

BAYESIAN SINGLE SAMPLING ACCEPTANCE PLANS

FOR FINITE LOT SIZES

by

Poul Thyregod

Technical Report No. 157

April 1971

University of Minnesota
Minneapolis, Minnesota

ERRATA to Technical Report No. 157

Bayesian Single Sampling Acceptance Plans For Finite Lot Sizes.

Professor A. Hald has kindly pointed out to me that formula (27) on p. 13 is incorrect. The formula should be

$$(27) \quad p_c(n|\omega) = \frac{c+1}{n+1} f^{(n+1)}(c+1|\omega) .$$

The correction necessitates the following changes:

p. 13, line 8 from the bottom: The last sentence in the paragraph should be: In the subsequent example we shall prove that the result also holds for the Bernoulli distribution.

p. 14 : Formula (28) is changed to:

$$-\Delta_n d(n, c) = \frac{c+1}{n+1} \Delta_c d(n+1, c),$$

p. 14 : Formula (29) is changed to:

$$d(n+1, c+1) - d(n, c) = \Delta_c d(n+1, c) + \Delta_n d(n, c) = \frac{n-c}{n+1} \Delta_c d(n+1, c).$$

p. 14, lines 8-14 (the end of the paragraph): Should read: from (29)

that $d(n_2+1, c+1) < d(n_2, c)$ and therefore we have

$$d(n_2+1, c_0(n_2+1)) \leq d(n_2+1, c+1) < d(n_2, c)$$

which shows that $d(n, c_0(n))$ is decreasing also for $n = n_2$. Thus it has been proved that $d(n, c_0(n))$ is a decreasing function of n .

p. 18 : The paragraph following formula (39) is replaced by: In table 1 we have listed the set of possible plans for $n < 30$. The table also gives the decision loss corresponding to these plans.

p. 19 : Table 1 should be replaced by page 2 of this note.

Finally, the Pascal distribution on p. 6 should be:

$$f(x|\omega) = \binom{x-1}{r-1} (1-\omega)^r \omega^{x-r} ,$$

TABLE 1

Table of $d(n, c_0(n))$ for $n < 30$. $k_s(\omega) = k_r(\omega) = 0.2$, $k_a(\omega) = \omega$
 and $dW(\omega) = 4(1-\omega)^3 d\omega$. $f(x|\omega) = \omega^x(1-\omega)^{1-x}$ for $x = 0, 1$.

n	$c_0(n)$	$d(n, c_0(n))$
0	0	.06553600
1	0	.03886933
2	0	.02744076
3	0	.02267886
4	0	.02109156
5	1	.02109156
6	1	.01705115
7	1	.01462691
8	1	.01332155
9	1	.01278875
10	2	.01278874
11	2	.01114040
12	2	.01007382
13	2	.00945757
14	2	.00918925
15	3	.00918924
16	3	.00829486
17	3	.00769442
18	3	.00733517
19	3	.00717350
20	4	.00717350
21	4	.00661232
22	4	.00622719
23	4	.00599184
24	4	.00588378
25	5	.00588378
26	5	.00549893
27	5	.00523090
28	5	.00506476
29	5	.00498744

1. Introduction and Summary.

We shall consider the following two-decision problem: Let X_1, X_2, \dots, X_N denote N independent identically distributed non-negative random variables with density $f(x|\omega)$ and let ω itself be a random variable with cumulative distribution function $W(\omega)$. Based on the observation of n of the variables, the observer chooses one of two possible terminal actions, accept or reject, say, with consequences depending on the remainder $N - n$ variables. The consequences will be measured in terms of the cost incurred per item and we shall assume that the cost of inspecting, accepting and rejecting an item are given as $k_s^*(X)$, $k_a^*(X)$ and $k_r^*(X)$, respectively, additive over the items. In quality control applications with lot by lot inspection N will usually denote the lot size and n the sample size.

A sampling plan or strategy is determined by the sample size n and a decision rule specifying the action as function of the sample outcome. A Bayesian sampling plan with respect to a distribution $W(\omega)$ is a sampling plan that minimizes the overall expected cost under $W(\omega)$.

Since the distribution of N exchangeable random variables allows a representation of the form

$$(1) \quad f_N(x_1, x_2, \dots, x_N) = \int \prod_{i=1}^N f(x_i|\omega) dW(\omega)$$

(see, e.g., the paper by Blum and Hanson (1960)) it follows that we may obtain a Bayesian sampling plan for an exchangeable prior distribution as the Bayesian sampling plan with respect to the distribution $W(\omega)$ in the representation (1) of the prior. Under the extra assumption of simple random sampling a Bayesian sampling plan with respect to $W(\omega)$ will be an optimal invariant Bayesian sampling plan for prior distributions $f_N(x_1, x_2, \dots, x_N)$ satisfying $f_N(\sum x_i) = \int f^{(N)}(\sum x_i|\omega) dW(\omega)$ with $f^{(N)}$ denoting the N -fold convolution of $f(x|\omega)$ with itself.

