MULTIPLE SAMPLING WITH CONSTANT PROBABILITY

By Walter Bartky

The University of Chicago

1. Introduction. In an attempt to reduce inspection costs, manufacturers have frequently resorted to sampling procedure in which the disposition of an aggregate or lot of similar units does not necessarily depend upon the results of a single sample. In practice, however, the number of permissible additional samples is limited to one or two; nevertheless, if the lot is very large, an appreciable reduction in the expected sample may be accomplished by allowing a greater number of additional samples. In this article probability formulae will be derived for an inspection procedure for infinite lots in which the number of additional samples is not limited and may be any number depending upon the results of the sampling. This development will be limited to the simple case of attribute inspection in which the units fall into two categories—satisfactory units or defective units. If p denotes the fraction defective in an infinite lot, then the probability of finding exactly m defective units or defects in a sample of n is

$$(1) \qquad P(m, n) = \binom{n}{m} p^m q^{n-m}, \qquad q = 1 - p.$$

Since $P(m, n)$ is the probability of m successes in n trials with constant probability of success p, though the terminology of commercial inspection will be used in this article, the results are applicable to other situations involving repeated trials with constant probability of success.

In contrast with multiple sampling, a single sample inspection procedure for lots of the type here considered is one in which a lot of units is accepted or rejected on the basis of the number of defective units found in the sample. Thus a lot is accepted if the number of defects is at most an integer c the "acceptance number," and rejected if the number exceeds c. For an infinite lot containing a fraction p of defects and a sample of n units, the probability of accepting is by (1)

$$(2) \qquad \prod_s (c, n) = \sum_{m \leq c} P(m, n),$$

and the probability for rejection is the difference between this sum and unity.

2. Multiple sampling. The procedure in multiple sampling is to examine first an initial sample of n_0 units. If the number of defects in this initial sample is at most c the lot is accepted and if the number of defects exceeds $c + k$ (k an integer) the lot is rejected. But if the number of defects is greater than c and less than $c + k + 1$ an additional sample is removed and examined. In the latter case similar criteria determine whether the lot is to be accepted or rejected or this method of sampling continued. With an infinite lot this scheme of sampling has an infinite variety of forms but there are certain advantages in limiting this discussion to the following type of multiple sampling procedure.

I. *Sample Sizes:* The initial sample is of n_0 units but all additional samples are of the same size, namely n units.

II. *Condition for Acceptance:* The lot is accepted if the number of defects in initial sample of n_0 units is at most c or if after taking r additional samples of n the total number of defects in the $n_0 + rn$ units examined equals $c + r$.

III. *Condition for Rejection:* The lot is rejected if the number of defects in initial sample of n_0 units exceeds $c + k$ or if after taking r additional samples of n the total number of defects exceeds $c + r + k$.

IV. *Condition for an Additional Sample:* An additional sample of n is taken only if neither condition II nor condition III is realized.

Thus in this sampling scheme the level for acceptance as well as the level for rejection increases by unity for each additional sample of n. If at the r-*th* additional sample a lot is neither accepted nor rejected then the total number of defects in initial plus additional samples must equal one of the k numbers

$$c + r + 1, c + r + 2, \cdots, c + r + k.$$

Denote the probabilities for obtaining these numbers by

$$P_1(r),\ P_2(r),\ \cdots,\ P_k(r) \tag{3}$$

respectively, the subscript indicating the number of defects in excess of the acceptance level.

To be accepted on the $(r + 1)$-*st* additional sample, (a) no defect must be found in the $(r + 1)$-*st* additional sample and (b) a total of $c + r + 1$ defects must be found in previous samples. The probability of (a) is given by (1), taking m equal to zero, and the probability of (b) is the first one in the set (3). Consequently the probability of accepting a lot on the $(r + 1)$-*st* additional sample is

$$P_0(r + 1) = q^n P_1(r).$$

If Π denotes the probability of eventually accepting the lot

$$\prod = \sum_{m \leq c} P(m, n_0) + q^n[P_1(0) + P_1(1) + P_1(2) + \cdots], \tag{4}$$

where the first term on the right is the probability of accepting on the initial sample and may be evaluated by means of (1). Furthermore

$$P_i(0) = P(c + i, n_0) \tag{5}$$

and is by (1) the probability of finding $c + i$ defects in initial sample.

