ECONOMIC DESIGN OF DOUBLE SAMPLING PLANS

A THESIS

Presented to

The Faculty of the Division of Graduate Studies

By

Ralph Dennis Stewart

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in the School of

Industrial and Systems Engineering

Georgia Institute of Technology

June, 1976

ECONOMIC DESIGN OF DOUBLE SAMPLING PLANS

Approved:

Douglas C. Montgomery, Chairman

Russell G. Heikes

Harrison M. Wadsworth

Date approved by Chairman: 5-24-76

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION AND LITERATURE SURVEY

1.1 Sampling

Acceptance sampling is one of the major fields of statistical quality control. It can protect consumers from receiving and producers from distributing a product of low quality. Variables sampling is a test to see whether or not a parameter related to quality falls within a certain range. Attribute sampling is a procedure that determines whether or not a particular quality characteristic is present or absent in a unit of product.

The type of sampling investigated in this research will be attribute sampling. That is, each item will be judged either defective or nondefective on the basis of examination of a single attribute. The purpose of the sampling is not to determine lot quality but to make a decision about the disposition of lots. A single sampling plan is defined by a sample size n and a rejection number c. A sample of size n is taken from a lot; and the lot is rejected if there are more than c defectives in the sample. The discriminatory power of a sampling plan can be measured by its operating characteristic curve. That is, a plot of the probability of lot acceptance versus the fraction defective of the lot. Traditionally, sampling plans have been chosen by finding the plan whose operating characteristic curve passes through two predetermined points. For instance, the probability of accepting a lot that has a fraction defective of 0.01 might be chosen to be 0.95. While we might wish the probability of accepting a lot with fraction defective of 0.07 to be 0.05. These two requirements specify two points on the operating characteristic curve, namely, (0.01,0.95) and (0.07,0.05). These two points are usually called the producer's risk point and the consumer's risk point respectively. Figure 1 shows a graph of an operating characteristic curve with these two points.

A double sampling plan specifies two sample sizes n_1 and n_2, two acceptance numbers c_1 and c_2 and two rejection numbers r_1 and r_2. A sample of size n_1 is taken and if c_1 items or less are found defective, the lot is immediately accepted. If r_1 or more items are defective, it is rejected. If the number of defectives is more than c_1, but less than r_1, a second sample of n_2 items is taken. If the total number of defectives is c_2 or less the lot is accepted. If there are r_2 or more defectives the lot is rejected. In this study r_1 and r_2 are both considered equal to $c_2 + 1$ so that the four numbers n_1, n_2, c_1, and c_2 will completely specify the sampling procedure.

1.2 Fully Economic Schemes

This research is concerned with double sampling not in the traditional sense but in the economic sense. By this, we mean that economic criteria are used in the design of the sampling plan, rather than only statistical considerations. Schemes of this type are generally based on estimates of costs involved in the sampling process and costs incurred as a result of errors in the decision process. In addition, a distribution of lot quality must be specified. The general procedure is

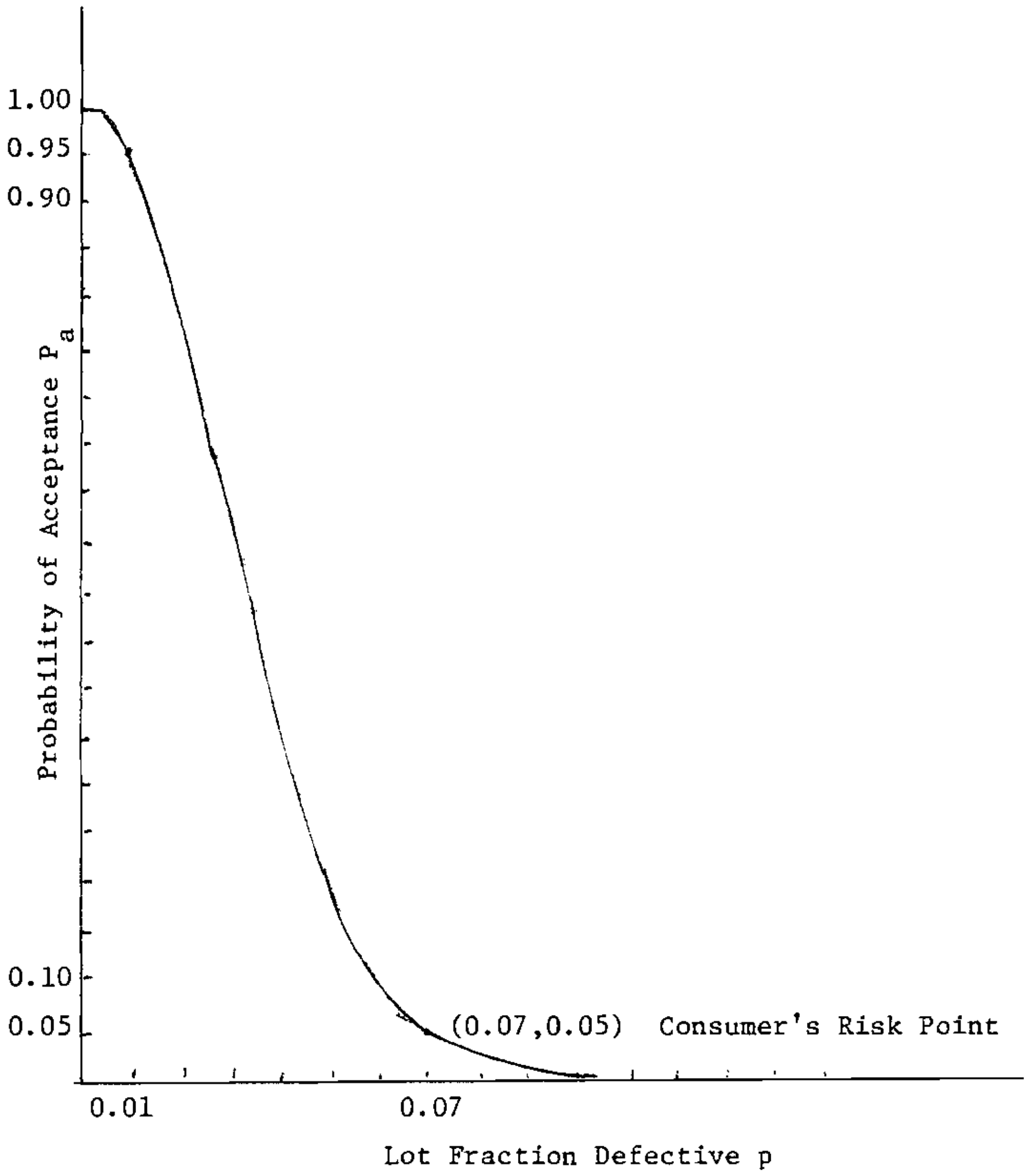


Figure 1. Operating Characteristic Curve

to decide on a risk function of the form:

$$R(p) = (\text{cost of sampling}) + (\text{cost of loss due to decision}).$$

Here p is a variable used to designate the lot fraction defective. The function R gives the expected cost if the fraction defective is actually p. Integrating R over the prior distribution of lot quality gives an average cost which is independent of p. Some of the work done previously on fully economic schemes is discussed in this chapter. Wetherill and Chiu (20) have compiled a thorough bibliography of papers dealing with acceptance sampling schemes with emphasis on the economic aspect. Numbers in parentheses refer to the bibliography.

Anscombe (2) gives a general discussion of the economic approach to sampling. He describes simple hypothetical process curves, inspection cost curves, and decision loss curves. He also relates this procedure to the concepts of AOQL and lot tolerance.

Champernowne (3) considers the problem of deriving sequential sampling plans that minimize the sum of decision and inspection costs. He uses the beta distribution as the prior distribution of lot quality. His plans are based on a critical fraction defective, p_0, where decision costs are zero and a multiplicative factor that relates the decision costs to the difference between p_0 and the fraction defective.

Guthrie and Johns (5) introduce a general linear cost model. They then find asymptotic characterizations for large lot sizes of the decision procedure and sample sizes, which are optimal in the Bayes sense for various prior distributions.

Hald (6) and Hald (8) give a very detailed discussion of the problems involved in single sampling fully economic schemes. His paper will be discussed in detail in Section 1.3. He presents numerical results using the beta prior distribution and studies the results obtained by varying parameters. He also studies the behavior of the optimum plan with respect to large lot sizes.

Pfanzagl (15) investigates some consequences that occur when assumptions concerning the prior distributions are modified. He also considers the efficiency of optimum double sampling procedures compared to optimum single sampling procedures. This paper is primarily based on the results of Hald (6).

Schüler (18) investigates the Bayesian design of a k-stage sampling procedure. At each stage j (j=1,...,k-1) the size of the (j+1)st sample is given as a function of the outcomes of the first j samples. If a sample size is ever determined to be zero or after the kth sample, a decision is made regarding the disposition of the lot.

Schmidt and Bennett (16) study the case of multiple attributes in single sampling under the assumption that rejected lots are scrapped. A mathematical model is presented and optimized by a search technique. They also present sensitivity analysis of the optimum sample size to changes in the assumed prior distribution.

Bennett, Case and Schmidt (17) present two cost models for acceptance sampling by variables. The models differ in the disposition of rejected lots, which can either be scrapped or totally inspected. The sensitivity of the model to errors in the prior distributions is examined.

Ailor, Schmidt, and Bennett (1) derive an economic acceptance sampling model for the case where several attributes and several variables are simultaneously subjected to acceptance sampling. Testing can either be destructive or nondestructive. In the case of non-destructive testing rejected lots can either be scrapped or screened.

Hald (9) presents optimum double sampling tests of given strength for a normally distributed random variable with unknown mean and known variance. The two optimality criteria used are Bayes and minimax average sample number. Hald also discusses approximations to optimum double sampling tests for Poisson and binomially distributed variables.

1.3 The Hald Model

In 1960 Hald gave a very detailed description of the problems involved in sampling inspection schemes based on estimates of prior distribution of lot quality and costs. His work is the foundation for most of the work subsequently done in sampling based on economic considerations. Hald considered the average loss caused by an accepted defective item as his economic unit. His cost parameters then were the cost of sampling a single item, k_s and the cost of scrapping a single item, k_r. Rejected lots are scrapped. Let N be the size of the lot, X be the number of defectives in the lot, n be the sample size, and x the number of defectives in the sample; then the cost for accepted lots is:

$$nk_s + (X - x).$$

The cost for rejected lots is:

$$nk_s + (N - n)k_r.$$

The probability of getting x defectives from a sample of n items is

$$p(x|X) = C(X,x) \quad C(N - X, n - x)/C(N,n).$$

Where $C(\mu,v)$ denotes the number of combinations of μ things taken v at a time. The average costs for lots of quality $p = X/N$ becomes

$$K(n,c,p) = nk_s + \sum_{x=0}^{c}(X - x)p(x|X) + (N - n)k_r \sum_{x=c+1}^{n} p(x|X).$$

Then the probability of acceptance $P_a(p)$ is

$$P_a(p) = \sum_{x=0}^{c} P(x|X).$$

Introducing this and dividing by N gives

$$K(n,c,p)/N = pP_a(p) + k_r(1-P_a(p) + (n/N)(k_s-k_r + \sum_{x=0}^{c}(k_r-x/N)p(x|X)).$$

To find the expected final costs the cost function is averaged over the prior distribution. Let $f_N(X)$, $X = 0,1,2,3,\ldots,N$, be the probability that a lot of N items contains X defective items. Then the corresponding cumulative distribution is

$$F_N(Np) = \sum_{x=0}^{[N_p]} f_N(X).$$

The related distribution for p is

$$\Phi_N(p) = F_N(Np).$$

After this derivation, Hald studies the distribution of defectives in a sample of size n given the distribution of defectives in the lot, $f_N(X)$. His major contribution is the idea of reproducibility and the major theorem is as follows.

> Let X denote the number of elements having a certain attribute in a population of N elements and let x and y=X-x denote the corresponding numbers of elements in a random sample (drawn without replacement) of size n and in the remainder of the population respectively. If the distribution of X is a hypergeometric, a binomial, a rectangular, a Polya, or a mixed binomial distribution, or any weighted average of these distributions with weights independent of N and X, then for any N the distribution of x is the same as the distribution of X with n substituted for N, and the distribution of y for any given x is also of the same type but with parameters depending on x and n.

This theorem shows that the lot distribution reproduces itself in the sample distribution for a certain class of prior distributions. Hald uses this result to find approximately optimal solutions for the various prior distributions and their properties are studied under changes of lot size and prior parameters.

The major differences in approach between Hald and this research occur in the approach to the prior distribution and the procedure for optimizing the cost function. Hald uses as a prior distribution $f_N(X)$, the distribution of defectives in the lot. Here a prior distribution for the process quality p, say $\Phi(p)$, is used. Hald optimizes the cost function by finding a set of inequalities in n anc c that bound a region that contains the optimal solution. Here a search technique utilizing a computer algorithm is used to evaluate the cost function and find the optimum sampling plan.

CHAPTER II

DEVELOPMENT OF THE MODEL

2.1 The Cost Structure

Economic schemes are based on the costs that occur because sampling is used. This model contains four cost components:

k_i = variable cost of sampling or testing one item for the presence of an attribute,

k_I = fixed cost of sampling inspection,

k_s = unit cost of action taken on rejected items,

k_a = unit cost of accepting a defective item.

The variable cost of sampling, k_i, includes all direct sampling costs attributed to a unit of product. The fixed cost, k_I, contains all direct and indirect costs that result from the existence of a sampling procedure which are independent of the sample sizes n_1 and n_2.