The model was originally introduced by Guthrie and Johns (1959) who, in their paper, gave asymptotic expansions for the Bayesian sampling plans under linear costs. Hald (1960) discussed the determination of the optimal (Bayesian) sampling plans in the linear case with Bernoulli random variables. In a later paper Hald (1967) gave asymptotic expansions for the Bayesian sampling plans under weak assumptions on the cost and the sampling distribution $f(x|\omega)$. Raiffa and Schlaiffer (1961) discussed the case where $f(x|\omega)$ is normal and gave a graphical solution to the optimization problem for a normal distribution of ω . A set of tables of optimal sampling plans for the Bernoulli case under a two-point distribution were published by Hald (1965), but, apart from these relatively simple situations, the literature does not contain many examples of optimal sampling plans. A possible explanation is that the computational effort involved in the determination of these plans has seemed too great even for a computer.

In a previous paper by Hald and Thyregod (1971) it was shown that the Bayesian sampling plans for the mean of a Poisson process enjoyed certain monotonicity properties, a result which greatly reduced the amount of computation necessary to tabulate the optimal sampling plans. The purpose of the present paper is to show that these properties are shared by the optimal sampling plans for a more general class of sampling distributions $f(x|\omega)$.

Throughout the paper we shall assume that $f(x|\omega)$ belongs to the exponential family of distributions, and moreover that f is a Polya frequency function with monotone likelihood ratio. The cost difference

$k_r^*(X) - k_a^*(X)$ is assumed to have only one change of sign. It is shown in Section 2 that the Bayesian sampling plan for this problem is of the form: take a sample of size n and base the decision on the event $X_1 + \dots + X_n \leq c$.

In Section 3 we discuss the decision loss for a decision rule characterized by the sample size and the acceptance number (n, c) , say. The main result is given in Theorem 1 which states that the optimal acceptance number is a monotone function of the sample size. Section 4 contains a proof of the intuitively reasonable result that the optimal sample size is an increasing function of N .

It is shown that the theory covers the following common situations:

- a. Sampling for defectives, X Bernoulli.
- b. Sampling for number of defects, X binomial or Poisson.
- c. Life testing, X Pascal or gamma.

Finally, it is shown in an example how the results of the paper may be used to simplify the tabulation of the optimal sampling plans as a function of lot size.

The fundamental results in Section 2 and 3 are simple consequences of basic results in the general theory of distributions with monotone likelihood ratio, or sign regular functions (of order 2), and therefore we shall conclude with a brief characterization of these distributions.

A non-negative function $f(x, y)$ is said to be sign regular if there exists a number ϵ either $+1$ or -1 such that

$$(2) \quad \epsilon \{f(x, y)f(x', y') - f(x, y')f(x', y)\} \geq 0 \quad \text{for } x \leq x', y \leq y'.$$

If ϵ in (2) is positive the function will be called totally positive otherwise, i.e., for $\epsilon = -1$, f is said to be reverse sign regular. If

a totally positive function may be written as a function $g(x-y)$ of the difference of x and y where $g(u)$ is a density of a random variable, we shall call $g(u)$ a Polya frequency.

The following properties of sign regular functions will be used without further reference (proofs may be found in the monograph by Karlin (1968)).

a. Let f and g be sign regular and let W denote a sigma-finite measure. Then the composition

$$h(x, z) = \int f(x, y)g(y, z)dW(y)$$

is reverse sign regular if either f or g (but not both) is reverse sign regular and totally positive otherwise.

b. Let f be sign regular and let W denote a sigma-finite measure such that $\int f(x, y)dW(y)$ exists for every x . Let $h(y)$ have at most one sign change and assume further that $g(x) = \int f(x, y)h(y)dW(y)$ exists for every x . Then $g(x)$ has at most one sign change and moreover, if g actually changes sign, h and g will exhibit the same sequence of signs when f is totally positive and the reverse sequence of signs when f is reverse sign regular.

c. If $g(u)$ is a one-sided Polya frequency, i.e., $g(u) = 0$ for $u < 0$, then the n -fold convolution of g with itself, $g^{(n)}(u) = g * g * \dots * g(u)$, is a Polya frequency and moreover $g^{(n)}(u)$ is totally positive in n and u .