According to the notation (3) the probability of finding a total of $c + r + 1 + i$ defects in initial plus $r + 1$ additional samples, that is i more defects than the acceptance level, is $P_i(r + 1)$. These probabilities may be expressed as linear combinations of the probabilities (3) with coefficients that are probabilities of the type (1). Thus

$$P_i(r + 1) = \sum_j P(i - j + 1, n)P_j(r) \tag{6}$$

where the sum may be made to extend for $j = 1, 2, \cdots, k$, provided one defines (1) as equal to zero for negative m. By repeated application of this linear transformation it is possible to express the probabilities (3) for additional samples in terms of the probabilities (5) for the initial sample. Thus if M denotes the $k \times k$ square matrix with elements

$$M_{ij} = P(i - j + 1, n) \qquad (i, j = 1, \cdots, k), \tag{7}$$

by omitting subscripts and regarding $P(r)$ as a vector with elements given by (3), the linear transformation may be written

$$P(r + 1) = MP(r). \tag{8}$$

Hence by repeated application of (8)

$$P(r) = M^r P(0) \qquad (r = 0, 1, 2, \cdots) \tag{9}$$

provided the zero power of the matrix M is defined as the identity matrix I.

The probability $P_i(r)$ cannot exceed the probability of finding exactly $c + r + i$ defects in a single sample of $n_0 + rn$ units, that is, in the notation of (1), the probability $P(c + r + i, n_0 + rn)$. Since the latter probabilities approach zero as r approaches infinity it follows that the limit of the elements of $P(r)$ as r approaches infinity is zero. Thus with this multiple sampling procedure a lot is eventually either accepted or rejected. Furthermore since the matrix M contains no negative elements and $P(0)$ may be chosen with all positive elements it follows that the elements of M^r approach zero as r approaches infinity or

$$\lim_{r \to \infty} M^r = \text{“0”}, \quad \text{the zero matrix.} \tag{10}$$

It can be demonstrated that since the limit (10) is the zero matrix the sum of the infinite geometrical series in the matrix M

$$I + M + M^2 + \cdots = (I - M)^{-1}, \tag{11}$$

where the right member is the reciprocal of the matrix $I - M$. Consequently the infinite sum of vectors

$$V = \sum_{r=0}^{\infty} P(r) = (I - M)^{-1} P(0). \tag{12}$$

This infinite sum of vectors has elements $V_1, V_2, \cdots, V_k$ of which the first element is the sum in brackets occurring in the right member of (4). Hence the probability of eventually accepting the lot

$$\prod = \sum_{m \leq c} P(m, n_0) + q^n V_1, \tag{13}$$

and is thus by (12) and (5) expressible in terms of probabilities for the initial sample, equations (1), and the reciprocal of the matrix $I - M$.

In addition to the probability for acceptance one is also interested in the expected number, E, of additional samples. Since

$$\sum_i P_i(r-1) \qquad (r = 1, 2, 3, \cdots),$$

where the sum extends over all $i = 1, 2, \cdots, k$ is the probability of continuing to the r-th sample, it follows that

$$\sum_i P_i(r-1) - \sum_i P_i(r)$$

is the probability that lot will be either accepted or rejected on the r-th sample. Therefore the expected number of additional samples

$$\begin{aligned} E &= \sum_{r>0} r[\sum_i P_i(r-1) - \sum_i P_i(r)] \\ &= \sum_{r \geqq 0} \sum_i P_i(r), \end{aligned}$$

or, on interchanging the order of summation and applying (12),

$$E = \sum_i V_i. \tag{14}$$

That is, the expected number of additional samples equals the sum of the elements of the vector V.