The acceptance cost arises because of the presence of defective items in even lots of high quality. The cost, k_a, contains all costs, such as failure of the finished product, that arise when a defective item is used for its intended purpose.

The cost of action taken on rejected items can take many forms depending on the destination of items found defective or of lots found to be of inferior quality. It may be the cost of diverting a lot or item to a less stringent job requirement, the cost of repairing defectives or replacing them from a known stock of nondefective items, or it may be the scrapping costs minus the salvage value. In this model the cost of rejection is paid by the tester. Therefore the case of a rejected lot that is to be returned to a vendor can be handled by setting k_s equal to the cost incurred by the tester because of delay minus penalty incurred by the vendor because of the rejected lot.

2.2 The Probabilities of Lot Acceptance and Rejection

As a result of applying a double sampling plan, four actions may be taken and the corresponding probabilities of these actions must be evaluated. These actions are either acceptance or rejection on either the first or second sample. Let N be the lot size, n_1 and n_2 the first and second sample sizes, c_1 and c_2 the first and second acceptance numbers. Then let $X = [pN + 1/2]$, where $[Z]$ is the greatest integer contained in Z. The probability of acceptance on the first sample, P_{a1}, is

$$P_{a1}(p) = \sum_{b=0}^{Min(c_1,X)} \frac{\binom{X}{b}\binom{N-X}{n_1-b}}{\binom{N}{N_1}}$$

The probability of rejection on the first sample is

$$P_{r1}(p) = \begin{cases} \displaystyle\sum_{b=c_2+1}^{\mathrm{Min}(n_1,X)} \frac{\binom{X}{b}\binom{N-X}{n_1-b}}{\binom{N}{n_1}} & \text{if } X \geq c_2 + 1 \\ 0 & \text{if } X \leq c_2 \end{cases}$$

The probability of acceptance on the second sample is

$$P_{a2}(p) = \begin{cases} \displaystyle\sum_{b=c_1+1}^{\mathrm{Min}(c_2,X)} \frac{\binom{X}{b}\binom{N-X}{n_1-b}}{\binom{N}{n_1}} \left(\sum_{j=0}^{\mathrm{Min}(c_2-b,\ X-b)} \frac{\binom{X-b}{j}\binom{N-n_1-(X-b)}{n_2-j}}{\binom{N-n_1}{n_2}} \right) & \text{if } X \geq c_1 + 1 \\ 0 & \text{if } X < c_1 + 1 \end{cases}$$

Finally the probability of rejection on the second sample is

$$P_{r2}(p) = \begin{cases} \displaystyle\sum_{b=c_1+1}^{\mathrm{Min}(c_2,X)} \frac{\binom{X}{b}\binom{N-X}{n_1-b}}{\binom{N}{n_1}} \left(\sum_{j=c_2-b+1}^{\mathrm{Min}(n_2,X)} \frac{\binom{X-b}{j}\binom{N-n_1-(X-b)}{n_2-j}}{\binom{N-n_1}{n_2}} \right) & \text{if } X \geq c_1 + 1 \\ 0 & \text{if } X < c_1 + 1 \end{cases}$$

2.3 The Expected Cost

Let the number of defective items found on the first sample be d_1 and the number of defective items found on the second sample be d_2. Then the cost of a lot containing X defectives that is accepted on the first sample can be divided into the cost of inspection, the cost of scrapping items found defective during sampling, and the cost of accepting defective items in the remainder of the lot. The inspection cost is the fixed cost of sampling, k_I, plus the unit cost of sampling, k_i, times the number of items sampled, n_1. The cost of scrapping items found defective during sampling is the cost of scrapping one item, k_s, times the number of defective items d_1. The cost of accepting defective items in the remainder of the lot is the unit cost, k_a, times the number of defective items in the rest of the lot, $X - d_1$. The cost of a lot accepted on the first sample is then

$$C_{a1} = k_I + k_i n_1 + k_s d_1 + k_a(X - d_1)$$

If a lot is accepted on the second sample then the number of items sampled is $n_1 + n_2$, the number of defective items found during sampling is $d_1 + d_2$ and the number of defectives accepted is $X - d_1 - d_2$. The cost of a lot accepted on the second sample is then

$$C_{a2} = k_I + k_i(n_1 + n_2) + k_s(d_1 + d_2) + k_a(X - d_1 - d_2).$$

If a lot is rejected, then it is assumed that all non-defective items found in testing are kept and all other items are disposed of in some way. The cost of inspection and the cost of scrapping items found defective during sampling are identical to the corresponding costs in acceptance on the first sample. Instead of the cost of accepting defectives in the remainder of the lot, there is the cost of disposing of the remainder of the lot. The cost of a lot which is rejected on the first sample is

$$C_{r1} = k_I + k_i n_1 + k_s d_1 + k_s(N - n_1).$$

If a lot is rejected on the second sample, the number of items sampled is $n_1 + n_2$, the number of defectives found during sampling is $d_1 + d_2$ and the number of items in the portion of the lot to be scrapped is $N - n_1 - n_2$. The cost of a lot rejected on the second sample is

$$C_{r2} = k_I + k_i(n_1 + n_2) + k_s(d_1 + d_2) + k_s(N - n_1 - n_2).$$

It would be easier to evaluate the cost equations if d_1 and d_2 could be approximated by functions of p. If the probability that an item is defective is p and the size of the sample is n then the average number of defectives in the sample will be np. Therefore d_1 can be approximated by $n_1 p$ and d_2 by $n_2 p$. Also X will be approximated by pN. The four cost functions are then

$$C_{a1}(p) = k_I + k_i n_1 + k_s n_1 p + k_a(N - n_1)p,$$

$$C_{a2}(p) = k_I + k_i(n_1 + n_2) + k_s(n_1 + n_2)p + k_a(N - n_1 - n_2)p,$$

$$C_{r1}(p) = k_I + k_i n_1 + k_s n_1 p + k_s(N - n_1),$$

$$C_{r2}(p) = k_I + k_i(n_1 + n_2) + k_s(n_1 + n_2)p + k_s(N - n_1 - n_2).$$

These cost functions are then weighted with the probabilities on the actions of acceptance and rejections. This gives cost as a function of p

$$C(p) = C_{a1}(p)P_{a1}(p) + C_{a2}(p)P_{a2}(p) + C_{r1}(p)P_{r1}(p) + C_{r2}(p)P_{r2}(p).$$

The expected cost is given by C(p) averaged over the distribution f(p)

$$E(\text{cost}) = \int_0^1 C(p)f(p)dp.$$

Here p represents the process quality and f(p) is a function that describes the behavior of that random variable in a particular lot. The general form of f(p) used in this research is the beta distribution. The beta distribution is a good continuous approximation of the discrete distributions of lot quality, and it has been used by other researchers. Writing the expected cost in expanded terms gives

$$E(\text{cost}) = \int_0^1 ((k_I + k_i n_1 + k_a p(N-n_1) + k_s p n_1)P_{a1}(p) +$$

$$(k_I + k_i(n_1+n_2) + k_a p(N-n_1-n_2) + k_s p(n_1+n_2))P_{a2}(p) +$$

$$(K_I + k_i n_1 + k_s p n_1 + k_s(N-n_1))P_{r1}(p) + \qquad [1]$$

$$(k_I + k_i(n_1+n_2) + k_s(n_1+n_2)p + k_s(N-n_1-n_2))P_{r2}(p))f(p)dp.$$

2.4 Default Options

If one had perfect information about the quality of the lot, then there would be no need for sampling inspection. Two cases of interest might be the acceptance or rejection of the lot without sampling. In the case of lot acceptance without sampling, the conditional expected cost given p would be

$$E(\text{cost of acceptance without sampling}|p) = k_a Np.$$

The cost in the rejection case is

$$E(\text{cost of rejection without sampling}) = k_s N.$$

Figure 2a shows a graph of these two functions. The two lines intersect at a value of p equal to k_s/k_a. Hald (4) calls this point the break-even quality. For p less than this value it is less expensive to accept the lot without sampling and for p greater than this value it is less expensive to reject it. The average expected cost for acceptance is

$$E(\text{cost of acceptance without sampling}) = \int_0^1 k_a N p f(p)\, dp$$

$$= k_a N\bar{p},$$

where $\bar{p}$ is the mean of the prior distribution on p.

Another alternative might be a decision to totally inspect the lot. This decision might result in a lower per unit sampling cost than any individual sampling plan might produce. Letting k_t represent this sampling cost the expected cost is

$$E(\text{cost of total inspection}|p) = k_I + k_t N + k_s pN.$$

If the inequality

$$k_I + k_t N < (k_s/k_a)(k_a - k_s)N$$

is true, then there will be values for p which result in total inspection as the least expensive default procedure. The average expected cost of total inspection is

$$E(\text{cost of total inspection}) = \int_0^1 (k_I + k_t N + k_s Np) f(p)\, dp$$

$$= k_I + k_t N + k_s N\bar{p}.$$

In Figures 2b and 2c the curves labeled C_a, C_r and C_i designate cost curves for acceptance, rejection, and inspection respectively. Figure 2b shows the cost of total inspection as a function of p and gives the points of intersection with the other two curves. Figure 2c gives a numerical example for the default costs. For values of the lot fraction defective p less than $p_1 = 0.045$ acceptance without sampling gives the smallest expected cost. For values of p between 0.045 and 0.667 total inspection is optimal. For values of p greater than 0.667 the optimal default option is rejection without sampling.

2.5 Evaluation of the Cost Function

The first step in optimizing the cost function is solving [1] for an arbitrary sampling plan n_1, n_2, c_1, c_2. Examining P_{a1}, P_{a2}, P_{r1}, and P_{r2} we note that their general form is a step function whose interval width is determined by N and is in fact 1/N excepting the first and last interval whose width is 1/2N. The reason for this is the relationship assumed between X and p, namely

$$X = [pN + 1/2]$$

Then [1] may be viewed as a sum of Stiljes integrals with weighting functions P_{a1}, P_{a2}, P_{r1}, and P_{r2}.

The original intent was to apply numerical integration techniques to the model, but that method proved intractable in both the required computer time and the number of points required per integration. Since the intervals in the weighting functions were, for N equal to 1000, say

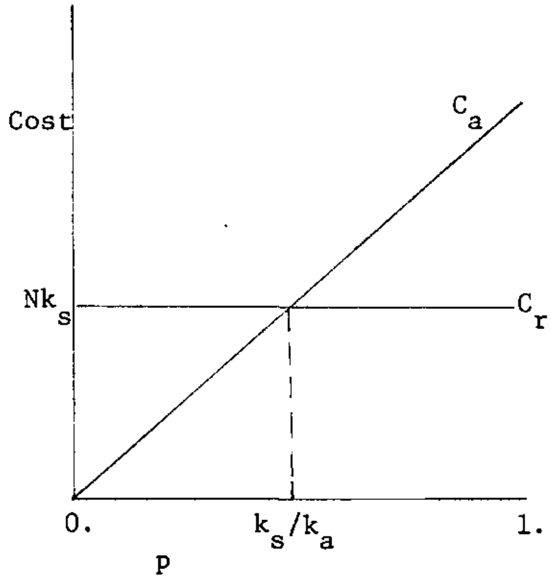
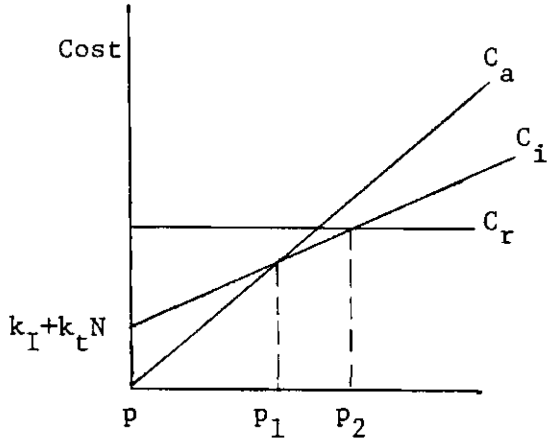


Figure 2a. Costs of Acceptance and Rejection Without Sampling



$$p_i = (k_I+k_tN)/(N(k_a-k_s))$$

$$p_2 = (Nk_s-(k_I+Nk_t))/(Nk_s)$$

Figure 2b. Cost of Total Inspection

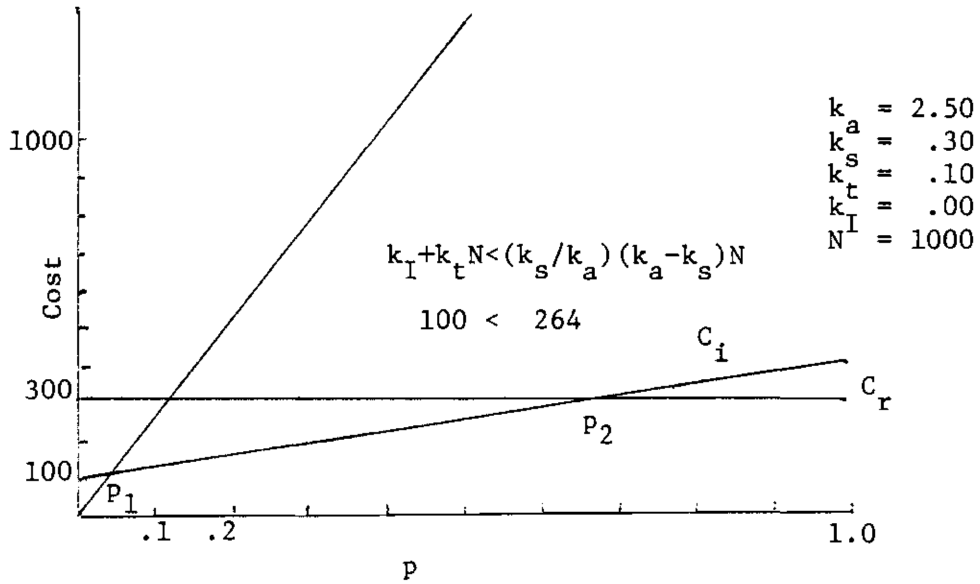


Figure 2c. Numerical Example of Default Costs

on the order of .001, more than 2000 points on the interval zero to one were required to achieve even an accuracy of one ten-thousandth. However to evaluate 2000 points required well over 15 minutes of central processing unit time. One source of computational inefficiency was the use of a library hypergeometric function. This was remedied by developing a hypergeometric function that made use of recursive relationships among the weighting functions. The number of points needed per evaluation was reduced by a different approach to the integration.

