603801

SAMPLING INSPECTION PLANS
FOR CONTINUOUS PRODUCTION

M. A. Girshick

P-17

May 10, 1948

The RAND Corporation
1700 MAIN ST • SANTA MONICA • CALIFORNIA

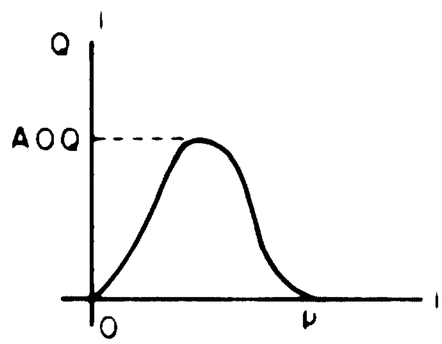
SAMPLING INSPECTION PLANS FOR CONTINUOUS PRODUCTION

The sampling inspection plans [redacted] discussed [redacted] are applicable only to a production process in which the units of product are classified as either defective or non-defective. These plans are designed to set a prescribed upper limit to the average outgoing quality of the final product entering into consumption channels. This is accomplished through a sequential process of partial and complete inspection of the finished product which will be described in detail later.

Sampling inspection schemes whose aim is to improve the outgoing quality of the finished product is not new in the field of industrial quality control. Various methods are in vogue with varying degrees of statistical validity.

The first step towards putting the use of sampling inspection for controlling the outgoing quality on a scientific basis was taken by H. F. Dodge and H. G. Romig of the Bell Telephone system. Dodge and Romig made their inspection plan an adjunct to lot-by-lot acceptance inspection. The basic idea involved in their scheme ~~can be~~ is briefly described, [redacted]

In lot-by-lot acceptance inspection a criterion is given for judging the quality of a lot on the basis of a random sample selected from it. The application of this criterion results either in the acceptance or the rejection of the lot. If the lot is accepted it goes into consumption channels as is. If it is rejected, however, it is first inspected 100% and the defective items removed and replaced by non-defective items before it is released for consumption. Thus, since rejected lots are presumably free of defectives after the 100 inspection, the average outgoing quality of the product is better than the incoming quality. In fact, it can be seen that given any reasonable lot-by-lot inspection scheme, the worse the incoming quality, the better will be the average outgoing quality when this inspection procedure is employed. This phenomenon can be illustrated by plotting in a graph the average outgoing quality Q (known as the AOQ) against the lot fraction defective p. This curve will usually have the following shape.



The value of Q at the peak of this curve is known as the AOQL (the average outgoing quality limit). Thus the AOQL represents the worst possible outgoing quality which can be expected on the average no matter what the incoming lot quality is.

The above method for controlling the outgoing quality of the finished product through sampling inspection has several limitations which it would be desirable to remove. In the first place, the plan is applicable to products which are formed into lots and the sampling which it presupposes is a random selection from the lot. However, it is often desirable to make the sampling inspection from the sequence of goods as they are produced and this their plan does not provide for. Another limitation of this plan is the high cost of inspection which it involves. It will be noted that the more the lot fraction defective exceeds the AOQL, the smaller the AOQ. But this is attained at a high cost in inspection which is inefficient if a given AOQL can be tolerated.

In 1943 Dodge published in the Annals of Mathematical Statistics a very interesting inspection plan which is applicable to a production flow without any regard to lot sizes but which, however, requires that the production process be in a state of statistical control. Unfortunately the limitation of time will not permit me to describe this plan. My main discussion will revolve around a class of inspection plans which were developed by A. Wald and J. Wolfowitz and which were published in the Annals of Mathematical Statistics in 1945. I shall also touch upon some original research in this field which I had undertaken while employed in the Bureau of the Census, Washington, D. C.

The sampling plans which I shall describe can be applied to a sequence of products as they are produced and they guarantee that the fraction defective remaining in the product after inspection will not exceed, on the average, a preassigned limit regardless of the size of the fraction defective in the product prior to inspection and regardless of whether or not the production process is in a state of statistical control.