Though it is possible to develop a general expression for the reciprocal matrix $I - M$, to determine the acceptance probability, Π, as well as the expected number of additional samples it is only necessary to evaluate V. Now by (12) this vector is the solution of the linear system of equations

$$(I - M)V = P(0). \tag{15}$$

Though for k small this system could be solved directly, in order to find a form of the solution applicable for any value of k, let the expansion in power series in x of

$$[(px + q)^n - x]^{-1} = g_1 + g_2 x + g_3 x^3 + \cdots, \tag{16}$$

where the coefficients, g, are functions of p and q. On clearing of fractions and equating coefficients of like powers of x it is found that

$$g_1 = q^{-n} \tag{17}$$

and, by equating the coefficients of the first k powers of x and using the notation (7),

$$g_i - \sum_{j=1 \cdots k} M_{ij} g_j = \begin{cases} 0 & (i = 1, 2, \cdots, k - 1), \\ g_{k+1} & (i = k). \end{cases} \tag{18}$$

Similarly, if the expansion in power series of

$$\frac{\sum_i P_i(0)x^i}{(px+q)^n - x} = h_1 + h_2 x + h_3 x^3 + \cdots, \tag{19}$$

where the sum is for all $i = 1, \cdots, k$, then by clearing of fractions and equating coefficients of like powers of x it is found that

$$h_1 = 0, \tag{20}$$

and

$$h_i - \sum_{j=1\cdots k} M_{ij} h_j = \begin{cases} -P_i(0) & (i = 1, \cdots, k-1), \\ -P_k(0) + h_{k+1} & (i = k). \end{cases} \tag{21}$$

It follows from equations (18) and (21) that if

$$V_i = g_i h_{k+1}/g_{k+1} - h_i \qquad (i = 1, \cdots, k), \tag{22}$$

then V, the vector with these elements, will satisfy equation (15). Since by (17) and (20)

$$V_1 = q^{-n} h_{k+1}/g_{k+1}, \tag{23}$$

the probability for eventually accepting the lot is by (13) expressible as

$$\Pi = \sum_{m \leqq c} P(m, n_0) + h_{k+1}/g_{k+1}, \tag{24}$$

while the expected number of additional samples is the sum of elements (22) of V_i.

These results will now be summarized and simplified formulae derived for special cases. In the summary all probabilities are expressed by means of (5) in terms of the probabilities (1).

3. Summary of multiple sampling formulas. For this multiple sampling procedure the initial sample is n_0 and the additional samples are n. A lot is accepted if on the r-th additional sample the total number of defects found is at most $c + r$ and rejected if the total exceeds $c + r$. An infinite lot containing a fraction p of defects is either accepted or rejected, the probability of acceptance being given by

$$\Pi = \sum_{m \leqq c} \binom{n_0}{m} p^m q^{n-m} + h_{k+1}/g_{k+1} \qquad (q = 1 - p), \tag{25}$$

and the probability of rejection is $1 - \Pi$. The expected number of additional samples is

$$E = \frac{h_{k+1}}{g_{k+1}} \sum_i g_i - \sum_i h_i, \tag{26}$$

where the sum extends over $i = 1, 2, \cdots, k$. The g_i and h_i are the coefficients in power series of x in the expansions of:

$$\frac{1}{(px+q)^n - x} = g_1 + g_2 x + g_3 x^2 + \cdots, \tag{27}$$

$$\sum_i \frac{\binom{n_0}{c+i} p^{c+i} q^{n_0-c-i} x^i}{(px+q)^n - x} = h_1 + h_2 x + h_3 x^2 + \cdots, \tag{28}$$

where the sum is for all $i = 1, 2, \cdots, k$. These formulae apply to all finite values of c and k provided the binomial coefficient is zero for values of the argument falling outside those occurring in the ordinary expansion of an integral power of a binomial.

4. Computation of coefficients g and h. If the denominator in (27) is first expanded in power series in

$$x(px+q)^{-n}$$

and then the resulting negative powers of binomials expanded in power series in x, it is found that

$$\begin{aligned} g_1 &= q^{-n}, \\ g_2 &= q^{-2n} - \binom{n}{1} p q^{-n-1}, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ g_k &= q^{-kn} - \sum_{m=1,\cdots,k-1} (-1)^{m+1} \binom{(k-m)n + m - 1}{m} \\ &\qquad \times p^m q^{-kn+mn-m}, \qquad k \neq 1. \end{aligned} \tag{29}$$

By (28) the coefficients h are expressible in terms of the g's,

$$\begin{aligned} h_1 &= 0, \\ h_k &= \sum_{i=1,\cdots,k-1} \binom{n_0}{c+i} p^{c+i} q^{n_0-c-i} g_{k-i}, \qquad k \neq 1. \end{aligned} \tag{30}$$