Consider the expected cost of acceptance on the first sample

$$C_1 = E(\text{cost of acceptance on the first sample})$$

$$\int_0^1 (k_I + n_1 k_i + k_s n_1 p + k_a (N-n_1)p) P_{a1}(p) f(p) dp.$$

Because of the relationship between X and p namely $X = [pN + 1/2]$, $P_{a1}(p)$ is a step function. Let R_x denote the region on the interval zero to one where X takes on the value x. If $N = 1000$, then R_0 is the interval [0.0, 0.0005) R_{55} is the interval [0.0545, 0.0555). The integral can then be rewritten

$$C_1 = \sum_{X=0}^{N} \int_{R_x} (k_I + n_1 k_i + k_s n_1 p + k_a (N-n_1)p) P_{a1}(p) f(p) dp.$$

Since X is constant over each interval R_x, $P_{a1}(p)$ is constant over the interval. Let the value of $P_{a1}(p)$ over the interval R_x be designated $W_{a1}(x)$. This quantity is independent of p and can be brought outside of the integral. The integral is

$$C_1 = \sum_{X=0}^{N} W_{a1}(x) \int_{R_x} (k_I + k_i n_1 + k_s n_1 p + k_a(N-n_1)p) f(p)dp.$$

Now given the form of the distribution $f(p)$, the integral can be evaluated directly. If the form of $f(p)$ is beta then the integrand is merely a weighted sum of beta functions. Similar results are obtained for the cases of acceptance on the second sample, rejection on the first sample, and rejection on the second sample. Letting $W_{a2}(x)$, $W_{r1}(x)$, and $W_{r2}(x)$ denote the value of $P_{a2}(p)$, $P_{r1}(p)$ and $P_{r2}(p)$ over the interval R_x respectively and noting that k_I and $k_i n_1$ appear in all four cost segments, equation [1] can be rewritten as

$$\begin{aligned} E(\text{cost}) = k_I + k_i n_1 + \sum_{X=0}^{N} & W_{a1}(x) \int_{R_x} (k_s n_1 p + k_a(N-n_1)p) f(p)dp + \\ & W_{a2}(x) \int_{R_x} (k_i n_2 + k_s(n_1+n_2) + k_a(N-n_1-n_2)p) f(p)dp \\ & + W_{r1}(x) \int_{R_x} (k_s n_1 p + k_s(N-n_1)) f(p)dp \\ & + W_{r2}(x) \int_{R_x} (k_i n_2 + k_s(n_1+n_2)p + k_s(N-n_1-n_2)) \\ & f(p)dp. \end{aligned}$$

This form proved computationally tractable and could be evaluated in a reasonable amount of computer time.

2.6 The Search Procedure

The original approach was to optimize the cost function using the Hook and Jeeves pattern search. This procedure consists of finding the function value at a base point, (n_1', c_1', c_2'). Then an exploratory move is made by evaluating the function at points corresponding to the base point incremented and decremented by a predetermined step size in each coordinate direction one at a time. When a decrease in the function value is detected, that value of the coordinate is substituted in a new base point, and the process is continued until all the coordinates have been tested. If the point found by the exploratory move is $(n_1^*, c_1^*, n_2^*, c_2^*)$, then the new base point is found by the formula

$$(n_1'', c_1'', n_2'', c_2'') = 2(n_1^*, c_1^*, n_2^*, c_2^*) - (n_1', c_1', n_2', c_2').$$

This is the pattern move. After this another exploratory move is made and the process is repeated. The optimization procedure continues until there is no decrease in function value following an exploratory move. The pattern search is described in more detail in Appendix 1.

The Hook and Jeeves pattern search works very well for a unimodal surface, but for a multimodal surface it must be restarted at several points in the factor space to insure that the global minimum is reached. This is very inefficient in terms of computer time and it is often hard to determine a stopping rule that guarantees that the global minimum is found. Alternatively, it is very inefficient to perform a full grid search of every point in the area in questions. The cost surface has many local minimum and is not well suited to a pattern search in all four variables.

From examination of grid search data, it was determined that the surface defined by

$$S = \{(n_1^*, c_1, n_2^*, c_2): n_1^* \text{ and } n_2^* \text{ are optimal for } c_1 \text{ and } c_2\}$$

was unimodal. This is a surface in only two variables, c_1 and c_2. If c_1 and c_2 were fixed then the remaining two variable surface was unimodal. Then the procedure evolved to a two variable pattern search. For each pair, (c_1, c_2) values of n_1 and n_2 are found that give minimum expected cost by a pattern search. Finally a pattern search in c_1 and c_2 will produce a global optimum.

CHAPTER III

NUMERICAL ANALYSIS OF THE SCRAPPING MODEL

3.1 Introduction

The expected cost is dependent on the cost vector by

$$E(\text{cost}|ak_I,\ ak_i,\ ak_a,\ ak_s) = aE(\text{cost}|k_I,\ k_i,\ k_a,\ k_s),$$

and

$$E(\text{cost}|k_I,\ k_i,\ k_a,\ k_s) = K_I + E(\text{cost}|0.0,\ k_i,\ k_a,\ k_s).$$

Both these equations follow directly from equation [2]. The second relationship allows us to assume any value for k_I without loss of generality. Then for any values of the two ratios k_i/k_a and k_s/k_a a unique point in the cost space is determined. Thus the optimal point in the sampling plan space is the same for both cost vectors

$$(k_I,\ k_i,\ k_a,\ k_s) = (0.,\ .1,\ .25,\ .1)$$

and

$$(k_I,\ k_i,\ k_a,\ k_s) = (0.,\ 1.,\ 2.5,\ 1.).$$

Also the expected cost for the first case is one-tenth that of the second.

If the default options of accepting and rejecting without sampling are considered the boundary of the solution space and all other sampling plans the interior of this space, then a necessary and sufficient condition for the model to have a solution in the interior of the space is that there exists one point that has an expected cost lower than that of the boundary. However if k_I is considered zero, then the condition can be simplified. If k_I is 0. and if and only if the point (N_1,C_1,N_2,C_2) = (1,0,0,0) has an expected cost less than that of both accepting and rejecting without sampling, then the optimal point lies in the interior of the sampling plan space. The point (1,0,0,0) corresponds to taking a sample of one and accepting the lot if it is not defective and rejecting the lot if it is defective. For this plan, the probability of acceptance given a population of N and a defective population of X is

$$P_a = 1 - X/N.$$

and the probability of rejection is

$$P_r = X/N.$$

Noting that X is a function of p, the expected cost is

$$E_{1,0,0,0}(\text{cost}) = \int_0^1 ((1-X(p)/N)(k_i+(N-1)pk_a+pk_s) + (X(p)/N)$$
$$k_i + (N-1)k_s + k_s p))f(p)dp$$

Let R_x be the interval of p where X=x, F(R) be the integral of the prior over the region R, and F'(R) be the integral of p times the prior over the region R. Then the expected cost can be written

$$E_{1,0,0,0}(\text{cost}) = k_i + \bar{p}(N-1)k_a + \bar{p}k_s + ((N-1)/N)\sum_{X=0}^{1000} x(F(R_x)$$
$$k_s - F'(R_x)k_a).$$

This expression, while complicated in form is easily evaluated; and if it is less than both the default costs, gives both a necessary and sufficient condition for an optimal sampling plan to exist at an interior point. Sufficiency is obvious and necessity follows from the description of the surface in 3.2. That is, if as one moves along the surface away from the boundary the cost increases, then the cost will continue to increase as one moves farther away from the boundary.

3.2 Description of the Surface

In Figures 3 to 8 the cost vector is

$$(k_I, k_i, k_a, k_s) = (1.00, .30, 2.50, .30).$$

The prior distribution is the Beta distribution described in Appendix 2 with parameters a = 1.15, b = 18.35. The mean of the distribution is .1 and the variance is .004. These figures provide an approximate idea of the surface surrounding the optimum point for these costs. It can be seen from Figure 4 the optimum is