2. The Probability Model.

We shall consider the situation described in the introduction and assume that the distribution of the random variable X_1 is absolutely

continuous with respect to counting measure or Lebesgue measure on the positive real axis with a density of the form

$$f(x|\omega) = B(\omega)h(x)\exp[xA(\omega)]$$

where $A(\omega)$ is a non-decreasing function of ω and $h(x)$ is a Polya frequency function satisfying $h(x) = 0$ for $x < 0$. The assumptions imply that $f(x|\omega)$ is totally positive in x and ω and that in addition $f(x|\omega)$ is a Polya frequency function.

Defining $f^{(n)}(x|\omega)$ as the n -fold convolution of f with itself we find

$$(3) \quad f^{(n)}(x|\omega) = [B(\omega)]^n h^{(n)}(x)\exp[xA(\omega)],$$

where $B(\omega)$ satisfies

$$(4) \quad [B(\omega)]^{-n} = \int h^{(n)}(x)\exp[xA(\omega)]d\mu(x),$$

μ denoting counting measure or Lebesgue measure. Since $h(x)$ is a Polya frequency we find that $h^{(n)}(x)$ is totally positive in n and x and hence the composition (4) with $\exp[xA(\omega)]$ yields a function which is totally positive in n and ω . Thus $[B(\omega)]^n$ and subsequently $f^{(n)}(x|\omega)$ are reverse sign regular in n and ω .

It follows that the set

$$(5) \quad I_n = \{x \in \mathbb{R} | h^{(n)}(x) > 0\}$$

is an interval on the real line or on the set of non-negative integers.

Furthermore, the marginal density of x ,

$$(6) \quad f^{(n)}(x) = \int f^{(n)}(x|\omega)dW(\omega),$$

is positive if and only if $x \in I_n$.

Example 1.

It is easy to verify that the following distributions satisfy the assumptions above.

- a. The Bernoulli distribution, $f(x|\omega) = \omega^x(1-\omega)^{1-x}$ for $x = 0,1$ and $0 \leq \omega \leq 1$.
- b. The geometric distribution, $f(x|\omega) = (1-\omega)\omega^{x-1}$ for $x = 1,2,\dots$ and $0 \leq \omega \leq 1$.
- c. The Poisson distribution, $f(x|\omega) = \omega^x e^{-\omega}/x!$ for $x = 0,1,\dots$ and $\omega \geq 0$.
- d. The exponential distribution, $f(x|\omega) = \omega^{-1} e^{-x/\omega}$ for $x \geq 0$ and $\omega > 0$.

From (3) we find that the exponential family is closed under convolution and moreover that $h^{(r)}(x)$, $r = 2,3,\dots$ is a Polya frequency and therefore $g(x|\omega) = f^{(r)}(x|\omega)$, $r = 2,3,4,\dots$ satisfies the assumptions if f does.

Thus a, b and d above may be extended to

- e. The binomial distribution, $f(x|\omega) = \binom{r}{x} \omega^x (1-\omega)^{r-x}$, $x = 0,1,\dots, r$, $0 \leq \omega \leq 1$.
- f. The Pascal distribution, $f(x|\omega) = \binom{x}{r} (1-\omega)^r \omega^{x-r}$, $x = r, r+1,\dots$, $0 \leq \omega \leq 1$ and
- g. The gamma distribution, $f(x|\omega) = x^{r-1} \omega^{-r} e^{-x/\omega}/r!$, $0 \leq x$, $0 < \omega$, integer r .

3. The Cost Function and the Decision Loss.

We shall assume that the cost of inspecting, accepting or rejecting an item of quality X is given by $k_s^*(X)$, $k_a^*(X)$ and $k_r^*(X)$, respectively, and that the cost difference $k_a^*(X) - k_r^*(X)$ changes sign at most once.

For simplicity we shall assume that the actions have been labeled such that the change is from negative to positive values of $k_a^* - k_r^*$. Defining $k_a(\omega) = E_{X|\omega}[k_a^*(X)]$ and $k_r(\omega) = E_{X|\omega}[k_r^*(X)]$ we deduce from the total positivity of $f(x|\omega)$ that $\underline{l}(\omega) = k_a(\omega) - k_r(\omega)$ has at most one sign change, from negative to positive values of \underline{l} .

Let n be fixed and consider a decision rule specified by the region of acceptance A_n , i.e., we accept if $(X_1, \dots, X_n) \in A_n$, reject otherwise. Since, for given ω , the first n observations are independent of the remaining $N - n$ variables we find the following expression for the expected cost

$$(7) \quad K = n \int k_s(\omega) dW(\omega) + (N-n) \int [P(A_n|\omega)k_a(\omega) + (1-P(A_n|\omega))k_r(\omega)] dW(\omega),$$

where

$$(8) \quad k_s(\omega) = E_{X|\omega}[k_s^*(X)]$$

and $P(A_n|\omega)$ denotes the conditional probability of acceptance, the operating characteristic. We shall assume that, in the long run, sampling is at least as costly as the terminal actions, i.e.,