Other expressions for the coefficients may be derived from the theory of functions of a complex variable. Thus by Cauchy's Integral Formula

$$\begin{aligned} g_{k+1} &= \frac{1}{2\pi\sqrt{-1}} \int_C \frac{dx}{x^{k+1}[(px+q)^n - x]}, \\ h_{k+1} &= \frac{1}{2\pi\sqrt{-1}} \int_C \frac{S(x)\,dx}{x^{k+1}[(px+q)^n - x]}, \end{aligned} \tag{31}$$

where

$$S(x) = \sum_{i=1,\cdots,k} \binom{n_0}{c+i} p^{c+i} q^{n_0-c-i} x^i, \tag{32}$$

and the closed path of integration C in the complex plane only includes the pole at the origin. Since the integrands are rational functions and the point at infinity is not a singularity for either integrand, these integrals taken about the origin are equal to the negative sum of the corresponding integrals taken about the zeros of

$$(px + q)^n - x.$$

If $p \neq n^{-1}$ it can be demonstrated that there are n distinct zeros $x_1, x_2, \cdots, x_n$ corresponding to the solutions of the algebraic equation

$$(px_s + q)^n = x_s \qquad (s = 1, \cdots, n). \tag{33}$$

One solution is obviously

$$x_1 = 1, \tag{34}$$

and for $p = n^{-1}$ this solution is a double root.

The integrals about these zeros are obtainable from Cauchy's Integral Formula and after integrating and simplifying the resulting sum by means of (33) it is found that for the case $p \neq n^{-1}$,

$$\begin{aligned} g_{k+1} &= \frac{1}{1 - np} + \sum_{s=2,\cdots,n} \frac{px_s + q}{x_s^{k+1}[q - (n-1)px_s]}, \\ h_{k+1} &= \frac{S(1)}{1 - np} + \sum_{s=2,\cdots,n} \frac{(px_s + q)S(x_s)}{x_s^{k+1}[q - (n-1)px_s]}. \end{aligned} \tag{35}$$

If the power series (27) is multiplied by the series

$$(1 - x)^{-1} = 1 + x + x^2 + x^3 + \cdots,$$

the resulting product

$$\frac{1}{(1-x)[(px+q)^n - x]} = g_1 + (g_1 + g_2)x + (g_1 + g_2 + g_3)x^2 + \cdots,$$

so that, by Cauchy's Integral Formula,

$$G_k = \sum_{i=1,\cdots,k} g_i = \frac{1}{2\pi\sqrt{-1}} \int_C \frac{dx}{x^k(1-x)[(px+q)^n - x]}. \tag{36}$$

Similarly the sum of the coefficients h that occur in the right member of (26) may be written

$$H_k = \sum_{i=1,\cdots,k} h_i = \frac{1}{2\pi\sqrt{-1}} \int_C \frac{S(x)\,dx}{x^k(1-x)[(px+q)^n - x]}. \tag{37}$$

The integrals (36) and (37) are of the same type as (31), and by employing the same method of integrating used in deriving (35), the following expressions for the sums of coefficients g and h occurring in (26) are obtained:

$$G_k = \sum_i g_i = \frac{k}{1-np} - \frac{n(n-1)p^2}{2(1-np)^2} + \sum_{s=2,\cdots,n} \frac{px_s+q}{x_s^k(1-x_s)[q-(n-1)px_s]},$$

$$(38)\quad H_k = \sum_i h_i = \frac{kS(1)-S'(1)}{1-np} - \frac{n(n-1)p^2 S(1)}{2(1-np)^2} + \sum_{s=2,\cdots,n} \frac{(px_s+q)S(x_s)}{x_s^k(1-x_s)[q-(n-1)px_s]}$$

provided $np \neq 1$. Here $S'(1)$ is the derivative of (32) with respect to x evaluated for $x = 1$. For the special case $np = 1$, two of the roots of (33)

$$x_1 = x_2 = 1,$$

and the integrals (36), (37) and (38) become respectively

$$(n-1)g_{k+1} = 2kn + \tfrac{8}{3}n - \tfrac{4}{3} + \sum_{s\geqq 3} \frac{n+x_s-1}{x_s^{k+1}(1-x_s)},$$