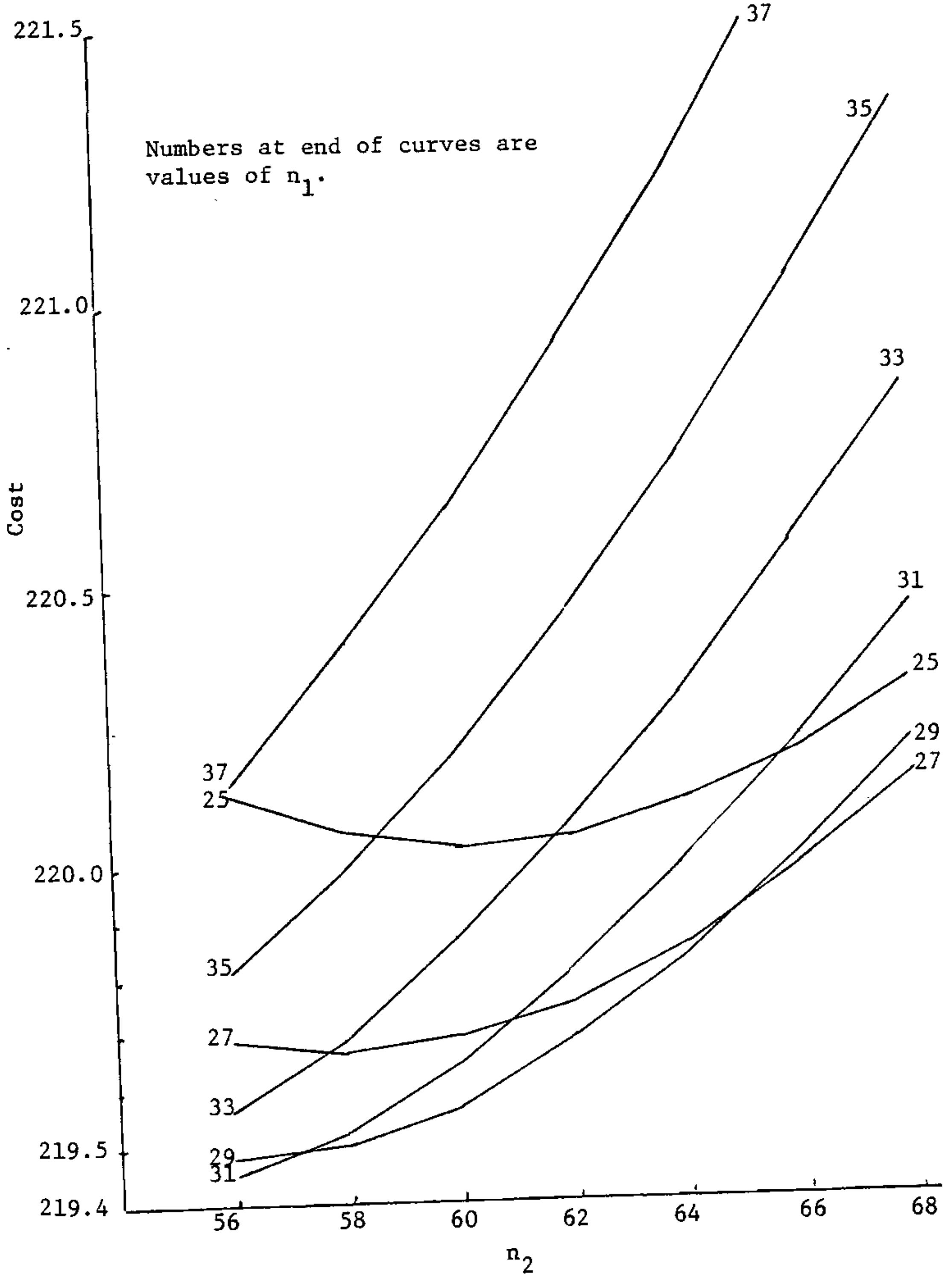


Figure 3. Expected Cost Curves for C_1=2 and C_2=10

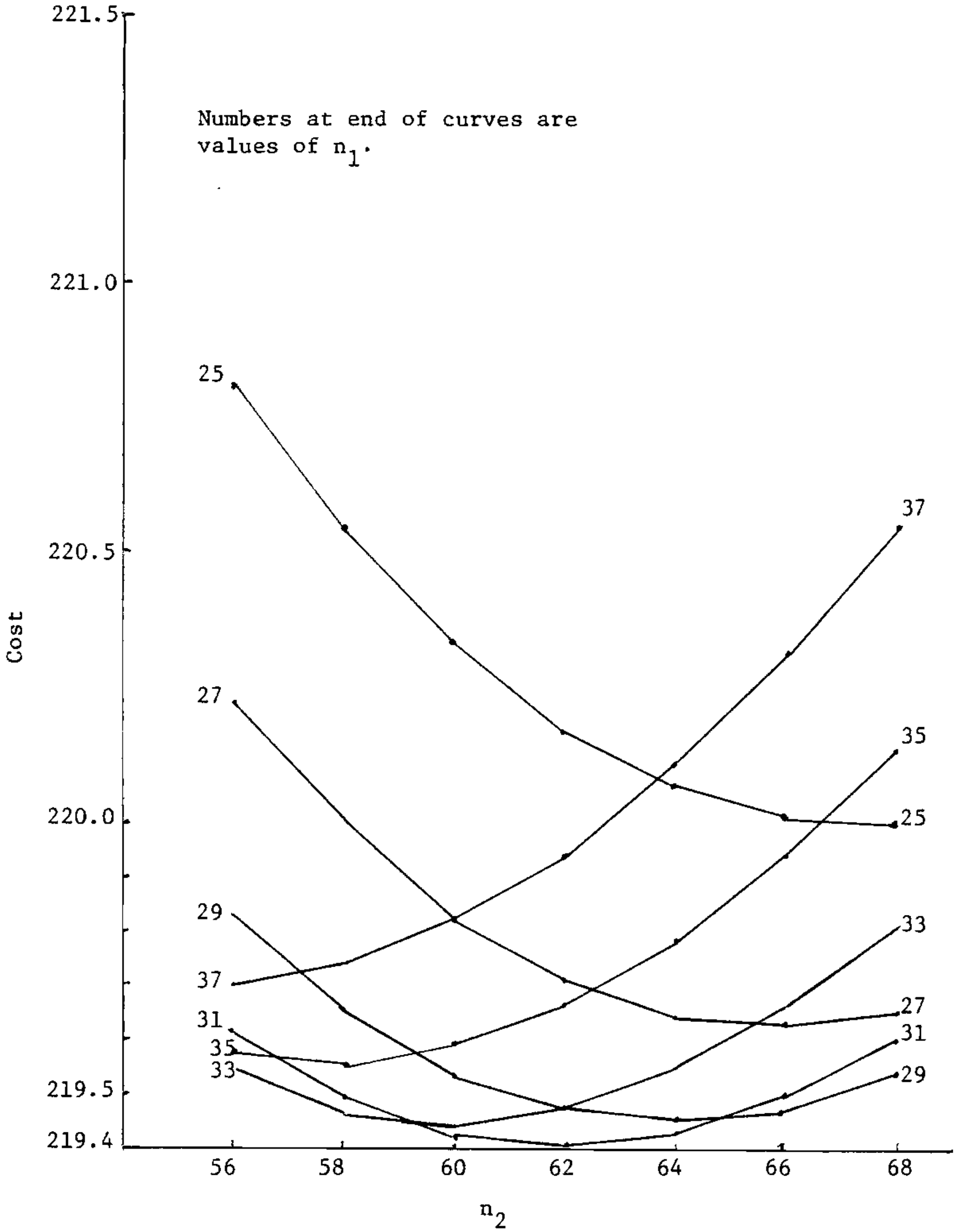


Figure 4. Expected Cost Curves for C_1=2 and C_2=11

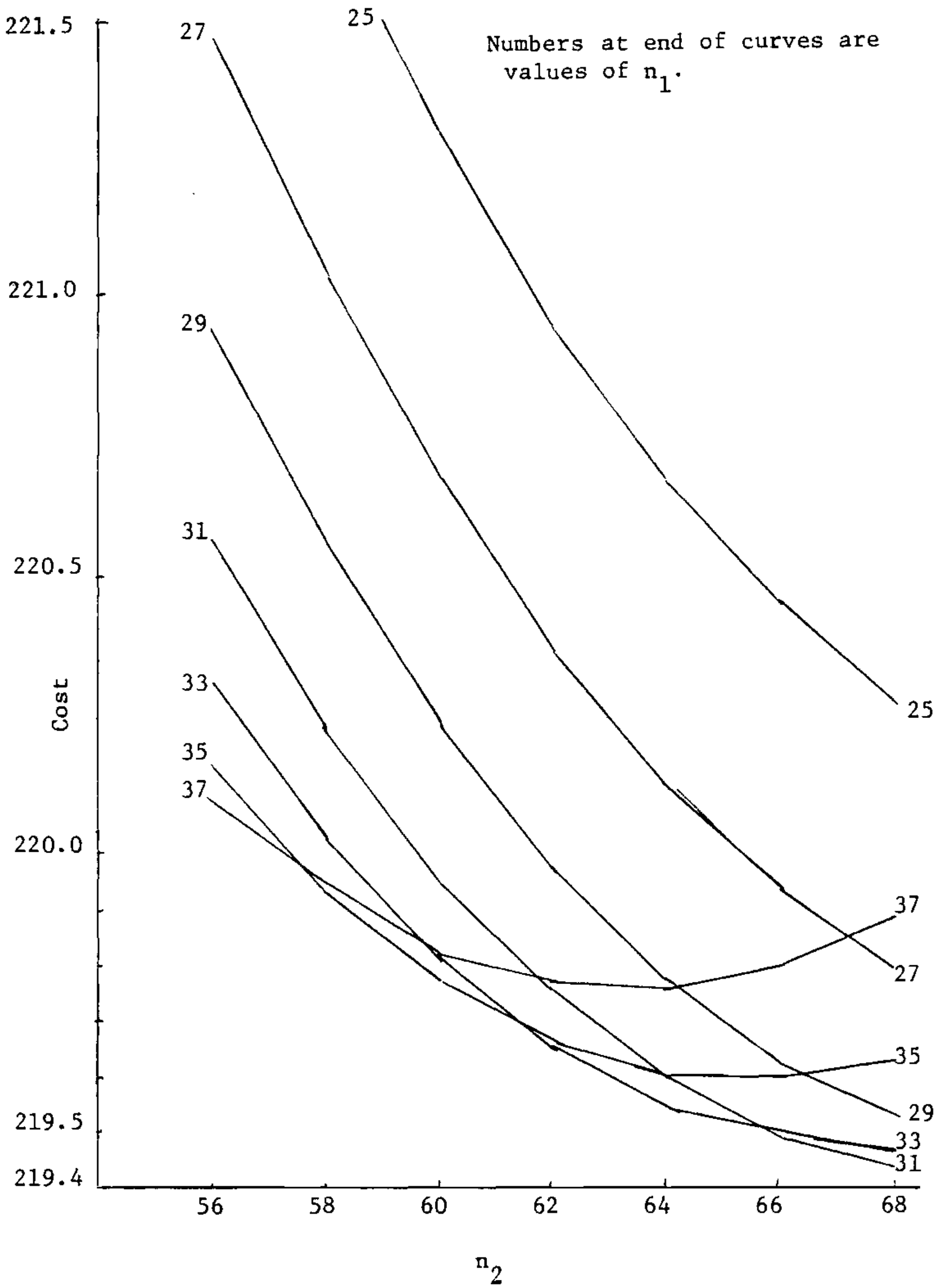


Figure 5. Expected Cost Curves for C_1=2 and C_2=12

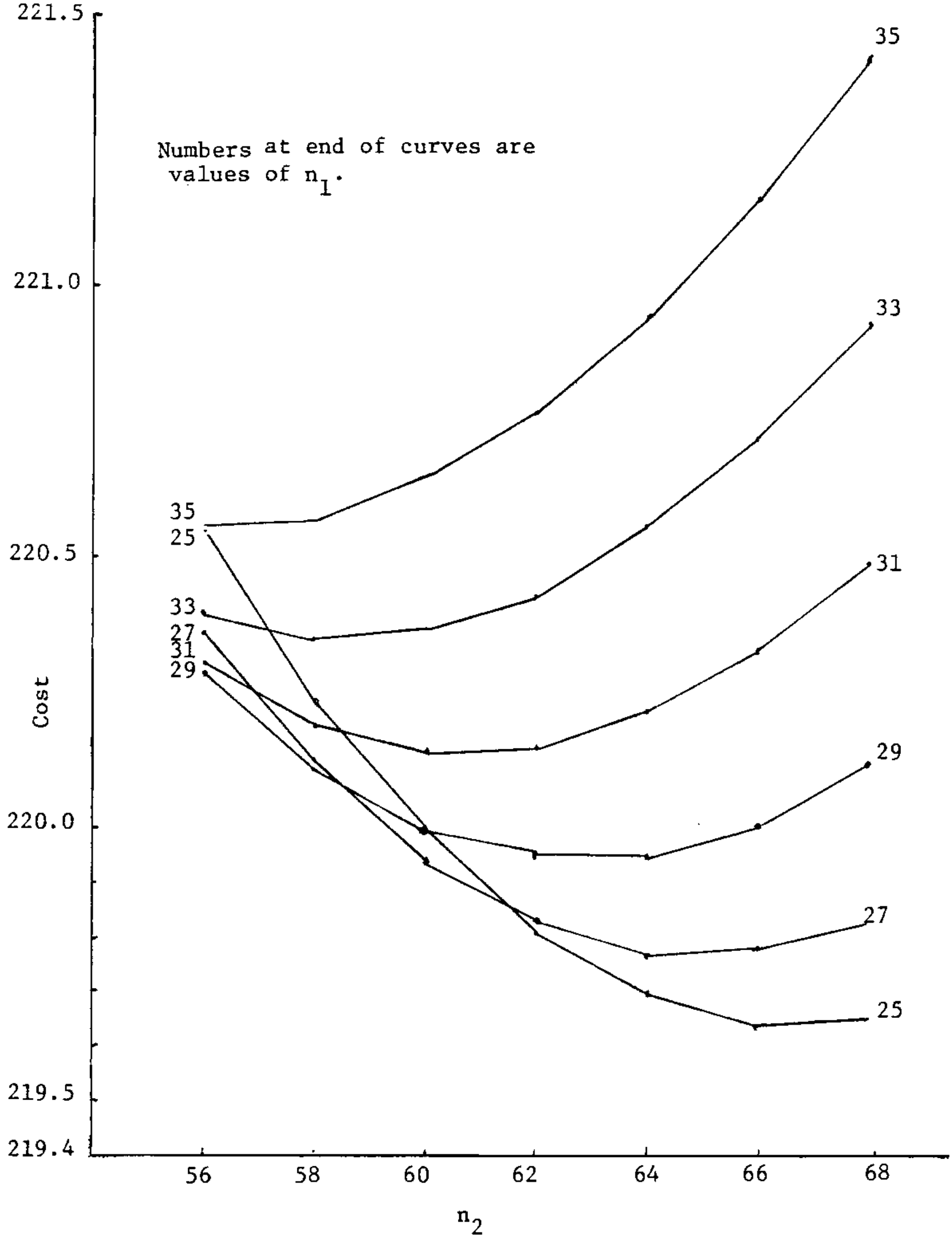


Figure 6. Expected Cost Curves for $C_1=1$ and $C_2=11$

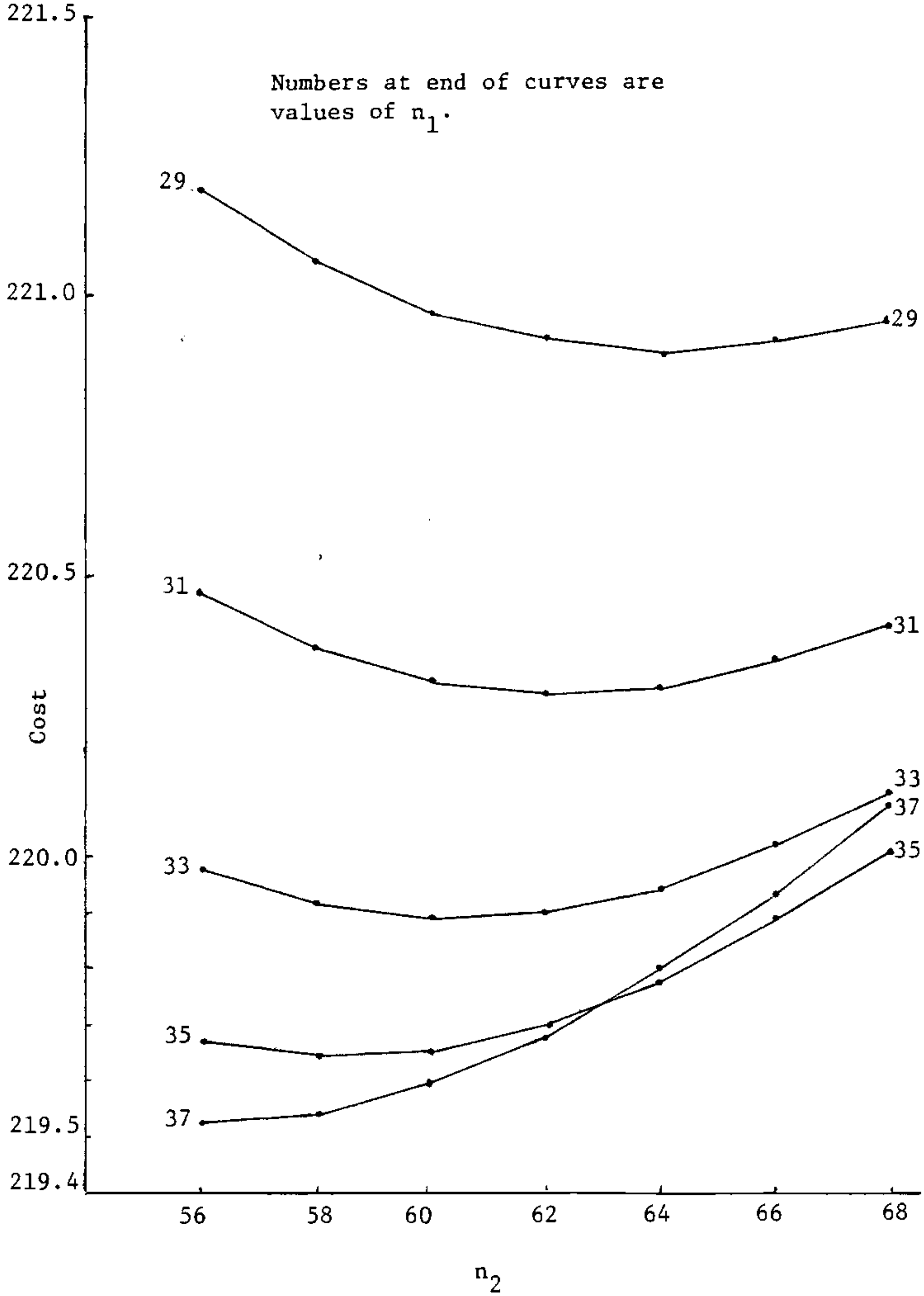


Figure 7. Expected Cost Curves for C_1=3 and C_2=11

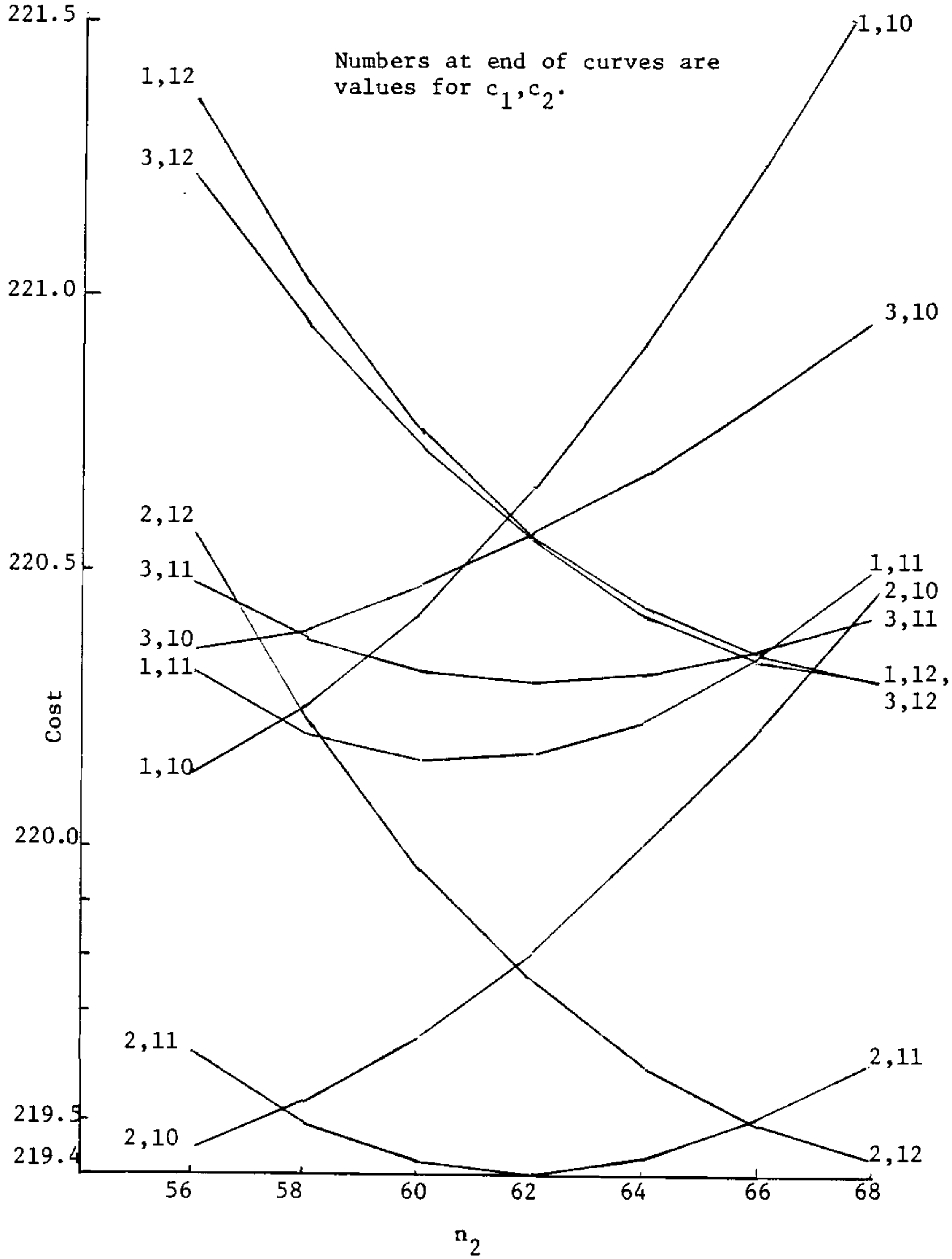


Figure 8. Expected Cost Curves for N_1=31

$$(n_1, c_1, n_2, c_2) = (31, 2, 62, 11)$$

with an expected cost of 219.39938. If three of the variables, n_1, n_2, c_1, c_2, are fixed and the fourth allowed to vary over its range of values, the individual points form a unimodal curve. Figure 4 shows a set of these curves with $c_1 = 2$ and $c_2 = 11$ for values of n_2 from 56 to 68 and n_1 from 25 to 37. Figure 9 shows an exaggerated section of these curves. If, in Figure 9, the scale for n_2 is increasing, then the value of n_1 for each curve decreases when moving from left to right. If we consider a set of these curves to be defined by fixed values of c_1 and c_2 and each individual curve to be defined by a fixed value of n_1, then the optimum point on each of these curves also form a unimodal curve when joined together. This curve of optimum points is defined by specific values of c_1 and c_2. Each point on it is an optimal value of n_1 for that value of n_2. This curve is relatively flat near the optimum but rises sharply away from it. Figure 4 shows curves surrounding the global optimum. Figures 3 and 5 show curves for (c_1,c_2) pairs (2,10) and (2,12). As expected the optimal value of n_2 increases and decreases as c_2 increases and decreases while the optimal value of n_1 is nearly unaffected. Figures 6 and 7 show curves for $c_2 = 11$, and c_1 equal to 1 and 3. Here, for $(c_1,c_2) = (1,11)$ the optimal value for n_2 is increased while that of n_1 is decreased. The inverse is true for $(c_1,c_2) = (3,11)$.