$$(9) \quad E[k_s(\omega)] \geq \max\{E[k_r(\omega)], E[k_a(\omega)]\}.$$

If the true process parameter ω was known the best procedure would be to accept all items if $k_a(\omega) \leq k_r(\omega)$ and reject otherwise, which procedure would lead to the minimal cost

$$(10) \quad K_{\min} = N \int k_m(\omega) dW(\omega)$$

with

$$(11) \quad k_m(\omega) = \begin{cases} k_a(\omega) & \text{for } \omega \in \Omega_a \\ k_r(\omega) & \text{for } \omega \in \Omega_r \end{cases}$$

and $\Omega_a = \{\omega \in \Omega : k_a(\omega) \leq k_r(\omega)\}$, $\Omega_r = \Omega \setminus \Omega_a$. To avoid trivialities we shall assume that Ω_a and Ω_r are non-empty and that the distribution $W(\omega)$ assigns positive mass to both sets.

We note that if the average quality of the items is known the minimal cost will be $K_{\min}^* = N \int E_x | \omega k_m^*(x) dW(\omega)$ with $k_m^*(x) = k_a^*(x)$ if $k_a^*(x) \leq k_r^*(x)$ and $k_m^*(x) = k_r^*(x)$ otherwise. Since both K_{\min} and K_{\min}^* are independent of the sampling plan we may determine the optimal sampling plan by minimizing $K - K_{\min}$ or $K - K_{\min}^*$ instead of K .

In the following discussion we shall work with the loss relative to K_{\min} rather than the cost incurred. From (7) and (11) we find the regret

$$(12) \quad R = K - K_{\min} = n\delta_s + (N-n)\{\delta_r + \int \underline{1}(\omega)P(A_n | \omega)dW(\omega)\}$$

with

$$(13) \quad \delta_s = \int (k_s(\omega) - k_m(\omega))dW(\omega)$$

and

$$(14) \quad \delta_r = \int (k_r(\omega) - k_m(\omega))dW(\omega)$$

denoting the loss due to inspection and rejection, respectively. The acceptance loss is found analogously as

$$(15) \quad \delta_a = \int (k_a(\omega) - k_r(\omega))dW(\omega) .$$

The form of the regret (12) permits us to reduce the optimization problem considerably as shown in the following lemma.

Lemma 1.

The optimal decision rule is equivalent to a rule of the form:
Accept if $S_n \leq c$ with $S_n = X_1 + \dots + X_n$, reject otherwise.

Proof:

Since S_n is sufficient for ω in the conditional distribution of X given ω , an optimal decision rule can be based on S_n . Observing that the loss is monotone and that the density of S_n is totally positive in S and ω we find that the class of monotone procedures form a complete class for the decision problem and therefore the optimal rule must be monotone, i.e., based on the event $S_n \leq c$.

In view of the above result we may restrict ourselves to consider sampling plans characterized by two numbers (n, c) , say, where n is the sample size, the acceptance number is c and the decision rule is as given in the lemma. The operating characteristic for the sampling plan (n, c) is $P(n, c, \omega) = \int_0^c f^{(n)}(x|\omega) d\mu(x)$ and the expected decision loss per item is

$$(16) \quad d(n, c) = \delta_r + \int \underline{l}(\omega) P(n, c, \omega) dW(\omega) .$$

Introducing the posterior risk,

$$(17) \quad \lambda(n, x) = \begin{cases} \int \underline{l}(\omega) f^{(n)}(x|\omega) dW(\omega) / f^{(n)}(x) & \text{for } x \in I_n \\ 0 & \text{otherwise,} \end{cases}$$

we find

$$(18) \quad d(n, c) = \delta_r + \int_0^c \lambda(n, x) f^{(n)}(x) d\mu(x) .$$

The following theorem gives an important property of $d(n, c)$.

Theorem 1.

Let $c_0(n) = \inf\{x \in R : x \in I_n\}$ if $\lambda(n, x) \geq 0$ for all $x \in I_n$ and $c_0(n) = \sup\{x \in I_n : \lambda(n, x) < 0\}$ otherwise. Then the following results hold.

- a. $d(n, c)$ is non-increasing for $c < c_0(n)$ and non-decreasing for $c > c_0(n)$ with a minimum when $c = c_0(n)$.
- b. $\min_c R(n, c, N) = R(n, c_0(n), N)$ for $n \leq N$.
- c. $c_0(n)$ is non-decreasing with n .