The mechanical procedures involved in the sampling plans are not complicated. The sequence of items produced are divided into seqments of size k. One observation, chosen at random from each segment, in sequence is inspected and the number of defective items found is cumulated. This operation is known as partial inspection. A rule (usually based on the cumulative number of observations) is given for terminating inspection. At the termination of inspection, a specified number (which might be zero) of units of product are inspected 100 per cent. This operation is known as complete inspection. The inspection procedure is then repeated on new material. It is, of course, assumed that during both partial and complete inspection, the defective items found are removed and replaced by non-defective items.

The statistical principles on which these plans are based, are also fairly simple. They can be briefly formulated as follows: Let x_1, x_2, ..., be the sequence of observations obtained from the first, second, etc., segment partially inspected from the beginning of an inspection operation. Here $x_i = 0$ if the item in the i-th segment was non-defective and $x_i = 1$ if the item was defective. At the n-th stage of partial inspection we consider an estimate $\hat{p}$ of the fraction defective in the product which has passed through inspection from the beginning of the inspection operation. This estimate is defined as

$$\hat{p} = \frac{(k-1)\sum_{i=1}^{n} x_i}{kn}$$

As long as partial inspection continues, this estimate is restricted by the plan to satisfy the inequality

$$\hat{p} \leq U + \phi(n)$$

where U is the desired AOQL and $\phi(n)$ is a nonnegative function of n which approaches zero as n approaches infinity. When inspection terminates, as specified by the rules of the plan, this inequality may no longer hold. However, in that case, enough additional items are inspected 100 per cent and the defectives replaced so as to make the estimate of the total less than (or equal to) U. It is not difficult to prove by employing the Strong Law of Large Numbers that any plan which keeps the above estimate bounded (as above), also guarantees that the AOQL will not be exceeded in the long run regardless of whether or not the production process is in a state of statistical control. Thus, stability in the production process is not a necessary requirement for the attainment of a given average outgoing quality limit. However, as is the case with most statistical procedures, achieving an average result is only part of the goal. What is often just as important is achieving small variability around that average. I shall return to this important question later.

Due to lack of time I shall restrict myself to a discussion of only a few of the plans which might be devised to achieve a given AOQL. The simplest one is the Fixed Lot Size Plan. Here the production sequence is assumed to be divided into lots of size M and each lot is further subdivided into N segments of size k. An integer $m < N$ is selected. Inspection begins by selecting at random one item from each consecutive segment of k items. These items are inspected in sequence and the number of defective items found is cumulated. Inspection terminates as soon as m defectives are found or alternatively all the N segments have been partially inspected and fewer than m defectives have been found. If the number of defectives found is m with a sample size $n < N$, then the remaining $N - n$ segments are inspected 100 per cent. Otherwise no 100 per cent inspection is required. This inspection procedure is then repeated on a new lot. It can be shown that the expected fraction of defective items left in the lots thus inspected is less than or equal to

$$U = \frac{k-1}{k}\,\frac{m}{N}\,.$$

The proof of this is short and it may be instructive to examine it here. Let $x_1, x_2, \ldots, (x_i = 0 \text{ or } x_i = 1)$ represent the quality of items partially inspected from the first, second, etc., segment and let D_i represent the number of defectives left in the i-th segment after partial inspection. Moreover, let p_i represent the fraction defective in the i-th segment before inspection. Then, since $Ex_i = p_i$ and $x_i + D_i = kp_i$ we have $E(k-1)x_i = ED_i$. Let $z_i = (k-1)x_i - D_i$. We write

$$(1) \qquad \sum_{i=1}^{N} z_i = \sum_{i=1}^{n} z_i + \sum_{i=n+1}^{N} z_i$$

where n is a random variable and represents the sample size at which partial inspection terminates. Taking expectations on both sides of the above equation we get

(2) $$E \sum_{i=1}^{n} z_i + E \sum_{i=n+1}^{N} = 0$$

But if in the second expression to the left we take the expectation first for a fixed n and then with respect to n, we see that this expression is zero also. Hence we have

(3) $$(k-1)\; E \sum_{i=1}^{n} x_i = E \sum_{i=1}^{n} D_i$$

or

(4) $$\frac{k-1}{kN}\; E \sum_{i=1}^{n} x_i = \frac{1}{kN} \left[E \sum_{i=1}^{n} D_i \right]$$

But at the termination of inspection $\sum_{i=1}^{n} x_i \leq m$. This proves the theorem.