Table 1 indicates the sensitivity of the expected cost to errors in the final optimum point. The costs are dependent on the values given to the four cost coefficients, but the percent differences are only dependent on the ratios k_s/k_a and k_i/k_a. The first line gives the

C_1 and C_2 are fixed and each curve represented a different value of n_1. The value of n_1 for each curve decreases from left to right. The scale for n_2 increases from left to right.

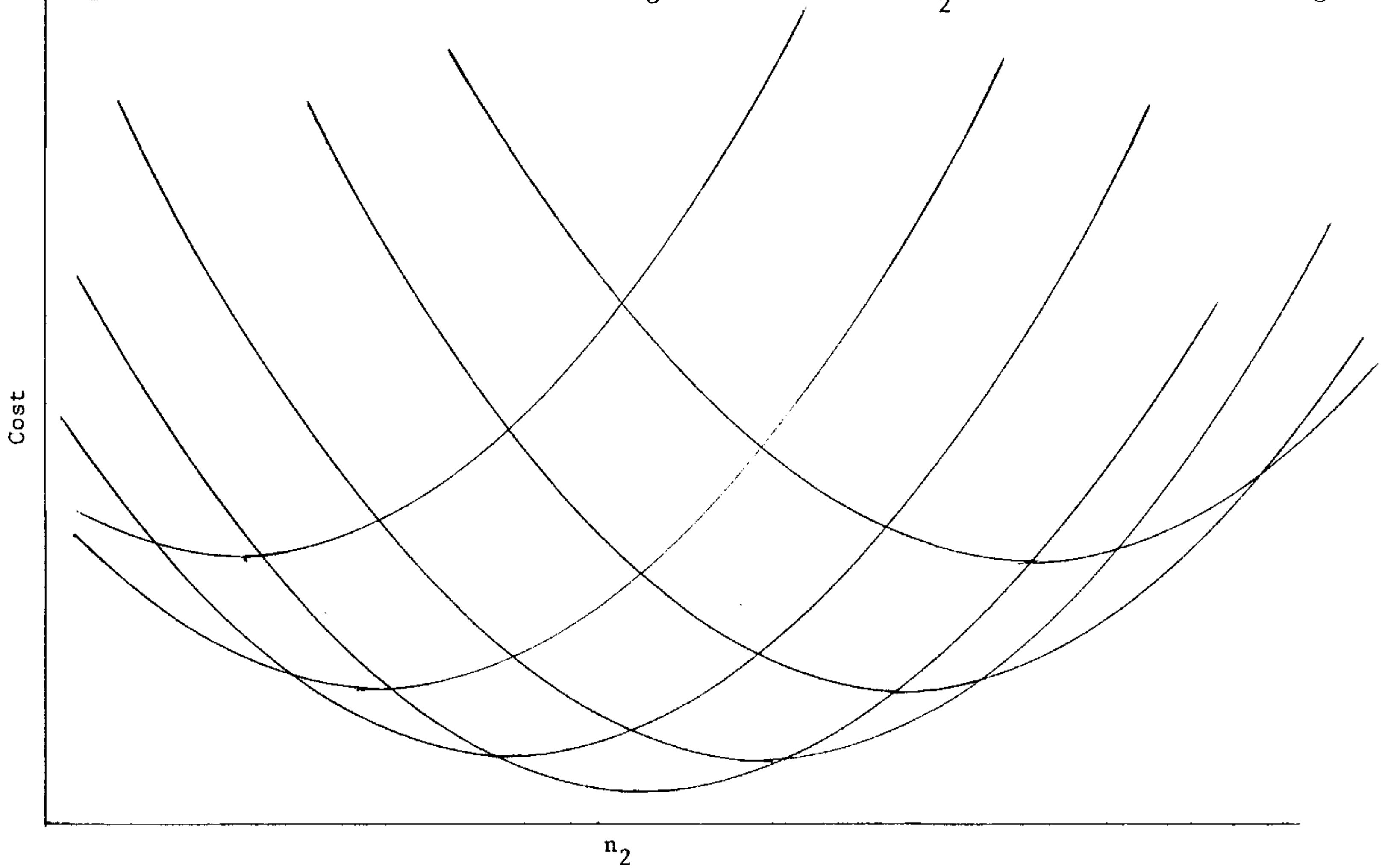


Figure 9. Expected Cost Curves

Table 1. Sensitivity of the Expected Cost to Errors in the Optimal Point

Sampling Plan n_1	c_1	n_2	c_2	Expected Cost	Difference Between Optimal and Exp. Cost
31	2	62	11	219.39938	-
23	1*	69	11	219.59324	0.18386
31	2	54	10*	219.42715	0.02777
39	3*	54	11	219.51268	0.11330
31	2	70	12*	219.42244	0.02306
30*	2	63	11	219.41024	0.01086
32*	2	61	11	219.41191	0.01253
31	2	61*	11	219.40300	0.00362
31	2	63*	11	219.40763	0.00825
25	2	62	11	220.17467	0.77529
27	2	62	11	219.71246	0.31308
29	2	62	11	219.46796	0.06858
30	2	62	11	219.41395	0.01457
32	2	62	11	219.42034	0.02096
33	2	62	11	219.47333	0.07395
35	2	62	11	219.66334	0.26396
37	2	62	11	219.94844	0.54906
31	2	56	11	219.61577	0.21639
31	2	58	11	219.48929	0.08991
31	2	60	11	219.41879	0.01941
31	2	61	11	219.40300	0.00362
31	2	63	11	219.40736	0.00825
31	2	64	11	219.42636	0.02698
31	2	66	11	219.49524	0.09586
31	2	68	11	219.60173	0.20235

optimal point and corresponding cost. The next section gives the optimal points if the variables marked with asterisks are forced to that value and the others are allowed to move to their optimum values. Table 1 shows that the surface is very flat in the vicinity of the optimum. The expected cost is least sensitive to changes in n_2, then n_1, then c_2 and most sensitive to c_1. The rest of Table 1 shows non-optimal points that result if first n_1 and then n_2 are varied, while the other, along with c_1 and c_2, if fixed.