Proof:

The existence of $c_0(n)$ is immediately verified if I_n is finite. If I_n is infinite we may choose a value $\omega^* \in \Omega_r$ carrying positive prior probability and construct a sequence $(n_1, x_1), (n_2, x_2), \dots$ with $n_i \rightarrow \infty$ such that the posterior density tends to concentrate around ω^* . Thus we find $\lambda(n_i, x_i) > 0$ for i sufficiently large. Now we note that since $f^{(n)}(x|\omega)$ is reverse sign regular in n and ω it follows that $\lambda(n, x)$ has at most one sign change (from positive to negative values) as n varies and hence $\lambda(n_i, x_i) > 0$ implies $\lambda(n, x_i) > 0$ for $n \leq n_i$. We may thus determine a value $x_i \in I_n$ satisfying $\lambda(n, x_i) > 0$ and hence find $c_0(n) < x_i$ which proves the existence of $c_0(n)$.

Restricting ourselves to $x \in I_n$ we find that $f^{(n)}(x|\omega)$ is totally positive in x and ω and hence we have

$$(19) \quad \lambda(n, x) \begin{matrix} \leq \\ > \end{matrix} 0 \text{ for } x \begin{matrix} \leq \\ > \end{matrix} c_0(n) \text{ and } x \in I_n.$$

Using (18) we find that an increase by Δ of the acceptance number c yields a contribution $\lambda(n, c+\Delta)f^{(n)}(c+\Delta)$ to the decision loss and therefore, in view of (19) we conclude the proposition a.

To prove b. we simply note that the regret depends on c only through the decision loss $d(n, c)$.

Using again the fact that $\lambda(n, x)$ varies from positive to negative values as n increases we may finally verify Proposition c.

The following corollary concludes the discussion of the decision loss as a function of the acceptance number only.

Corollary 1.

The decision loss satisfies the relation

$$(20) \quad d(n, x) \leq \max\{\delta_a, \delta_r\}$$

with δ_a and δ_r given by (14) and (15) respectively.

Proof:

Consider first $x < c_0(n)$. We then find from Proposition a and the form of the decision loss (18) that $d(n, x) \leq \delta_r$. For $c \geq c_0(n)$ it follows analogously that $d(n, x) \leq \lim_{c \rightarrow \infty} d(n, c)$. Letting $c \rightarrow \infty$ in (16) we find

$$(21) \quad d(n, c) \rightarrow \delta_r + \int \underline{1}(\omega) dW(\omega)$$

since $P(c, n, \omega) \rightarrow 1$. The proof is completed by the observation that $\delta_r + \int \underline{1}(\omega) dW(\omega) = \delta_a$.

We shall conclude this section with a discussion of the decision loss as a function of the sample size.

To facilitate the discussion the following notation is introduced,

$$(22) \quad p_c(n|\omega) = P(n, c, \omega) - P(n+1, c, \omega), \quad n = 1, 2, \dots$$

Note that p_c gives the distribution of the waiting time until the first crossing of the level $S_n = c$ in an infinite series of sample results, i.e., we have $p_c(n|\omega) = P[T(c) = n]$ where the random variable $T(c)$ denotes the largest index k for which $S_k \leq c$ in an unlimited series of independent observations of a variable with density $f(x|\omega)$.

Noting that

$$(23) \quad \Delta_n d(n, c) = - \int \underline{1}(\omega) p_c(n|\omega) dW(\omega) \quad \text{for } n = 1, 2, \dots,$$

where Δ_n denotes the usual forward difference operator, $\Delta_n f(n) = f(n+1) - f(n)$, we obtain

Theorem 2.

Let $p_c(n|\omega)$ be reverse sign regular in n and ω and totally positive in c and ω . Then there exists an integer $n_0(c)$ such that $d(n, c)$ is non-increasing for $n < n_0(c)$ and non-decreasing for $n > n_0(c)$. Moreover $n_0(c)$ is non-decreasing with c .

Proof:

The proof will be omitted since it is analogous to the proof of Theorem 1.

It is easy to verify that the assumptions of the theorem are satisfied for the distributions a) - e) in Example 1. In particular, the assumptions are satisfied for these distributions when n is restricted to integer multiples of r , $r = 2, 3, \dots$, only. Thus it follows that the conclusions of the theorem hold for all the distributions discussed in Example 1.

It should be noted that the assumptions of Theorem 2 are automatically fulfilled in the case where observation of the waiting time $T(c) = n$ gives rise to the same likelihood function as the event $S_n = c$. As the following lemma indicates we may obtain even stronger results in this particular case.

Lemma 2.

Let $f(x|\omega)$ be absolutely continuous with respect to Lebesgue measure. If $p_c(n|\omega) = k(n, c) f^{(n)}(c|\omega)$ with $k(n, c)$ independent of ω , then $d(n, c_0(n))$ is non-increasing.