$$(n-1)h_{k+1} = (2kn + \tfrac{8}{3}n - \tfrac{4}{3})S(1) - 2nS'(1) + \sum_{s\geqq 3} \frac{n+x_s-1}{x_s^{k+1}(1-x_s)} S(x_s),$$

$$(39)\quad (n-1)\sum_i g_i = k^2 n + \tfrac{5}{3}kn + \tfrac{1}{18}n - \tfrac{4}{3}k - \tfrac{1}{18} - \tfrac{1}{9}n^{-1} + \sum_{s\geqq 3} \frac{n+x_s-1}{x_s^k(1-x_s)^2},$$

$$(n-)1\sum_i h_i = (k^2 n + \tfrac{5}{3}kn + \tfrac{1}{18}n - \tfrac{4}{3}k - \tfrac{1}{18} - \tfrac{1}{9}n^{-1})S(1)$$

$$- (\tfrac{2}{3}n - \tfrac{4}{3} + 2kn)S'(1) + nS''(1) + \sum_{s\geqq 3} \frac{n+x_s-1}{x_s^k(1-x_s)^2} S(x_s),$$

where the sum extends over all roots of (33) that are not equal to unity. Here $S'(1)$ and $S''(1)$ are the first and second derivatives of (32) with respect to x for $x = 1$ and $p = n^{-1}$.

Formulas (35), (38) and (39) require for their evaluation the solutions of equation (33). For n greater than unity there are just two positive real solutions, say $x_1 = 1$ and x_2. If n is even all other roots are complex numbers, while if n is odd they are complex with the exception of one negative real root. Consequently by (33) for $s = 3, 4, \cdots, n$ the absolute values of the roots satisfy the inequality

$$(p \mid x_s \mid + q)^n > x_s,$$

and consequently the $\mid x_s \mid$ cannot be between x_1 and x_2. But equation (33) may be written

$$\frac{(px_s+q)^n - 1}{(px_s+q) - 1} = \frac{1}{p} \qquad (s \neq 1)$$

so that for $s = 2, 3, \cdots, n$

$$\sum_i (px_s + q)^i = 1/p \tag{40}$$

where the sum is taken for $i = 0, 1, \cdots, n - 1$ and therefore

$$\sum_i (p \mid x_s \mid + q)^i > 1/p \qquad (s = 3, 4, \cdots, n).$$

Now x_2 is the only real and positive solution of (40), consequently, in order to satisfy the inequality, the absolute values of roots corresponding to $s = 3$, $4, \cdots, n$ must exceed x_2. On combining this result with the former, it follows that

$$\mid x_s \mid > 1 \quad \text{and} \quad x_2. \tag{41}$$

Consequently for large values of k the most important terms in the right members of (35), (38) and (39) correspond to the real positive roots $x_1 = 1$ and x_2 of equation (33). By omitting the terms corresponding to $s = 3, \cdots, n$ one can derive approximations to the g and h and their sums applicable for large k values. In fact for np near unity the roots corresponding to $s = 3, 4, \cdots$ are considerably greater than unity as is illustrated in the following table of roots for the case $np = 1$:

$n = 2, \quad p = 1/2; \qquad x_s = 1, 1;$

$n = 3, \quad p = 1/3; \qquad x_s = 1, 1, -8;$

$n = 4, \quad p = 1/4; \qquad x_s = 1, 1, -7 \pm 4\sqrt{-2};$

$n = 5, \quad p = 1/5; \qquad x_s = 1, 1, -12.2531 \cdots,$

$$-4.8734 \cdots \pm 7.7343 \cdots \sqrt{-1}$$

and for $s = 3, 4, 5, \cdots$, $\mid x_s \mid$ is greater than 8.

For very large values of n and small values of p one can find approximate values for the roots by solving the limit equation obtained from (33) by putting

$$a = np$$

and letting n approach infinity. This equation is

$$e^{a(x_s - 1)} = x_s, \tag{42}$$

where e is the base of the natural logarithms. For the case $a = 1$, the roots are $1, 1, 3.0891 \cdots \pm 7.4602 \cdots \sqrt{-1}, 3.66 \cdots \pm 13.88 \cdots \sqrt{-1}$ and

$$x_s = \frac{b(1 + \log_e b)}{b^2 + 1} (b - \sqrt{-1}) + b\sqrt{-1} \quad \text{approximately,}$$

where

$$b = (2u + 1/2)\pi, \qquad u = 4, 5, 6, \cdots.$$

From equation (39) and these numerical results it follows that even with k as small as 3 the percentage error for the case $np = 1$ introduced in g_4 by omitting the terms in the indicated sum is less than .002%. Consequently for all practical purposes one may omit the complex and negative roots for values of k greater than 3 in computing the g's for np in the neighborhood of unity. For smaller values of k the exact values of the g's are readily obtainable from (29).