If the production process is not in a state of statistical control and if the items in the lot are not shuffled prior to inspection, then the AOQ, which is given by

$$\frac{k-1}{kN}\; E \sum_{i=1}^{N} x_i \; ,$$

will be a function of all the individual p_i's representing the fractions defective in each segment of the lot. However, it is of interest to note that the AOQ is a symmetric function of the p_i's so that as long as the segments are held intact, it is immaterial in what order they are presented for inspection as far as the value of the AOQ is concerned.

We have seen that the maximum value that the AOQ can have is $(k-1/k)(m/N)$. We might inquire, given a fixed average lot quality p, for what values of the p_i's will the AOQ be a minimum? The answer is; that AOQ is a minimum when all the p_i's are equal. But this case reduces itself to the Bernoulli case, i.e., to the case where the production is in a state of statistical control.

For controlled production, the AOQ can be obtained as follows: Let p represent the fraction defective inherent in the production process and L(p) represent the probability that partial inspection will terminate with fewer than m defectives. Then

$$(5) \qquad L(p) = \sum_{j=0}^{m-1} C_j^N p^j (1-p)^{N-j}$$

and

$$(6) \quad AOQ = Q(p) = \frac{k-1}{kN} E \sum_{i=1}^{n} x_i = \frac{k-1}{kN} \left[L_p E_1 \sum_{i=1}^{n} x_i + (1-L_p) E_2 \sum_{i=1}^{n} x_i \right]$$

where E_1 represents the expectation under the condition that fewer than m defectives are found and E_2 represents the expectation under the condition that exactly m defectives are found during partial inspection. Now

$$(7) \qquad E_2 \sum_{i=1}^{n} x_i = m$$

and

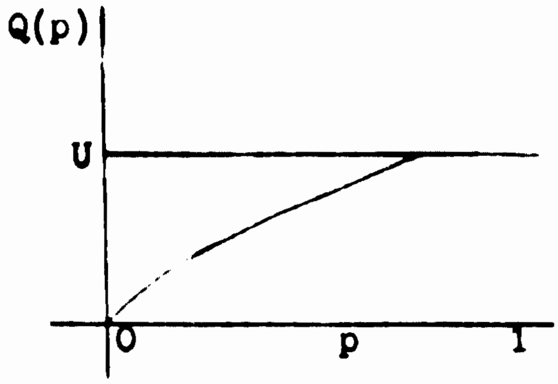
$$(8) \qquad \mathrm{LpE}_1 \sum_{i=1}^{n} x_i = \sum_{j=0}^{m-1} jC_j^N \, p^j (1-p)^{N-j} \; .$$

Hence

$$(9) \qquad Q(p) = \frac{k-1}{kN}\left[\sum_{j=0}^{m-1} jC_j^N \, p^j(1-p)^{N-j} + m\left(1 - \sum_{j=0}^{m-1} C_j^N\right)\right]$$

$$= \frac{k-1}{k}\,\frac{m}{N}\left[1 - \frac{1}{m}\sum_{j=0}^{m-1}(m-j)C_j^N p^j (1-p)^{N-j}\right]$$

when $Q(p)$ is plotted against p, the resulting curve has the following shape.



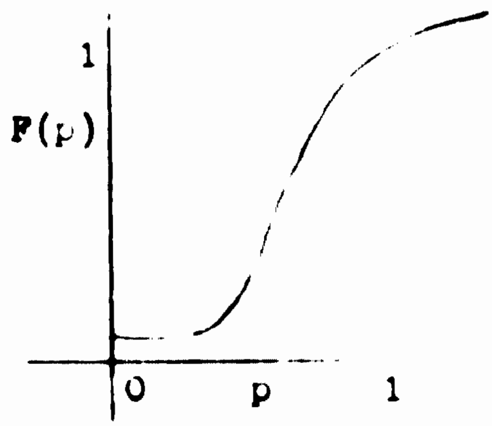
In addition to the AOQ curve, there is another curve which measures the fraction of the material inspected (AFI) and which I shall designate by $F(p)$. The $F(p)$ curve is obtained from the relationship

$$(10) \qquad p\left[1 - F(p)\right] = Q(p)$$

or

$$(11) \qquad F(p) = 1 - \frac{Q(p)}{p} \; .$$

The AFI curve when plotted against p has the following shape.