Proof:

From

$$(24) \quad \Delta_n d(n, c) = -k(n, c)d'_c(n, c)$$

we have

$$(25) \quad d(n+1, c_0(n)) \leq d(n, c_0(n))$$

since $d'_c(n, c) \geq 0$ for $c = c_0(n)$. But $d(n+1, c)$ is minimized when $c = c_0(n+1)$ and hence

$$(26) \quad d(n+1, c_0(n+1)) \leq d(n+1, c_0(n)) .$$

Combining (25) and (26) we get $d(n+1, c_0(n+1)) \leq d(n, c_0(n))$ which proves the proposition.

The lemma applies when $f(x|\omega)$ is the exponential distribution and therefore the conclusion of the lemma also holds when $f(x|\omega)$ is a gamma distribution with integer shape parameter. Moreover it follows from a result by Hald and Thyregod (1971) that $d(n, c_0(n))$ is non-increasing when $f(x|\omega)$ is the Poisson distribution. The following example shows, however, that the result is not generally true.

Example 2.

Let $f(x|\omega)$ be the Bernoulli distribution, $f(x|\omega) = \omega^x(1-\omega)^{1-x}$ for $x = 0, 1$. It is well known that the distribution of S_n is binomial, $f^{(n)}(x|\omega) = \binom{n}{x}\omega^x(1-\omega)^{n-x}$ for $x = 0, 1, \dots, n$, and that the waiting-time distribution is the Pascal distribution, $p_c(n|\omega) = \binom{n}{c}\omega^{c+1}(1-\omega)^{n-c}$ for $n = c, c+1, \dots$ such that we have

$$(27) \quad p_c(n|\omega) = \frac{n+1}{c+1} f^{(n+1)}(c+1|\omega)$$

and therefore,

$$(28) \quad -\Delta_n d(n, c) = \frac{n+1}{c+1} \Delta_c d(n+1, c),$$

which leads to

$$(29) \quad d(n+1, c+1) - d(n, c) = \Delta_c d(n+1, c) + \Delta_n d(n, c) = -\frac{n-c}{c+1} \Delta_c d(n+1, c).$$

Consider now $n_1 > 0$ and let $c = c_0(n_1)$. From (28) we find that $d(n, c)$ is decreasing if $\Delta_c d(n+1, c) > 0$, i.e., as long as $c_0(n+1) = c$. Letting n_2 denote the first value of n such that $\Delta_c d(n+1, c) < 0$ we conclude from (29) that $d(n_2+1, c+1) > d(n_2, c_0(n_2))$. Now it is easy to construct cases where $c_0(n_2+1) = c + 1$, (an example has been provided in Example 3), and therefore we conclude that $d(n, c_0(n))$ will not in general be decreasing. However, in the next section we shall show that the possible optimum sample sizes correspond to decreasing sections of $d(n, c_0(n))$ and thus, in the present example we may simply rule out n_2 as a candidate for an optimum sampling plan.

4. The Regret Function and the Optimal Sampling Plan.

From (12) and (16) we find the regret

$$(30) \quad R(n, c, N) = n\delta_s + (N-n)d(n, c) \quad \text{for } n = 0, 1, \dots, N,$$

where we have defined the loss corresponding to the best singular strategy, $n = 0$, as $d(0, c) = \min\{\delta_a, \delta_r\}$ and $c_0(0) = 0$.

To simplify the discussion of R as function of the lot size N , we shall formally extend the definition above to cover non-negative real values of N . We shall put

$$(31) \quad R(n, c, N) = \begin{cases} N\delta_s & \text{for } 0 \leq N \leq n \\ n\delta_s + (N-n)d(n, c) & n < N \end{cases} .$$

The graph of (31) is a polygonal line consisting of two segments with slopes δ_s and $d(n, c)$. Since we have assumed $\max\{\delta_r, \delta_a\} \leq \delta_s$ it follows from (20) that R is a concave function of N .

Now, define

$$(32) \quad R_0(N) = \inf_{n, c} R(n, c, N) \quad \text{for } N \geq 0 .$$

Obviously R_0 is a concave function of N and moreover R_0 may be obtained as the pointwise minimum of the family $R(n, c_0(n), N)$, $n = 0, 1, 2, \dots$. Since $R(n, c_0(n), N) = N\delta_s$ for $n \geq N$ and $R(n, c_0(n), N) \leq N\delta_s$ if $n < N$ it is clear that the minimum exists and that it is obtained for a sample size $n(N)$ satisfying $0 \leq n(N) \leq N$. The following lemma gives a necessary condition for the optimal sample size.

Lemma 3.