5. Special cases. Consider first the case in which $c < 0$ and $n_0 \leqq k + c$. With these conditions, under no circumstances could a lot be accepted or rejected on the initial sample and the indicated sum in the right member of (25) is zero. Furthermore for this case the sum (32) becomes

$$S(x) = (px + q)^{n_0} x^{-c}. \tag{43}$$

Consequently it follows from (33) that

$$S(x_s) = x_s^{t-c}, \tag{44}$$

where

$$t = n_0/n. \tag{45}$$

It should be noted however, that for t not an integer the right member of (44) is multiple valued and one must take that value for which

$$x_s^t = (px_s + q)^{n_0}. \tag{46}$$

Thus for real positive values of x_s, the right member of (44) is real. For integral values of t there is of course no ambiguity in the notation.

If (44) is substituted in the second equation of (35), the resulting expression for the h coefficient is of the same form as that for the g coefficient, in fact

$$h_{k+1} = g_{k-t+c+1},$$

so that by (25) the probability for acceptance is for this case

$$\Pi = g_{k-t+c+1}/g_{k+1}. \tag{47}$$

In similar manner it follows from (43) and (46) that the sum of the h coefficients, equation (38),

$$H_k = G_{k-t+c} + t$$

and hence by (26) the expected number of additional samples

$$E = \Pi G_k - G_{k-t+c} - t. \tag{48}$$

Since the initial sample is nt units and the additional samples are all equal to n units, the expected total number of units, sampled, that is, initial plus additional samples is

$$I = n_0 + nE = n(\Pi G_k - G_{k-t+c}). \tag{49}$$

Since for this case it is impossible to accept or reject on the initial sample one could combine the initial sample with the first additional sample. In fact one can continue combining initial and additional samples and thus increasing c and t provided the new initial sample n_0 and the new c value thus obtained are such that

$$c \leqq 0, \qquad n_0 = nt \leqq k + n - 1 + c. \tag{50}$$

In this process of combining samples t and c increase at the same rate and consequently formula (47), and the right member of (49) are unchanged. In other words formulas (47) and (49) may also be used under conditions (50).

It was demonstrated in Section 3 that for k sufficiently large one can omit those terms in (35) and (38) corresponding to complex or negative roots of (33). If this is done the following useful approximations for the g and G are obtained:

$$\begin{aligned} g_k &= (1 - np)^{-1} + [q - (n - 1)px]^{-1}x^{-k+(1/n)}, \\ G_k &= k(1 - np)^{-1} - \tfrac{1}{2}n(n - 1)p^2(1 - np)^{-2} \\ &\qquad + [q - (n - 1)px]^{-1}(1 - x)^{-1}x^{-k+(1/n)}, \end{aligned} \tag{51}$$

provided $np \neq 1$, $k \neq 1$ and x is the real positive root of

$$(px + q)^n = x \qquad (np \neq 1) \tag{52}$$

that is not equal to unity. For $np = 1$ these approximations become by (39)

$$\begin{aligned} (n - 1)g_k &= 2kn + 2n/3 - 4/3 \\ (n - 1)G_k &= k^2n + 5kn/3 + n/18 - 4k/3 - 1/18 - n^{-1}/9, \qquad k \neq 1. \end{aligned} \tag{53}$$

These formulae in conjunction with formulae (47) and (49) give quite satisfactory approximations for the probability for acceptance Π and the expected total number of units sampled even when values of the subscripts employed are as small as 3. Of course the larger the value of k in (51), (52) or (53) the better these approximations.