From the above discussion we see that the average fraction defective left in the lots after inspection cannot exceed $(k-1/k)(m/N)$ no matter how capricious the production process may be. To quote Wald and Wolfowitz: "If these schemes are employed, then, even if Maxwell's demon of gas theory fame were to transfer his activities to the production process, he would be unsuccessful in an effort to cause the AOQL to be exceeded." But if we were to consider this inspection scheme as a contest between the statistician and the demon, the demon could still outwit the statistician in a very important respect. He will have to grant the AOQL to the statistician, but by arranging the defectives in the sequence of product in a particular manner, he can make the quality of the outgoing material so highly variable and spotty as to nullify to a large extent the effect of the long run guarantee of acceptable quality. It is easy to show that if p_i represents the fraction defective in the i-th segment of the lot in order in which it is partially inspected, then the variability of the outgoing quality will be a maximum, if the demon arranges the segments in an increasing order of magnitude of the p_i's.

The reason why we can conceive of the demon having the freedom of arrangement of the defectives is due to the manner in which we perform the partial sampling. If during partial sampling the observations were obtained at random from the whole lot instead of selecting one at random from each segment, the demon could be thwarted even in this respect. However, even then, the problem of spotty material would remain. In fact this problem is serious even if the production process is assumed to be in a state of statistical control. However, there is good reason to believe that with further research this problem will be satisfactorily solved.

If we examine the particular plan under consideration (and the same situation holds for the other plans which I shall discuss later) it is defined by 3 constants, k, m, and N. However, the only quantity that is fixed by design is the AOQL which is essentially a function of the ratio of m to N. We see then that we have sufficient freedom in selecting these constants. Thus in addition to requiring a given AOQL, we might impose other conditions. For example, if the production process is in control, we might require that the variance of the outgoing quality ($\hat{Q}$) from lot to lot shall not exceed a given value. The variance in this case is given by

$$(12) \qquad \sigma^2_{\hat{Q}} = \frac{pq(k-1)E(n) + (k-1)^2p^2\sigma^2_n}{N^2}$$

where $E(n)$ and σ^2_n can be computed without difficulty.

However, even if the production process is not in control, various reasonable models of the possible distribution of defectives among the segments of the lots may be constructed and either the variance or an upper limit for the variance of $\hat{Q}$ obtained as a function of the parameters k, m, and N. For example, one model which appears reasonable is to assume that the defective items produced by the machine tend to occur in clusters but that the position of the clusters in the lot is random. The degree of clustering within the segments into which the lot is divided may be measured by the intraclass correlation.

$$(13) \qquad \rho = \frac{\sum_{i=1}^{N}(p_i-p)^2 - \frac{1}{k-1}\sum_{i=1}^{N}p_i q_i}{Npq}$$

where the p_i's are the fraction defective in the segments and p is the average fraction defective in the lot. When each p_i is either 0 or 1, $\rho = 1$ and the variance of $\hat{Q}$ is that case is zero. When all the p_i's are equal, ρ takes on its minimum value $-\frac{1}{k-1}$. But in that case the variance of $\hat{Q}$ is given by the expression (12). An important problem which is yet to be solved is this. For what value of ρ is the variance of $\hat{Q}$ a maximum? My guess is, that this occurs when ρ is a minimum. If this were true then we would have a convenient upper bound for the variability of the outgoing quality for this particular model.

Another inspection plan which I should like to mention briefly is one which is currently in use in the Bureau of the

Census for controlling the error rate in the coding and punch card operations. It is similar to the above plan except that the notion of a lot is dispensed with. It is defined, as above, by the three integers m, N, and k. The plan operates as follows: The units of product in the production sequence are divided into segments of size k. Inspection begins by selecting at random one item from each consecutive segment of k items. The items are inspected in sequence and the number of defectives found as well as the number of items examined is cumulated. Inspection terminates when, and only when, the cumulative number of defectives reaches m. At this point, the size of the sample n is compared with the integer N. If $n \geq N$, the product which has passed through inspection is considered acceptable and the inspection procedure is repeated on new product beginning with the segment next to the last segment partially inspected. If, on the other hand, $n < N$, the following actions are taken: (a) N-n segments (corresponding to k(N-n) items) are inspected 100 per cent. These segments are selected from the product beginning with the segment next to the last segment partially inspected; (b) the inspection procedure is repeated on new material, beginning the inspection with the segment following the last segment which was 100 per cent inspected.