Let $R_0(N) = R(n_1, c_0(n_1), N)$ for some $N > 0$. If $R_0(N) < N\delta_s$ then

$$(33) \quad d(n, c_0(n)) > d(n_1, c_0(n_1)) \quad \text{for } 0 \leq n < n_1 .$$

Proof:

Let $0 \leq n < n_1$ and consider

$$\begin{aligned} \rho &= R(n_1, c_0(n_1), N) - R(n, c_0(n), N) \\ &= (n_1 - n)\{\delta_s - d(n, c_0(n))\} + (N - n_1)\{d(n_1, c_0(n_1)) - d(n, c_0(n))\} . \end{aligned}$$

Since n_1 is optimal it follows that $\rho \leq 0$ and furthermore we find from $R_0(N) < N\delta_s$ that $n_1 < N$. Thus

$$(34) \quad d(n_1, c_0(n_1)) - d(n, c_0(n)) \leq (n-n_1)\{\delta_s - d(n, c_0(n))\}/(N-n_1) \leq 0.$$

It is easy to verify that the middle term in (34) is zero only when $d(n, c_0(n)) = \delta_s$ and since by assumption $d(n_1, c_0(n_1)) < \delta_s$ we finally get (33).

Theorem 3.

Let $(n(N), c(N))$ denote the optimal sampling plan for a lot of size N . Then there exists a version of $n(N)$ which is non-decreasing with N . Moreover $c(N) = c_0(n(N))$ is non-decreasing for this choice of $n(N)$.

Proof:

The graph of $R_0(N)$ is a polygonal line with the property that the slope of the line connecting successive vertices is decreasing. The slope of $R_0(N)$ is given by $d(n, c_0(n))$ with $n = n(N)$ if N is not a vertex.

Consider $N_1 < N_2$ not belonging to the set of vertices and let $n_1 = n(N_1)$, $n_2 = n(N_2)$. Since R_0 is concave we conclude that $d(n_1, c_0(n_1)) \geq d(n_2, c_0(n_2))$. If we tentatively assume that $n_2 < n_1$, the preceding lemma leads us to the contradictory result $d(n_1, c_0(n_1)) < d(n_2, c_0(n_2))$. Hence we must have $n_1 \leq n_2$. If $R_0(N_1) = N_1 \delta_s = N_1 d(0, c_0(0))$, we may choose $n(N_1) = 0$ such that the conclusion is also valid in this case.

The proof of the monotonicity of $n(N)$ is now immediate, noting that we may choose $n(N) = n(N-)$ if $(N, R_0(N))$ is a vertex.

Finally, the monotonicity of $c(N)$ follows from Theorem 1.

Example 3.

Assume that each item in a lot may be classified as either defective or acceptable and assume that lots originate from a Bernoulli-process where the process-parameter ω varies from lot to lot according to the beta-distribution $dW(\omega) = 4(1-\omega)^3 d\omega$ for $0 \leq \omega \leq 1$. Finally let the costs per item of inspection, rejection and acceptance be given by the following table

Action	Quality of item	
	acceptable	defective
inspect	0.2	0.2
reject	0.2	0.2
accept	0	1

We then find $k_a(\omega) = \omega$, $k_s(\omega) = k_r(\omega) = 0.2$ so the break-even quality $\omega_r = 0.2$ and moreover we have the same loss for all singular strategies $\delta_a = \delta_r = \delta_s = 0.065536$.

The marginal distribution of the number of defective items in the sample is the Polya distribution,

$$(35) \quad f^{(n)}(x) = \binom{n+3-x}{3} / \binom{n+4}{4} \quad \text{for } x = 0, 1, \dots, n,$$

and the posterior risk is simply the difference between the mean in the posterior distribution of ω and the cost of rejecting an item, i.e.,

$$(36) \quad \lambda(n, x) = \frac{1+x}{n+5} - 0.2 \quad \text{for } x = 0, 1, \dots, n.$$

From (18), (35) and (36) we obtain the decision loss

$$(37) \quad d(n, c) = 0.065536 + \sum_{x=0}^x \left(\frac{1+x}{n+5} - 0.2 \right) \binom{n+3-x}{3} / \binom{n+4}{4}$$

and the regret function

$$(38) \quad R(n, c, N) = \begin{cases} 0.065536n + (N-n)d(n, c) & \text{for } c = 0, 1, \dots, n \\ & n = 1, 2, \dots, N \\ 0.065536N & \text{for } n = 0 \text{ or } n > N. \end{cases}$$

It is easy to verify that $\lambda(n, x) < 0$ iff $x < n/5$ and therefore $c_0(n)$ is the largest integer strictly less than $n/5$ or, equivalently $c_0(n) = c$ for

$$(39) \quad 5(c-1) \leq n < 5c, \quad c = 0, 1, 2, \dots$$

It should furthermore be noted that the remark made in Example 2 applies in this case since $c_0(n)$ runs through consecutive integers, and therefore we may exclude the left endpoint of (39) from the set of possible optimal plans. In Table 1 we have listed the reduced set of possible plans for $n < 30$. The table also gives the decision loss corresponding to these plans.