3.3 Numerical Results

In all of the results presented the prior distribution used for p is the beta distribution presented in Appendix 2. The initial set of optimum sampling plans were found using a prior distribution with mean $\mu = .1$ and variance $v = .004$. The initial set of cost coefficients studied were

$$(k_i, k_a, k_s) = (.3, 2.5, .3) \text{ and } k_I = 1.00.$$

The reason for using $k_I = 1.00$ instead of zero is that this value enabled the computer program to sense the boundary more easily. The remaining optimum points in the first set are found by using plus and minus ten percent in each of the cost coefficients. The results are summarized in Table 2. Table 3 lists values of the ratios k_1/k_a and k_s/k_a. From Table 2 we can see that as k_a is increased while k_i and k_s are fixed, the sample sizes increase while the acceptance numbers decrease or remain the same. The optimum plans therefore accept fewer lots. If k_a and k_s are held

Table 2. Optimal Sampling Plans for Selected Sets of Cost Coefficients and a Beta Prior

$k_i = .27$

$k_s=$ / $k_a=$	.27	.30	.33
2.25	31,2,62,11 197.55944	27,2,84,15 204.64301	30,3,118,22 210.06051
2.50	37,2,59,10 208.38643	40,3,94,16 217.29608	36,3,130,22 224.24294
2.75	43,2,65,10 217.29933	46,3,95,15 227.85840	49,4,177,27 236.23612

$k_i = .30$

$k_s=$ / $k_a=$	.27	.30	.33
2.25	30,2,46,9 199.25451	26,2,63,12 206.3664	22,2,77,15 211.79430
2.50	35,2,43,8 210.31643	31,2,62,11 219.39938	28,2,77,14 226.57758
2.75	40,2,48,8 219.45321	36,2,59,10 230.33139	32,2,86,14 239.16815

$k_i = .33$

$k_s=$ / $k_a=$	.27	.30	.33
2.25	28,2,40,8 200.69802	25,2,49,10 207.7998	21,2,58,12 213.15867
2.50	34,2,35,7 211.93492	30,2,46,9 221.10921	26,2,64,12 228.31158
2.75	39,3,38,7 211.20672	35,2,42,8 232.26430	31,2,62,11 241.23932

Plans are presented as n_1,c_1,n_2,c_2. Prior Parameters a=1.15,b=18.35,u=.1 v=.004

Table 3. Ratios of k_i/k_a and k_s/k_a

k_a \ k_i or k_s	.27	.30	.33
2.25	.120	.133	.147
2.50	.108	.120	.132
2.75	.098	.109	.120

Table 4. Comparison of Sampling Plans by Operating Characteristic Curves

Prior Mean	Prior Variance	Cost Vector k_i	k_a	k_s	Optimal Sampling Plan	Value of p so that p_a= .95	.90	.10	.05
0.1	0.004	0.30	2.50	0.30	31,2,62,11	0.080	0.090	0.180	0.200
0.1	0.004	0.33	2.50	0.30	30,2,46,9	0.087	0.087	0.189	0.210
0.1	0.004	0.27	2.50	0.30	40,3,94,16	0.087	0.095	0.172	0.189
0.1	0.004	0.30	2.75	0.30	36,2,59,10	0.070	0.078	0.165	0.180
0.1	0.004	0.30	2.25	0.30	26,2,63,12	0.092	0.103	0.207	0.230
0.1	0.004	0.30	2.50	0.33	28,2,77,14	0.094	0.104	0.197	0.218
0.1	0.004	0.30	2.50	0.37	35,2,43,8	0.065	0.074	0.167	0.186
0.1	0.01	0.30	2.50	0.30	23,1,55,9	0.075	0.085	0.186	0.207
0.1	0.02	0.30	2.50	0.30	21,1,42,7	0.069	0.079	0.195	0.220
0.05	0.005	0.30	2.50	0.30	10,0,38,7	0.075	0.086	0.224	0.264

Table 5. Comparison of Optimal Plans for Different Beta Priors

$(k_i, k_a, k_s) = (.30, 2.50, .30)$ $k_I = 1.00$

Variance/Mean \ Mean	0.01	0.05	0.10
0.04	***	5,0,31,6 124.62236	31,2,62,11 219.39938
0.10	3,0,46,6 25.36731**	10,0,48,7 107.70313	23,1,55,9 183.16498
0.20	5,0,46,6 21.21707	11,0,42,6 84.96969	21,1,42,7 145.40971

$(k_i, k_a, k_s) = (.27, 2.75, .30)$ $k_I = 1.00$

Variance/Mean \ Mean	0.01	0.05	0.10
0.04	***	6,0,53,8 136.04039	46,3,95,15 227.85840
0.10	4,0,57,7 27.212181	12,0,61,8 113.28785	27,1,70,10 189.22116
0.20	4,0,57,7 22.1960	13,0,54,7 88.1573	17,0,52,7 14938616

$(k_i, k_a, k_s) = (.33, 2.25, .30)$ $k_I = 1.00$

Variance/Mean \ Mean	0.01	0.05	0.10
0.04	***	1,0,34,6 113.38509	25,2,49,10 207.7998
0.10	2,0,39,6 23.30479	8,0,36,6 101.15162	20,1,35,8 175.41711
0.20	4,0,41,6 20.071827	9,0,38,6 81.19038	13,0,55,6 140.18699

*** There exists no points with expected cost less than the cost of accepting without sampling

** Cost of acceptance without sampling < 25.36731 < cost of acceptance without sampling + 1.00

fixed and k_i increased, then fewer items are sampled. Both of these results are very intuitive. It is easier to see the effects of varying costs if the operating characteristic curves are studied.

Table 4 lists information about the operating characteristic curves for some of the optimum plans and some of the operating characteristic curves are shown in Figure 10. Increasing k_i holding k_a and k_s constant spreads the curve out while decreasing k_i makes the curve steeper. If k_s is increased with the other two costs constant, the curve becomes steeper. Decreasing k_s causes the curve to flatten out. Increasing k_a shifts the curve to the left; decreasing it shifts the curve to the right. For a decreased prior mean and the same ratio of mean to variance, the curve remains the same for a high probability of acceptance but includes a larger portion of the p axis for low probabilities of acceptance.

From Table 2 we can see that the expected cost is most sensitive to changes in the cost of accepting a defective and only slightly less sensitive to changes in cost of scrapping an item. A ten percent change in the cost of accepting a defective item causes about a 5 percent change in the expected cost. A ten percent change in the cost of scrapping an item causes about a 4 percent change in the expected cost, while a 10 percent change in the cost of inspecting an item causes only a 1 percent change in the expected cost.

Table 5 gives optimal sampling plans for the base point and two other points under other prior distributions of process quality. One interesting point noted from this table is that the expected cost decreases when the variance increases for a fixed mean. Once the prior distributions are examined, it is clear why this is true. Figure 11

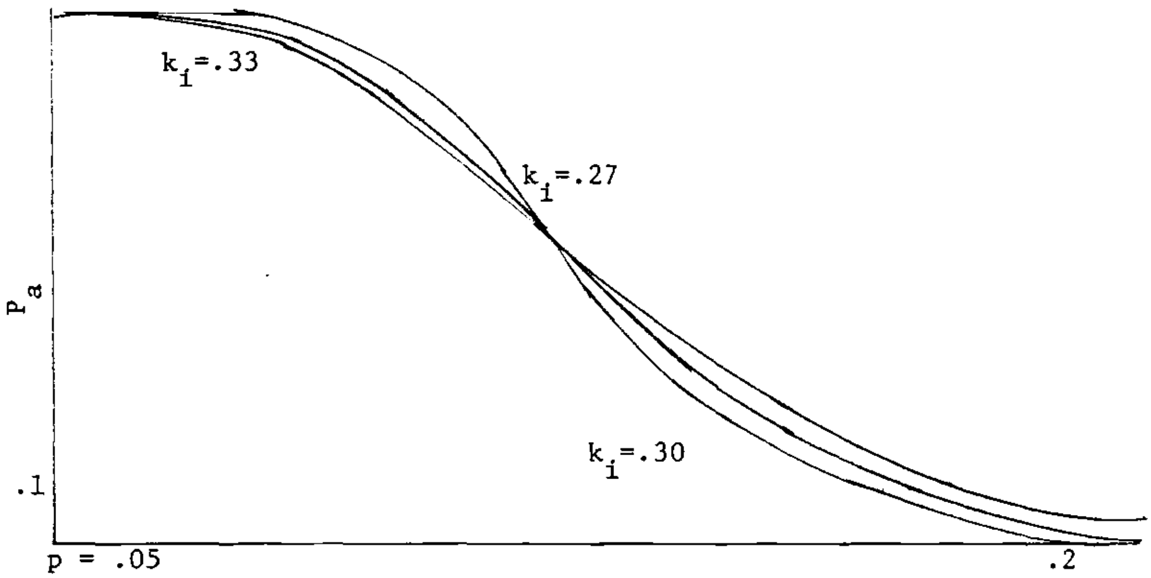


Figure 10a. Operating Characteristic Curves for Optimal Plans for Several k_i

31,2,62,11	k_a=2, k_s=.3
30,2,46,9	k_i=.30
40,3,94,16	k_i=.33
	k_i=.27

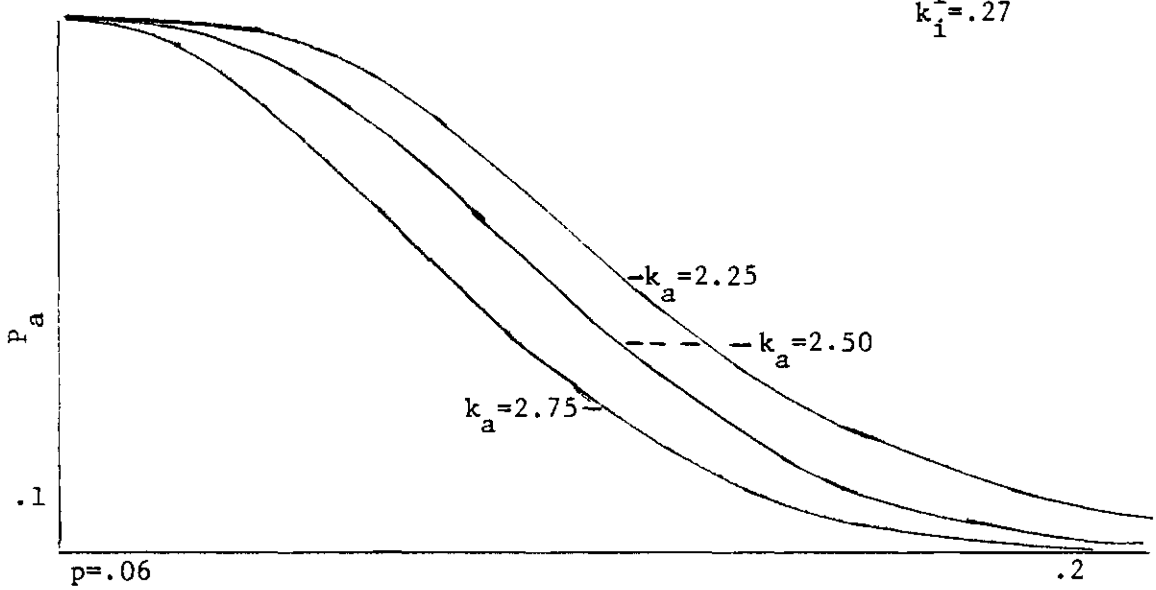
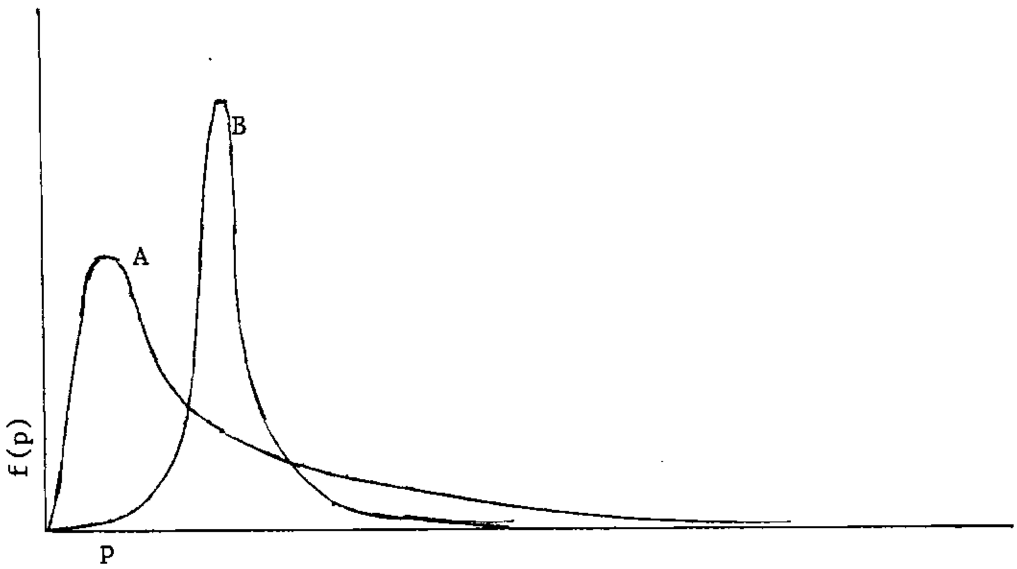


Figure 10b. Operating Characteristic Curves for Optimal Plans for Several k_a

31,2,62,11	k_i=.30, k_s=.30	36,2,59,10	k_a=2.75
	k_a=2.5	26,2,63,12	k_a=2.25



The means of curves A and B are the same, but the variance of A is greater than that of B.

Figure 11. Graph of Beta Priors

presents two beta distributions which have equal means. However, curve A has a higher variance than curve B. This means that it has more area towards the ends of the distribution. These are the areas where double sampling works most effectively and decisions can be made correctly more often on the first sample resulting in a lower cost. In the limit as the variance of the prior goes to zero, the expected cost goes either to the default cost of acceptance or rejection.

Table 6 gives the sensitivity of the expected cost to small changes in the prior. A ten percent change in the prior mean holding variance/mean constant produces about a 2 percent change in the expected cost and does not change the optimal sampling plan. A 20 percent change in the variance with a fixed mean produces about the same change in the expected cost and also leaves the optimal sampling plan unchanged.

Table 7 shows an extension of Table 2. If k_a is increased while k_s is held equal to k_i, both the expected cost and number of items sampled increase. If k_a is decreased then the opposite happens until the default cost of acceptance without sampling is reached. It is interesting to examine the costs to determine when it is less costly to accept a lot without sampling. If k_i is held equal to k_s, then Table 8 gives values of k_a/k_i such that the optimal point lies in the interior for all values of k_a/k_i greater than the ratio. For a fixed mean, the minimum ratio of k_a/k_i required for sampling increases as the variance decreases. If the ratio of variance to mean is held constant, the ratio of k_a/k_i required for sampling increases as the mean decreases.

