1.8]$$

where C_X^M is the number of possibilities of having defective products among the X products from the batch of size N , C_{N-X}^{n-M} is the number of possibilities of having compliant products in the size n sample among the compliant parts of the batch, and C_N^n is the number of possibilities of having n products among the N products.

The probability of detecting no non-compliant products in the M non-compliant products drawn in the test sampling is:

$$P_{n0MNt} = (1 - p_t)^M \quad [1.9]$$

Finally, let us denote by p_{n0NXt} the probability of having no non-compliant parts detected in a sample of size n , drawn in a batch of size N , in which X parts are non-compliant. Hence, according to the formula of total probabilities:

$$P_{n0NXt} = \sum_{M=0}^{\min(n,X)} [P_{nMNXt} \cdot P_{n0MNXt}] \quad [1.10]$$

or, based on equations [1.9] and [1.10]:

$$P_{n0NXt} = \sum_{M=0}^{\min(n,X)} \left[\frac{C_X^M \cdot C_{N-X}^{n-M}}{C_N^n} \cdot (1 - p_t)^M \right] \quad [1.11]$$

Let us denote by X_d the number of parts detected as non-compliant when all the parts are tested following the rejection of the batch, as a result of the detection of at least one non-compliant part in a sample of size n . The “cost” function associated with a sample of size n , knowing that X_d non-compliant parts were detected, is defined as follows:

$$C(n, N, X, C, K, t) = (C \cdot n + X \cdot K) \cdot P_{n0NXt} + [C \cdot N + (X - X_d) \cdot K] \cdot (1 - P_{n0NXt}) \quad [1.12]$$

Since X_d follows a binomial law and $E[X_d] = X \cdot p_t$, the average cost depending on X is:

$$C(n, N, X, C, K, t) = (C \cdot n + X \cdot K) \cdot P_{n0NXt} + [C \cdot N + X \cdot (1 - p_t) \cdot K] \cdot (1 - P_{n0NXt}) \quad [1.13]$$

where $(1 - p_t)$ is the additional cost due to the lack of detection capacity of tests p_t . The cost depends on X number of non-compliant products in the batch, unknown by definition. Given $p(X)$, the probability law of X for $X = 0$ to N , the expectation of the total cost of a test of n parts is therefore:

$$C(n, N, C, K, t) = \sum_{X=0}^N p(X) \cdot \left[(C \cdot n + X \cdot K) \cdot \sum_{M=0}^{\min(n,X)} \left[\frac{C_X^M \cdot C_{N-X}^{n-M}}{C_N^n} \cdot (1 - p_t)^M \right] + [C \cdot N + X \cdot (1 - p_t) \cdot K] \cdot \left(1 - \sum_{M=0}^{\min(n,X)} \left[\frac{C_X^M \cdot C_{N-X}^{n-M}}{C_N^n} \cdot (1 - p_t)^M \right] \right) \right]$$