Now the root x of (52) is greater or less than unity depending on whether the product $a = np$ is less than or greater than unity. Consequently it follows from (47) and (51) that for $c = 0$ and t finite

$$\begin{aligned} \Pi' &= \lim_{k\to\infty} \Pi = \lim_{k\to\infty} g_{k-t+1}/g_{k+1} \\ &= 1, \quad np < 1; \\ &= x^t, \quad np > 1; \end{aligned} \tag{54}$$

while by (49) and (51) the expected total number of units sampled has the limiting value

$$I' = \lim_{k\to\infty} I = \begin{cases} nt(1 - np)^{-1}, & np < 1; \\ \infty, & np > 1. \end{cases} \tag{55}$$

But k infinite implies that under no circumstance can a lot be rejected. Consequently Π' and I' are the exact values of the probability for acceptance and the expected total sample respectively for the following sampling procedure:

The initial sample is $n_0 = nt$ and all additional samples are n. The lot is accepted if on the initial sample no defects are found or if after taking r additional samples a total of exactly r defects is found.

In inspection problems p is usually small and n large so that the approximation (40) may be used to determine the real positive root x, thus

$$e^{a(x-1)} = x \qquad (a = np). \tag{56}$$

It then follows from (54) and (55) that for $np > 1$

$$\frac{-\log \Pi'}{1 - x} = n_0 p, \qquad \frac{-\log x}{1 - x} = np. \tag{57}$$

These relations are of course equivalent to (54) and (56). Suppose that the probability Π' and the fraction p are assigned. Then the initial sample n_0, and additional sample n, will depend on only the parameter x. Consider next the problem of sampling a number of lots that fall into two categories, namely those containing a fraction p of defects and those containing a fraction p^* of defects where $p^* < p$. If in addition the sampling procedure is to be such that lots with fraction p^* of defects are eventually accepted, but lots with fraction p of defects have a small assigned probability of acceptance Π', then whatever the value of x as long as the resulting $np \geqq 1$ these conditions are satisfied. Furthermore if one insists that the expected total sample for lots containing a fraction p^*, namely by (55)

$$I'(p^*) = n_0(1 - np^*)^{-1},$$

be a minimum, then it is found that

$$x = p^*/p. \tag{58}$$

This remarkably simple result is capable of still greater generalization. By an altogether different approach to the problem the author has succeeded in proving that of all possible multiple sampling procedures, the multiple sampling method here described and defined by equations (57) and (58) gives the minimum expected inspection for the problem under consideration provided n is sufficiently large.[1]

By letting both $-c$ and k approach infinity it is possible to derive probability formulae for sampling procedure in which a lot is either rejected or the sampling continues without end. These formulae are included in Table I along with other special cases derived from previously listed general formulae.

[1] *Note:* The author has postponed publication of this proof in the hope that it might be generalized to include sampling problems involving both acceptance and rejection of a lot.

TABLE I

Notation:

n = number of units in each additional sample
n_0 = number of units in initial sample
p = fraction defective in lot
$a = np$
$q = 1 - p$
c = maximum number of defects in initial sample for acceptance
$t = n_0/n$ = ratio initial sample to additional samples
$f = c + k + 1$ = minimum number of defects in initial sample for rejection
$c + r$ = number of defects in initial plus first r additional samples for acceptance
$f + r = c + k + 1 + r$ = minimum number of defects in initial plus first r additional samples for rejection
Π = probability of eventually accepting lot with fraction p defects
$1 - \Pi$ = probability of eventually rejecting lot with fraction p defects
I = expected total number of units sampled (i.e., initial plus whatever additional samples are sampled).
x = real positive root different from unity of the equation
$(px + q)^n = x.$

Conditions	Π	I
(a) $k = 1$, $c = 0$, $f = 2$	$q^{n_0} \times \frac{1 - (n - n_0)pq^{n-1}}{1 - npq^{n-1}}$	$n_0 \times \frac{1 - (q^{n-1} - q^{n_0-1})np}{1 - npq^{n-1}}$
(b) $k = 1$, $c = 0$, $f = 2$, $n_0 = n$	$q^n(1 - npq^{n-1})^{-1}$	$n(1 - npq^{n-1})^{-1}$
(c) $c = -k$, $f = 1$	q^{n_0-n}/g_{k+1}	$n_0 + nq^{n_0-n}G_k/g_{k+1}$
(d) $k = 1$, $c = -1$, $f = 1$	$q^{n_0+n}(1 - npq^{n-1})^{-1}$	$n_0 + nq^{n_0}(1 - npq^{n-1})^{-1}$
(e) $k = 2$, $e = -2$, $f = 1$	$\frac{q^{n_0+2n}}{1 - 2npq^{n-1} + \frac{n(n+1)}{2}p^2q^{2n-2}}$	$n_0 + \frac{nq^{n_0}(1 + q^n - npq^{n-1})}{1 - 2npq^{n-1} + \frac{n(n+1)}{2}p^2q^{2n-2}}$