Here, as in the previous case, the AOQL cannot exceed $(k-1/k)(m/N)$ regardless of whether or not the production process is stable.

The third and last inspection plan of this class which I should like to discuss in some detail is one in which the criterion for decisions is identical with that of the sequential probability ration rest developed by Wald for testing statistical hypotheses. These plans are of interest because, they require essentially a fixed amount of 100 per cent inspection when such inspection is called for, and moreover, the so-called one-sided sequential plan is the most efficient in the sense that for a prescribed AOQL it requires least inspection when the production process is in statistical control.

The one-sided sequential plan is defined by a constant h, an integer k and a positive fraction s. The integer k specifies the partial sampling rate to be employed and the fraction s specifies the desired AOQL. The sequence of items are divided into segments of size k and one item is selected at random from each segment and the number of defectives found is cumulated. At the j-th stage of partial inspection, the cumulative number of defectives is given by $\sum_{i=1}^{j} x_i$, where, as above, $x_i = 1$ or $x_i = 0$ depending on whether or not the item in the i-th segment was found defective or non-defective. Partial inspection continues as long as $\sum_{i=1}^{j} x_i < h + sj$. If for some $j = n$, $\sum_{j=1}^{n} x_i \geq h+sn$, partial inspection terminates and h/s segments are inspected 100 per cent (we assume that h/s is chosen to be an integer). The inspection procedure is then repeated on new material, beginning with the segment following the last segment which was 100 per cent inspected.

Except for a slight approximation which is negligible when h is not small, the AOQL for this plan is equal to $(k-1/k)s$.

When the production process is in statistical control and if $p \leq s$, we know from sequential theory that there exists a positive probability that partial inspection once begun will go on indefinitely. Thus, if $p \leq s$, the probability is 1 that eventually a sequence of items will be found in which only partial inspection will be employed. On the other hand, if $p > s$, partial inspection will always terminate with probability 1. But, in this case, the average outgoing quality $Q(p)$ is given by

$$Q(p) = \frac{(k-1)pE(n)}{kE(n) + k\frac{h}{s}} \tag{15}$$

and since,

$$E(n) = \frac{h}{p-s}, \tag{16}$$

we get

$$Q(p) = \frac{k-1}{k}s = \text{AOQL} \tag{17}$$

Thus for all $p > s$, the average fraction inspected is given by

$$F(p) = 1 - \frac{\text{AOQL}}{p} \tag{18}$$

which is minimum.

While the one-sided sequential inspection plan is optimum from the point of view of cost of inspection (at least in the case of controlled production) it is not a plan which can be recommended in cases where the excessive variability of the outgoing quality in finite batches of the product is an important factor. This becomes apparent when we consider the fact that in order for any material to be inspected 100 per cent, the point $(\sum_{i=1}^{j} x_i, j)$ must reach or exceed the line $y = h + sj$ for some sample size j. But if few defective items are found during a long stretch of partial inspection, the point can wander so far away from the line that a great deal of unacceptable material can pass by before this fact is noted and the quality of the product improved by the 100 per cent inspection. This situation can be remedied to some extent by employing a two-sided sequential inspection plan.

A two-sided sequential inspection plan is defined by two lines

$$y = h_2 + sj \tag{19}$$

and

$$y = -h_1 + sj \tag{20}$$

where h_1 and h_2 are positive constants and j represents the number of items inspected at the j^{th} stage of partial inspection. In this plan, partial inspection continues as long as the point $(\sum_{i=1}^{j} x_i, j)$ lies between the above two lines. Partial inspection terminates when, for some $j = n$, either

(a) $$\sum_{i=1}^{n} x_i \leq -h_1 + sn$$

or

(b) $$\sum_{i=1}^{n} x_i \geq h_2 + sn \ .$$

In the former case no 100 per cent inspection is called for and the inspection procedure is simply repeated on new material. In the latter case, the inspection is resumed on new material only after h_2/s segments (i.e., $k(h_2/s)$ items) have been inspected 100 per cent.