In order to tabulate $n(N)$ we shall determine $R_0(N)$ by successively determining

$$R_0^*(i, N) = \min_{n \leq i} \min_c R(n, c, N)$$

for $i = 0, 1, 2, \dots$. The graph of R_0^* is a polygonal line with decreasing slopes. Moreover $R_0^*(i, N)$ has at most one point of intersection with $R(i+1, c_0(i+1), N)$ and therefore we may easily determine $R_0^*(i+1, N)$ from its predecessor. The first few steps of this procedure lead to

$$\begin{aligned} R_0^*(0, N) &= R(0, 0, N), \\ R_0^*(1, N) &= R(1, 0, N) \quad \text{for } 1 \leq N \end{aligned}$$

and

TABLE 1

Table of $d(n, c_0(n))$ for $n < 30$. The points of increase have been excluded. $k_s(w) = k_r(w) = 0.2$, $k_a(w) = w$ and $dW(w) = 4(1-w)^3 dw$.
 $f(x|w) = w^x(1-w)^{1-x}$ for $x = 0, 1$.

n	$c_0(n)$	$d(n, c_0(n))$
0	0	.06553600
1	0	.03886933
2	0	.02744076
3	0	.02267886
4	0	.02109156
6	1	.01705115
7	1	.01462691
8	1	.01332155
9	1	.01278875
11	2	.01114040
12	2	.01007382
13	2	.00945757
14	2	.00918925
16	3	.00829486
17	3	.00769442
18	3	.00733517
19	3	.00717350
21	4	.00661232
22	4	.00622719
23	4	.00599184
24	4	.00588378
26	5	.00549893
27	5	.00523090
28	5	.00506476
29	5	.00498744

$$R_0^*(2, N) = \begin{cases} R_0^*(1, N) & \text{for } N \leq 4 \\ R(2, 0, N) & \text{for } 5 \leq N. \end{cases}$$

Continuing the process we finally obtain the list of optimal sampling plans displayed in Table 2.

TABLE 2

Optimal sampling plans for the Bernoulli distribution $k_r(\omega) = k_s(\omega) = 0.2$, $k_a(\omega) = \omega$ and $dW(\omega) = 4(1-\omega)^3 d\omega$.

N	n	c
1 - 4	1	0
5 - 11	2	0
12 - 28	3	0
29 - 47	7	1
48 - 76	8	1
77 - 102	12	2
103 - 144	13	2
145 - 179	17	3
180 - 232	18	3
233 - 275	22	4
276 - 340	23	4
341 - 390	27	5
391 - 467	28	5

It is obvious that the algorithm suggested in the above example will work for any combination of loss function, prior distribution and sampling distribution satisfying the assumptions of the present paper. The algorithm is well-suited for description in a programming language, allowing the use of subroutines to compute δ_a , δ_r , δ_s and $d(n, c_0(n))$. In a tentative FORTRAN-program we used one single subroutine to compute

the decision loss for the case with linear costs, Bernoulli sampling distribution and a beta prior, another subroutine to cover the analogous situation but with a two-point prior; the remaining parameters of the cost functions and the prior distribution were given as input parameters to the program. The program needed about 20 seconds of computing time on an IBM 7090 computer to produce a table of optimal sampling plans for $N \leq 200000$.

5. References.

- Blum, J. R. and Hanson, D. L. (1960). On invariant probability measures I. Pacific J. Math. 10 1125-1129.
- Guthrie, D. and Johns, M. V. (1959). Bayes acceptance sampling procedures for large lots. Ann. Math. Statist. 30 896-925.
- Hald, A. (1960). The compound hypergeometric distribution and a system of single sampling inspection plans based on prior distributions and costs. Technometrics 2 275-352 and 370-372.
- _____ (1965). Bayesian single sampling attribute plans for discrete prior distributions. Mat. Fys. Skr. Dan. Vid. Selsk. 3 No. 2, 88 pp. Munksgaard, Copenhagen.
- _____ (1967). Asymptotic properties of Bayesian single sampling plans. J. Roy. Statist. Soc. B 29 162-173. Corrigenda on p. 586.
- Hald, A. and Thyregod, P. (1971). Bayesian single sampling plans based on linear costs and the Poisson distribution. To appear in Mat. Fys. Skr. Dan. Vid. Selsk. 3 No. 7, Munksgaard, Copenhagen.
- Karlin, S. (1968). Total Positivity, Vol. I. Stanford Univ. Press, Stanford.
- Raiffa, H. and Schlaifer, R. (1961). Applied statistical decision theory. Division of Research, Graduate School of Business Administration, Harvard University, Boston.