TABLE I—*Concluded*

Conditions	Π		I	
(f) $k = -c$, $= \infty$, $f = 1$	0	for $np > 1$	$n_0 + nx(1-x)^{-1}$	for $np > 1$
	$q^{n_0-n}(1-np)$	for $np < 1$*	∞	for $np < 1$
(g) $c = 0$, $n = 2$, $n_0 = f$, $= k+1$	$\dfrac{1}{1+(p/q)^{n_0}}$		$\dfrac{n_0(2\Pi - 1)}{q-p}$	
(h) $c = 0$, $n = 2$, $n_0 = f$, $= k+1$, $p = 1/2$	0.5		n_0^2	
(i) $c = 0$, $n_0 = n$	g_k/g_{k+1}		$n(\Pi G_k - G_{k-1})$	
(j) $c = 0$, $k = \infty$	1	$(np < 1)$	$n_0(1-np)^{-1}$	$(np < 1)$
	$x^{n_0/n}$	$(np > 1)$	∞	$(np > 1)$

* In this sampling procedure a lot cannot be accepted so that Π is the probability that additional samples will be taken without end. The probability of rejecting lot is however 1 − Π.

TABLE II

Values of g and G for Limit $n = \infty$, $p = 0$

$np =$ $a =$	0.2558	0.4024	0.6931	1.0000	1.3863	2.0118	2.5584
$x =$	10	5	2	1	.5	.2	.1
g_1	1.292	1.495	2.000	2.718	4.000	7.477	12.915
g_2	1.338	1.634	2.614	4.671	10.455	40.86	133.76
g_3	1.3432	1.665	2.935	6.667	23.48	208.2	1343.2
g_4	1.3437	1.6717	3.097	8.667	49.55	1045.	13.4×10^3
g_5	1.3438	1.6729	3.178	10.667	101.70	5228.	134×10^3
g_∞	1.3438	1.6732	3.2589	∞	∞	∞	∞
G_1	1.292	1.495	2.000	2.718	4.000	7.477	12.915
G_2	2.629	3.130	4.614	7.389	14.45	48.34	146.7
G_3	3.972	4.795	7.549	14.05	37.93	256.5	1490
G_4	5.316	6.467	10.65	22.72	87.5	1301.	14.9×10^3
G_5	6.660	8.140	13.82	33.39	189.2	6529.	149×10^3

As an illustration of the method of application of these formulae, suppose that the sampling procedure is to be such that the probability, Π, of accepting a "p" value of $0.5 + \epsilon$ equals the probability of rejecting a "p" value of $0.5 - \epsilon$. This condition on probabilities is by Table I, formula (g), always satisfied if $c = 0$, $n = 2$, and $n_0 = k + 1$. This corresponds to a multiple sampling scheme in which additional samples are only two units each and a lot is accepted or rejected on initial sample if none or all units are defective. With $\epsilon = 0.1$ and $\Pi \leqq 1/6$, one can take $n_0 = 4$ and $k = 3$. The expected total number of units examined depends on "p" and varies for this numerical case from 4, for $p = 0$ or 1, to a maximum of 16, for $p = 0.5$. Nevertheless a single sample plan satisfying the same conditions would require a sample of 23 units whatever the value of p.

The previous problem is, however, not typical of those encountered in commercial inspection for in such situations p is usually very small. In practice one can generally replace the formulae in Table I by their limiting values for $n = \infty$, $p = 0$, and $np = a$. Table II gives the limiting values of the g and G as well as x for a small number of values of a.

Finally the justification for multiple sampling lies in the fact that a reduction in the expected total sample is possible. Though this paper is limited to the consideration of a very elementary type of sampling, it indicates that it might be worth while to investigate the possibility of utilizing the methods of multiple sampling in inspection for variables. Unfortunately serious mathematical difficulties are even encountered in so simple a problem as multiple sampling from a normal population for the mean.