It is instructive to compare the economically optimum double sampling plans to economically optimum single sampling plans and sampling

Table 6. Sensitivity to Small Changes in the Prior

Prior Mean, Variance

Optimal Plan

Expected Cost

$(k_I, k_i, k_a, k_s) = (1.00, .30, 2.50, 30)$

	$\mu = .05,\ v = .004$ 10,0,48,7 113.39455	
$\mu = .049,\ v = .005$ 10,0,48,7 105.34261	$\mu = .05,\ v = .005$ 10,0,48,7 107.70313	$\mu = .051,\ v = .005$ 10,0,48,7 110.06943
	$\mu = .05,\ v = .006$ 10,0,48,7 105.39207	

Table 7. Optimal Sampling Plans for Extended Ratios for k_i/k_a and k_s/ka

Prior Parameters
$\mu = .1$, $v = .004$

Costs			Ratios		Optimal Plan				Expected Cost
k_i	k_a	k_s	k_i/k_a	k_s/k_a	n_1	c_1	n_2	c_2	
0.27	3.00	0.27	0.09	0.09	48	2	61	9	224.66437
0.27	3.25	0.27	0.083	0.083	54	2	65	9	230.80455
0.27	3.50	0.27	0.077	0.077	59	2	71	9	235.9696
0.27	4.00	0.27	0.068	0.068	69	2	68	8	244.03064
0.27	5.00	0.27	0.054	0.054	71	1	85	7	254.44552
0.33	2.00	0.33	0.165	0.165	16	2	58	13	195.48302
0.33	1.75	0.33	0.189	0.189	7	1	44	11	174.8016
0.33	1.50	0.33**							
0.33	1.25	0.33**							
0.33	1.00	0.33**							
0.36	2.75	0.24	0.131	0.087	32	1	32	5	209.15156
0.39	2.75	0.21	0.142	0.076	25	0	17	2	194.11876

** The expected cost is greater than that of accepting without sampling.

Table 8. Rations of k_a/k_i with k_i=k_s so that the Optimal Expected Cost is Less than the Expected Cost of Acceptance without Sampling

Prior Parameters		Ratio of
Mean	Variance	k_a/k_i
0.1	0.004	5.877
0.1	0.010	3.707
0.1	0.020	2.410
0.05	0.005	4.803
0.01	0.001	8.657

plans from the Military Standards given in Duncan (4). Table 9 gives a table of optimal single sampling plans, their expected cost, the corresponding cost from the optimal double sampling plan, and the percent savings attained by using double sampling. The optimal single sampling plans were derived by forcing n_2 to be zero in the original model. The savings range from about one percent for a case of relatively large prior mean and small coefficient of variation to about 7 percent for the relatively small prior mean and large coefficient of variation case.

Let the prior with mean $\mu = .1$ and variance $v = .004$, and the cost vector

$$(k_I, k_i, k_a, k_s) = (1., .3, 2.5, .3),$$

be used to find the expected cost of the plans. From Military Standard 105D, the single sampling plan given for N = 1000 with an AQL of 0.1 under normal insepction is n = 80 and c = 14. The expected cost of this plan is about 230. The double sampling plan from Military Standard 105D given under the same conditions is

$$(n_1, c_1, n_2, c_2) = (50, 7, 50, 18).\ r_1 = 11,\ r_2 = 19$$

The cost for this plan is about 268. This comparison cannot be made directly because the Military Standard Plans use curtailment on the second sample and do not require rejection numbers to be equal.

Table 9. Comparison of Optimal Single Sampling Plans and Optimum Single Sampling Plans

Prior		Costs			Sampling Plan		Expected Cost	Optimal Double Cost	Percent Diff.
μ	v	k_i	k_a	k_s	n	c			
0.1	0.004	0.30	2.50	0.30	59	7	222.08544	219.39938	1.22
0.1	0.004	0.33	2.50	0.30	51	6	223.644447	221.10921	1.15
0.1	0.004	0.27	2.50	0.30	67	8	220.24410	217.29608	1.36
0.1	0.004	0.30	2.75	0.30	67	7	232.86303	230.33139	1.10
0.1	0.004	0.30	2.25	0.30	44	6	209.15113	206.36640	1.35
0.1	0.004	0.30	2.50	0.33	52	7	229.85326	226.57758	1.44
0.1	0.004	0.30	2.50	0.27	58	6	212.47219	210.31643	1.02
0.1	0.01	0.30	2.50	0.30	44	5	186.26519	183.16498	1.72
0.1	0.02	0.30	2.50	0.30	29	3	148.51286	145.40971	2.14
0.05	0.005	0.30	2.50	0.30	24	3	111.10403	107.70313	3.16
0.01	0.001	0.30	2.50	0.30	********				
0.01	0.002	0.30	2.50	0.30	8	1	22.747297	21.21707	7.21

******** Expected cost of all points greater than default cost of acceptance.

3.4 The Scrapping Model with Costs Considered as Random Variables

There are some cases where the costs cannot be considered as fixed quantities. The cost of scrapping may be dependent on a salvage value that changes over time. If a batch of items is to be used in more than one way, then the cost of accepting a defective will depend on the distribution of usage. If the costs can be considered as independent random variables whose distribution does not depend on the lot fraction defective, then the expected cost can be written as

$$E(\text{cost}) = \iiiint\limits_{K_I K_i K_a K_s} E(\text{cost}) \mid k_I, k_i, k_a, k_s)\; f_{k_I}(k_I)\, dk_I\; f_{k_i}(k_i)\, dk_i\; *$$

$$f_{k_a}(k_a)\, dk_a\; f_{k_s}(k_s)\, dk_s.$$

K_I, K_i, K_a, and K_s are the regions where k_I, k_i, k_a, and k_s take on values respectively. Also, f_{k_I}, f_{k_i}, f_{k_a}, and f_{k_s} are the distributions of the cost coefficients. The $E(\text{cost} \mid k_I, k_i, k_a, k_s)$ is the right hand side of equation (1). If any of the distributions are discrete then the corresponding integral signs are replaced by summations. Since all of the costs and p are assumed independent, the order of integration can be interchanged. The integrals corresponding to the costs can be brought inside of the $E(\text{cost}) \mid k_I, k_i, k_a, k_s)$ integral. These integrals can now be evaluated easily. The result then is the original expected cost equation with the cost vector replaced by the vector of the means of the distributions of the costs.

CHAPTER IV

OTHER VARIATIONS OF THE BASIC COST MODEL

4.1 The Inspection Model

An alternative to scrapping rejected lots is one hundred percent screening of the items remaining after a lot has been rejected. If this is the case, then the cost of accepting a lot on either sample is unchanged.

When a lot is rejected, the cost of disposing the remainder of the lot must be replaced by the cost of inspecting the remainder and scrapping only the defectives. The cost of a lot which is rejected on the first sample is then

$$C_{r1} = k_I + k_i n_1 + k_s d_1 + k_i(N - n_1) + k_s(X - d_1).$$

The cost of rejecting a lot on the second sample is

$$C_{r2} = k_I + k_i(n_1 + n_2) + k_s(d_1 + d_2) + k_i(N - n_1 - n_2) + k_s(X - d_1 - d_2).$$

Applying the same approximation as in the basic model C_{r1} and C_{r2} become

$$C_{r1}(p) = k_I + k_i n_1 + k_s n_i p + k_i(N - n_1) + k_s(N - n_1)p,$$

$$C_{r2}(p) = k_I + k_i(n_1 + n_2) + k_s(n_1 + n_2)p + k_i(N - n_1 - n_2) +$$
$$k_s(N - n_1 - n_2)p.$$

The cost as a function of p is

$$C(p) = C_{a1}(p)P_{a1}(p) + C_{a2}(p)P_{a2}(p) + C_{r1}(p)P_{r1}(p) + C_{r2}(p)P_{r2}(p).$$

From here the development proceeds as in the basic model

$$E(\text{cost}) = \int_0^1 E(\text{cost}|p)f(p)dp$$
$$= \int_0^1 [P_{r1}(p)(k_I + k_in_1 + k_sn_1p + (N-n_1)k_i + (N-n_1)k_sp) +$$
$$P_{r2}(p)(k_I + k_i(n_1+n_2) + k_s(n_1+n_2)p + (N-n_1-n_2)k_i +$$
$$(N-n_1-n_2)k_sp) +$$
$$P_{a1}(p)(k_I + k_in_1 + k_a(N-n_1)p + k_sn_1p) +$$
$$P_{a2}(p)(k_I + k_i(n_1+n_2) + k_a(N-n_1-n_2)p + k_s(n_1+n_2)p]f(p)dp$$

Table 10 gives some optimal sampling plans derived from the inspection model. The optimal expected cost is most sensitive to changes in the cost of accepting a defective item and least sensitive to the cost of scrapping an item. A 10 percent change in the cost of accepting a defective item results in about a 6.5 percent change in expected cost. A ten percent change in k_i or k_s results in a 3 percent or .5 percent change in expected cost, respectively.

Table 11 indicates the sensitivity of the inspection model to small changes in the prior. The inspection model is somewhat more sensitive to changes in the prior than the scrapping model. A ten percent change in mean results in a 2 percent change in expected cost and a 20 percent change in variance results in a 4 percent change in cost.

For a given set of cost coefficients these two models can be used to determine the optimum disposition of a lot. That is, the decision to either screen or scrap rejected lots can be made on the basis of lowest expected cost rather than as an arbitrary decision.

4.2 The Restricted Model

If we require sample sizes to satisfy a relationship such as $n_2 = 2n_1$ we obtain a higher expected cost. However, we may gain some efficiency in administration. Having the plan parameters related in a simple fashion reduced the liklihood of the inspector making an error in the sampling procedure. Table 12 gives some examples of optimal sampling plans resulting from requiring both the sample sizes and acceptance numbers to satisfy certain relationships. The plans from the Military Standard 105D require n_1 and n_2 to be equal, but they do not require any relationships between acceptance numbers. The plans here were found by requiring the specified relationship to hold during the search procedure. For a cost vector of

Table 10. Optimal Points from the Inspection Model

Prior Parameters		Costs			Sampling Plan	Expected Cost
μ	v	k_i	k_a	k_s	n_1,c_1,n_2,c_2	
0.05	0.005	0.30	2.50	0.30	14,1,80,13	113.64701
0.05	0.005	0.27	2.50	0.30	17,1,80,12	110.18388
0.05	0.005	0.33	2.50	0.30	12,1,73,13	116.51853
0.05	0.005	0.30	2.25	0.30	12,1,70,13	105.72691
0.05	0.005	0.30	2.75	0.30	17,1,81,12	120.79892
0.05	0.005	0.30	2.50	0.27	15,1,73,12	113.04800
0.05	0.005	0.30	2.50	0.33	14,1,79,13	114.24791
0.05	0.002	0.030	2.50	0.30	2,0,39,7	125.7450
0.05	0.010	0.30	2.50	0.30	9,0,58,9	94.058488
0.01	0.001	0.30	2.50	0.30	***	
0.10	0.010	0.30	2.50	0.30	20,1,77,13	200.36394

*** No point has a cost less than accepting without sampling.

Table 11. Sensitivity of the Inspection Model to Small Changes in the Prior

$(k_i,k_a,k_s) = (.30,2.50,.30) \qquad k_I = 1.00$

Prior Parameters		Optimal Sampling Plan	Expected Cost
μ	v	n_1,c_1,n_2,c_2	
0.051	0.005	14,1,80,13	116.06923
0.049	0.005	14,1,80,13	111.23377
0.050	0.006	15,1,80,13	109.33044
0.050	0.004	13,1,73,12	118.13061
0.050	0.005	14,1,80,13	113.64701

Table 12. Table of Optimal Plans with Constraints on the Relationships Between Sample Sizes and Rejection Numbers

Prior Parameters Mean = .1
Variance = .004

$(k_I,k_i,k_a,k_s) = (1.,.3,2.5,.3)$ Optimal Expected Cost 219.39938

n_2/n_1	c_2/c_1	n_1	c_1	n_2	c_2	Expected Cost
2	5	29	2	58	10	219.50570
2	4	35	3	70	12	219.69492
2	3	38	4	76	12	220.83688
2	6	33	2	66	12	219.48799
2	7	37	2	74	14	220.09671
1	3	39	3	39	9	219.71703
1	4	34	2	34	8	219.80758
1	2	47	5	47	10	220.69228

$(k_I,k_i,k_a,k_s) = (1.,.3,2.25,.3)$ Optimal Expected Cost 206.3664

n_2/n_1	c_2/c_1	n_1	c_1	n_2	c_2	Expected Cost
2	7	18	1	36	7	206.98228
2	6	29	2	58	12	206.53202
2	5	25	2	50	10	206.43778
2	4	30	3	60	12	206.60484

$(k_I,k_i,k_a,k_s) = (1.,.3,2.75,.3)$ Optimal Expected Cost 230.33139

n_2/n_I	c_2/c_1	n_1	c_1	n_2	c_2	Expected Cost
2	7	42	2	84	14	231.00561
2	6	37	2	74	12	230.42199
2	5	35	2	70	10	230.51574
2	4	37	3	74	12	230.76044

$$(k_I, k_i, k_a, k_s) = (1., .3, 2.5, .3)$$

the optimum value for the pair $(n_2/n_1, c_2/c_1)$ over the set of all integers is (2,6). In general, the loss in expected cost obtained by using the optimum restricted plan instead of the overall optimum plan is less than one percent. As k_a is increased while k_i and k_s are held constant, the optimal value of the c_2/c_1 increases while n_2/n_1 remains near constant.