Except for the action taken, the inspection procedure described above is identical with the sequential method of lot-by-lot acceptance inspection. Consequently, all the sequential theory which has been recently developed can be employed in studying the consequences of this plan. Thus, for example, if the production is in statistical control, the AOQ curve can be computed from the formula

(21) $$Q(p) = \frac{p(k-1)E(n)}{k\left\{E(n) + [1 - L)p)]\right\} \frac{h_2}{s}}$$

where L(p) is the Operating Characteristic of the sequential probability ratio test and E(n) is the Average Sample Number. The API curve is, of course, given by

(22) $$F(p) = 1 - \frac{Q(p)}{p} \quad .$$

In the two-sided sequential plan, Q(p) is always less than AOQL even if $p > s$ but approaches it asymptotically as p approaches 1. Thus introducing the lower line increases somewhat the cost of inspection but usually not to an appreciable extent unless h_1 is made exceedingly small.

With any of the above plans, there is a minimum amount of inspection which cannot be avoided regardless of the quality of the product. This minimum is given by the partial sampling rate employed. Generally speaking, the partial sampling rate has but slight influence on the AOQ but a great deal of influence in determining the variability of the outgoing quality. The smaller the segment size k, the smaller will be this variability. However, when it is small, the cost of inspection will be unduly large whenever the product happens to be of acceptable quality. Thus in cases where the production fluctuates periodically from good to bad quality, it would be desirable to have a plan in which the sampling rate automatically adjusts itself to the quality of the product. A possible approach to this problem is to introduce two sampling rates, one strict and one reduced.

To illustrate, suppose the second plan which I have discussed is in use. We choose two integers k_1 and k_2 with $k_1 > k_2$. The sequence of product is divided into either segments of size k_1 or k_2 depending on whether reduced or strict inspection is called for, respectively.

Inspection begins with strict inspection. That is, one item is chosen at random from each consecutive segment of k_2 items and the number of defective items found, as well as the number of items sampled, is accumulated. Inspection terminates when and only when m defectives have been found. If m defectives have been found with a sample size $n \geq N$, the inspection procedure is repeated on new material, but with the reduced sampling rate, that is, the sequence of items is now divided into segments of k_1 items each and partial inspection consists of selecting and examining one item out of each segment. If, however, m defectives have been found with a sample size $n < N$, the following actions are taken: (a) a batch of $k_2(N-n)$ consecutive items is completely inspected, (b) the inspection procedure is repeated on new material, beginning with the segment just after the last segment that was inspected 100 per cent.

In the second and subsequent inspection operations, a decision is always made when (but not before) m defectives have been found. If reduced inspection was in effect and the sample size $n < N$, a batch of $k_1(N-n)$ items is completely inspected and the inspection procedure is repeated on new material but the sampling rate is changed from reduced to strict. If reduced inspection was in effect and the sample size $n \geq N$, the inspection procedure is repeated on new material at the reduced sampling rate. If however, strict inspection was in effect and the sample size $n < N$, a batch of $K_2(N-n)$ items is completely inspected and the inspection procedure repeated on new material at the strict sampling rate. If strict inspection was in effect and the sample size $n \geq N$, the inspection procedure is repeated but at the reduced sampling rate. Whenever m defectives are found and the sample $n \geq N$, the inspection procedure is repeated on new material employing a reduced sampling rate regardless of the sampling rate in effect in the operation just preceding.

The criterion for employing either reduced or strict inspection can be summarized briefly as follows: Whenever an inspection operation terminates and 100 per cent inspection of some material is called for, it also calls for strict inspection to follow. Conversely, if no 100 per cent inspection is indicated, then reduced inspection is to follow.

It is easily seen that when the material submitted for inspection is of good quality, the partial sampling rate will most often be reduced, while if the material is of poor quality, the partial sampling rate will most often be strict. This will have the effect of reducing the cost of inspection when inspection is least needed and reducing the variability of the OQ when such a reduction is of significance.