4.3 Curtailment

Often when the number of defectives found during testing reaches the rejection number, the rest of the sample is rejected without testing. This process is called curtailment. If we curtail only on the second sample, the value of n_2 used in the cost model must be adjusted. The probabilities of lot acceptance and rejection are based on the original sample sizes n_1 and n_2 and are therefore independent of curtailment. From Duncan (4) we have

$$E(n_2|p) = \sum_{k=c_1+1}^{c_2} P(n_1:k)(n_2P''(n_2:C_2-k) + ((C_2-k+1)/P)P'(n_2+1:C_2-k+2))$$

with

$P(n:k)$ = Probability of exactly k defective items out of n

$P'(n:k)$ = Probability of k or more defective items out of n

$P''(n:k)$ = Probability of k or less defective items out of n.

The $E(n_2|p)$ was evaluated at each end of the intervals of integration for p and the largest value was used as an approximation for n_2 over that interval in $E(\text{cost}|p)$. When curtailment is used in the scrapping model, the main change occurs on the second sample. For instance, the optimal point from the basic model for the cases

$$(k_I, k_i, k_a, k_s) = (1., .3, 2.5, .3)$$

is

$$(n_1, c_1, n_2, c_2) = (31, 2, 62, 12) \text{ with } E(\text{cost}) = 43.39938$$

as compared to

$$(n_1, c_1, n_2, c_2) = (30, 2, 115, 17) \text{ with } E(\text{cost}) = 217.58505$$

for the optimum with curtailment. The second sample size for the optimal plan increases because the expected number of items sampled is smaller under curtailment. The average expected savings over the uncurtailed model was about one percent.

CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

The main purpose of this research was to derive a model for the design of double sampling plans based on an economic criterion. Secondary objectives were to develop methods allowing evaluation of the cost function in a reasonable amount of computer time and to find a search procedure that would find the global minimum of the associated cost surface. All of these were accomplished. The major assumption used in accomplishing these purposes was that the number of defectives in a lot could be represented as a step function of lot quality. The rejection numbers r_1 and r_2 were assumed to be $c_2 + 1$. The cost of sampling one item k_i was also assumed to be the same for both the first and second samples.

The surface around the optimum is relatively flat, but becomes steeper as we move away from the optimum. In general, as k_a is increased while k_s and k_i are held constant the optimum plan becomes more selective. The expected cost behaves linearly with respect to the fixed cost of sampling k_I and is uniquely determined by the two ratios k_1/k_a and k_s/k_a. The optimum sampling plan remains unchanged under small changes in the prior mean and variance. In general, double sampling is not much more efficient than single sampling in terms of total expected cost. This agrees with the conclusions of Pfanzagl (15). However, in some cases a 7 percent reduction in cost can be achieved independent of the costs involved. If the costs associated with sampling were very high, then the use of double sampling could result in a substantial savings.

5.2 Recommendations

The costs presented here and in most other research on economic acceptance sampling, are constants. In reality, the costs might be a function of the sample sizes, based on either the total number of items sampled or the maximum number of items on test at any point in time. For instance, the inspection procedure might require a test stand that would limit the number of items tested at one time. The cost of inspection might be a step function, say

$$k_I = K[M/\max(n_1,n_2)].$$

Here K is a constant, M the maximum number of items that can be tested at one time and [] the greatest integer contained within the function. Another variation would be to allow k_i to vary from the first sample to the second. In reality the cost of inspecting one unit would usually be higher on the second sample. This would bring down the number of items in the second sample.

Another extension of the model would be to include inspector error and bias. A distribution would be added to model the probability that an inspector would misclassify an item. The treatment might be similar to that of Mei, Case, and Schmidt (14).

The choice of prior is important to the accuracy of the results. A type of prior that has not seen much analysis is the mixed beta:

$$f(p) = af_1(p) + bf_2(p),\ a+b=1,\ f_1 \text{ and } f_2 \text{ beta distributions.}$$

This distribution is useful for modeling any process that is bimodal, or modeling output from a mixture of two production facilities with different fraction defectives. Some analysis should be conducted to determine model behavior under different prior distributions of lot quality.

Hald (9) has done some work on optimum double by variables sampling using average sample number as an optimality criterion. Bennett, Schmidt, and Case (17) analyzed single sampling by variables using fully economic criterion. However no fully economic treatment for double variables sampling by variables has been performed.

A last extension concerns changes in the basic structure of the sampling procedure. When curtailment on the second sample was tried, the first sample size was left almost unchanged. It would be interesting to see the results of applying curtailment to the first sample. Another assumption made was that the rejection number for the first sample and that for the total would be the same. If they were not forced to be equal, there would be an improvement in overall efficiency; and it would be interesting to determine its significance.

APPENDIX I

OPTIMIZATION METHODS

A.1 Pattern Search in n_1 and n_2 with c_1 and c_2 Fixed

This appendix gives an example of the pattern search used to optimize cost. Part A illustrates the search in n_1 and n_2 while part B shows the search in c_1 and c_2.

Cost Parameters

$k_I = 0.00$

$k_i = .30$

$k_a = 2.50$

$k_s = .30$

Rejection Numbers

$c_1 = 2$

$c_2 = 11$

Beta Prior

$a = 1.15$

$b = 18.35$

$\mu = .1$

$v = .004$

Step Size

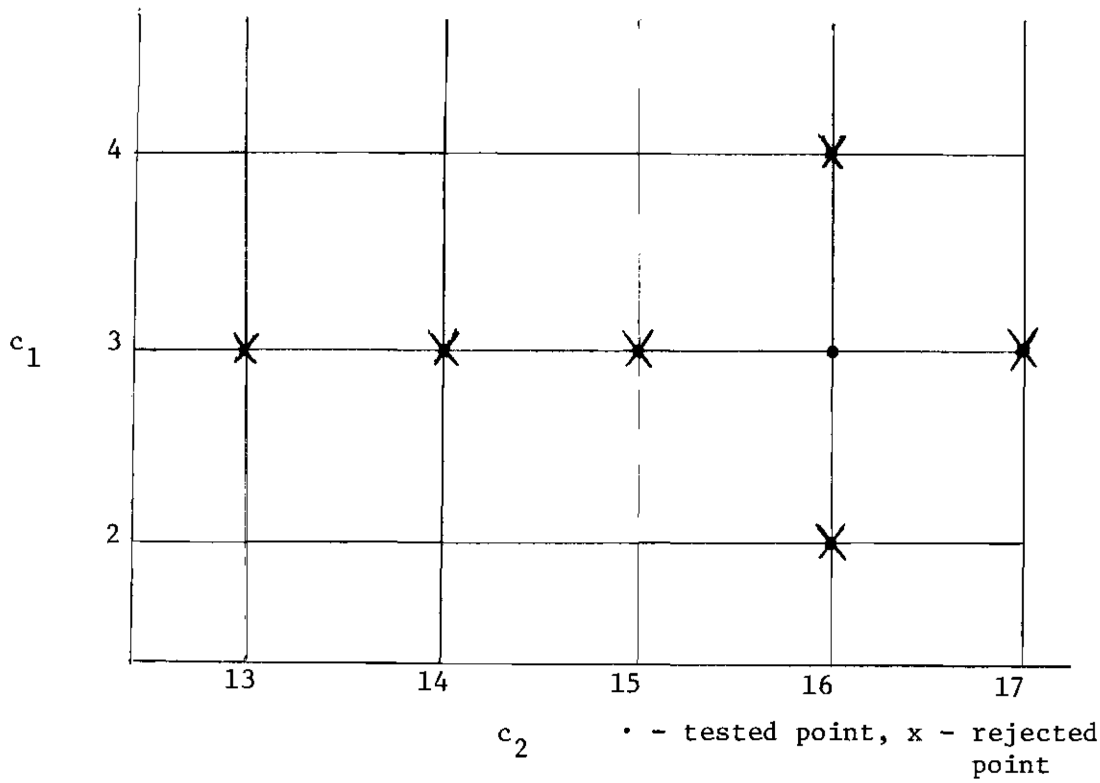
$d = 1$

Step Number	n_1	n_2	Function Value	Comments	
1	29	65	218.45619	Starting Point	
2	30	65	218.43606	(n_1+d_1, n_2)	Steps 2-4 are
3	30	66	218.47821	(n_1+d_1, n_2+d)	exploratory moves from
4	30	64	218.41780	(n_1+d, n_2-d)	the base point
5	31	63	218.40763	(n_1+2d, n_2-2d)	Pattern move from last base point
6	32	63	218.44010	(n_1+d, n_2)	Exploratory moves
7	30	63	218.41024	(n_1-d, n_2)	from the new base
8	31	64	218.42636	(n_1, n_2+d)	points
9	31	62	218.39938	(n_1, n_2-d)	
10	31	61	218.40300		Pattern move
11	32	61	218.41191	The combination of pattern move and	
12	30	61	218.42950	exploratory moves has not produced any	
13	31	62	218.39938	functional improvement so the pattern move has failed and the algorithm returns to the previous base prior (31,62)	
14	32	62	218.42034	Exploratory moves from the last base	
15	30	62	218.41395	point also fail so the algorithm is	
16	31	61	218.40300	finished	
17	31	63	218.40736		

A.2 Search Procedure in c_1 and c_2

Search Values		Optimal Values		Expected Cost
c_1	c_2	n_1	n_2	
3	13	40	70	217.37248
3	14	40	79	217.31961
3	15	40	86	217.29410
3	16	40	94	217.29068
3	17	40	102	217.30590
4	16	48	87	217.44219
2	16	32	102	217.33452

The initial search procedure is along the plus direction in c_2. When a decrease in function value is found the plus and minus directions in c_1 are tested. When these fail, the optimum is found.



APPENDIX II

BETA PRIOR DISTRIBUTION

The form of the Beta Distribution used in this study is

$$F(p) = (G(a+b+2)/(G(a+1)G(b+1)))p^{a}(1-p)^{b}, \; 0 > p > 1, \; a,b > -1.$$

The mean is $\mu = (a+1)/(a+b+2)$

and the variance is $v = (a+1)(b+1)/((a+b+2)^{2}(a+b+3))$.

The function G is the gamma function

$$G(n) = e^{-x}x^{n-1}dx.$$

The parameters a and b can also be written in terms of μ and v:

$$a = \frac{\mu^{2}(1-\mu) - \mu v}{v} - 1$$

$$b = \frac{(a+1)(1-\mu)}{\mu} - 1$$

These forms are useful in finding a prior with specified mean and variance.

BIBLIOGRAPHY

1. Ailor, R. B.; Schmidt, J. W.; Bennett, G. K. (1975). The Design of Economic Acceptance Sampling Plans for a Mixture of Variables and Attributes, AIIE Transactions.

2. Anscombe, F. J. (1951). The Cost of Inspection. Statistical Methods in Industrial Production, R. S. S., London.

3. Champernowne, D. G. (1953). The Economics of Sequential Sampling Procedures for Defectives. Applied Statistics, Vol. 2, 118-130.

4. Duncan, A. J. (1974). Quality Control and Industrial Statistics. Richard D. Irwin, Inc., Homewood, Illinois.

5. Guthrie, D.; Johns, M. V. (195). Bayes Acceptance Sampling Procedures for Large Lots. Ann. Math. Statist., Vol. 30, 896-925.

6. Hald, A. (1960). The Compound Hypergeometric Distribution and a System of Single Sampling Inspection Plans Based on Prior Distributions and Costs. Technometrics, Vol. 2, 275-340.

7. Hald, A. (1965). Bayesian Single Sampling Attribute Plans for Discrete Prior Distributions, Mat. Fys. Skr. Dan. Vid. Selsk., Vol. 3, 1-88.

8. Hald, A. (1968). Bayesian Single Sampling Attribute Plans for Continuous Prior Distributions. Technometrics, Vol. 10, 667-683.

9. Hald, A. (1975). Optimum Double Sampling Tests of Given Strength I. The Normal Distribution. J. Amer. Statist. Ass., Vol. 70, 457-462.

10. Hald, A.; Keiding, N. (1969). Asymptotic Properties of Bayesian Decision Rules for Two Terminal Decisions in Multiple Sampling. J. R. Statist. Soc., B, Vol. 34, 55-74.

11. Hamaker, H. C. (1951). Economic Principles in Industrial Sampling Procedures. Bull. Int. Statist. Inst., Vol. 33, 105.

12. Hamaker, H. C. (1958), Some Basic Principles of Sampling Inspection by Attributes, Appl. Statist., Vol. 7, 149-159.

13. Horsnell, G. (1957), Economic Acceptance Sampling Plans, J. R. Statist. Soc., A, Vol. 120, 148-201.

14. Mei, W.; Case, K. E.; Schmidt, J. W. (1975). The Effects of and

Compensation for Bias and Imprecision on Variables Acceptance Sampling Plans. AIIE Transactions.

15. Pfanzagl, J. (1963). Sampling Distribution Based on Prior Distributions and Costs. Technometrics, Vol. 5, 47-61.

16. Schmidt, J. W.; Bennett, G. K. (1972). Economic Multiattribute Acceptance Sampling. AIIE Trans., Vol. 4, 194-199.

17. Schmidt, J. W.; Case, K. E.; Bennett, G. K. (1974). The Choice of Variables Sampling Plans Using Cost Effective Criteria, AIIE Trans., Vol. 6, 178-184.

18. Schüler, W. (1967). Multistage Sampling Procedures Based on Prior Distributions and Costs. Ann. Math. Stat., Vol. 38, 464-470.

19. Sittig, J. (1951). The Economic Choice of Sampling System in Acceptance Sampling. Bull. Inst. Statist. Inst., Vol. 33, 51-84.

20. Wetherill, G. B.; Chiu, W. K. (1975). A Review of Acceptance Sampling Schemes with Emphasis on the Economic Aspect. Int. Stat. Rev., Vol. 43, 191-210

21. United States Department of Defense (1963) Military Standard Sampling Procedures and Tables for Inspection by Attributes (MIL-STD-105D). Washington, D.C. Government Printing Office.