INFLUENCE OF STRATUM CONSTANTS ON A TWO STAGE BAYESIAN SAMPLING SCHEME FOR A FINITE, STRATIFIED AND DICHOTOMOUS POPULATION.

by

ALFREDO GABRIEL PARDO

Ingeniero Mecánico, Universidad Central de Venezuela

Caracas-Venezuela, 1966

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Industrial Engineering

KANSAS STATE UNIVERSITY

Manhattan, Kansas

1972

Approved by:

Doris Grosh

Major Professor

With love,
To
My wife Alicia, our children Diana and Alfredo Andres, my father Alfredo Pardo S., my mother in law Mercedes Perez Arevalo, my brothers and sisters.

In Memoriam of my mother Carmen de Pardo and my father in law F. M. Perez Arevalo.

ACKNOWLEDGEMENT

I want to express my sincere appreciation to my major professor Dr. Doris L. Grosh for her unvaluable advice and guidance through all the phases of the present work.

I also wish to express recognition to Dr. L. E. Grosh Jr. and Mr. John Devore for their help in the computer programming part as well as to Ms. Marie Jirak for their fine job on the typing of the manuscript.

Special thanks and everlasting gratitude to my family for their patience, moral and economical support during the last two and a half years, without them and their help it would be impossible to conclude satisfactorily this step in my life. Special recognition to my wife Alicia and children, Diana and Alfredo Andres, for their moral support and sacrifices each of them did inside and above their own possibilities.

To all those persons and institutions that in one way or the other gave me the proper advice, moral or economical support at the right moment. They were many of them and I do not like to list them here because I am afraid to forget some of them involuntarily. However, I want to mention special gratitude to Ministerio de Obras Publicas de la Republica de Venezuela and its Directors for giving the time required and the necessary economical support which made this work possible.

Finally special acknowledgement to Dr. Frank A. Tillman, Dr. Doris Grosh, Dr. L. E. Grosh Jr. and Dr. Bob Smith for their willingness to be part of my advisory committee.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

CHAPTER I

INTRODUCTION

1.1. Bayes Theorem

Given a set of elements that can be categorized according to two different criteria, say sample outcomes A_j on one hand, and states of nature (hypothesis or parameter values) B_i on the other hand.

It is assumed that an event A_j may occur as a result of any of the states described by the B_i hypotheses. The prior probability of these B_i states are $P(B_i)$ $(i = 1, \ldots, K)$ and their sum $\sum_{i=1}^{K} P(B_i) = 1$, i.e. the states $\{B_i\}$ are mutually exclusive and exhaustive.

We assume that the conditional probabilities are known, that is, the probability that an event A_j occurs, given that the hypothesis B_i is true, is called the conditional probability

$$P(A_j|B_i) \qquad i = 1, \ldots, K.$$

We are interested in knowing: How does the probability of state B_i change when the additional information is available that the event A_j has actually happened? Bayes Theorem will provide an answer to this question.

The problem is to find the conditional (posterior) probability $P(B_i|A_j)$.

The joint probability of the compound event (A_j, B_i) is

$$\begin{aligned} P(A_j,B_i) &= P(B_i)\, P(A_j|B_i), \text{ or} \\ &= P(A_j)\, P(B_i|A_j), \end{aligned} \tag{1.1}$$

the right hand side being a consequence of the multiplicative law of probabilities.

Using the second form we have the alternative form

$$P(B_i|A_j) = \frac{P(A_j,B_i)}{P(A_j)} \tag{1.2}$$

The decomposition rule for compound events allows the denominator to be expressed in the form

$$P(A_j) = \sum_{k=1}^{K} P(A_j,B_k), \tag{1.3}$$

and using the first expression of (1.1) in equations (1.2) and (1.3) we have

$$P(B_i|A_j) = \frac{P(A_j,B_i)}{\sum_{k=1}^{K} P(B_k)P(A_j|B_k)} = \frac{P(B_i)P(A_j|B_i)}{\sum_{k=1}^{K} P(B_k)P(A_j|B_k)} = \frac{\text{Joint}}{\text{Marginal}} \tag{1.4}$$

Formula (1.4) is known as Bayes Theorem.

To summarize:

$P(B_i)$ = Prior probability of the hypothesis B_i before experimentation

$P(A_j|B_i)$ = Conditional probability for a given sample result A_j under the specified hypothesis B_i

$P(B_i|A_j)$ = Posterior probability of the hypothesis B_i after the outcome A_j has been observed.

In the literature most often the A_j represents a sample datum and the B_i a hypothesis to be tested, or a parameter to be estimated.

1.2. The Bayesian Procedure

The Bayesian procedure is based on formula (1.4), where the prior probability $P(B_i)$ associated with the hypothesis B_i represents the experimenter's preconceptions about the population being studied, and the probability $P(A_j|B_i)$ associated with the event A_j is the conditional probability.

The posterior distribution obtained by this method is used to estimate some parameter of interest or to decide which hypothesis B_i to accept.

The procedure of focusing the attention on the posterior distribution has the following advantages:

1) Allows the experimenter to introduce any preconceptions he may have about the population in study. These preconceptions may be the results of past experience or the conclusion of a theoretical study.

2) This approach agrees with the human mind because $P(B_i) = 0$ leads to $P(B_i|A_j) = 0$. In words, if a person does not give any credence to the fact that state B_i may obtain, no matter what the sample outcome may be, the posterior probability of that particular state always will be zero. This fact can be proved by Equation (1.4). Similarly $P(B_i) = 1$ leads to $P(B_i|A_j) = 1$. That is, if a person feels himself with a very strong idea about the absolute validity of a particular state and does not accept any argument in favor of other states, the posterior probability

of that particular state will be always total and equal to one whatever the experiment's outcomes are.

To prove this results, remember that states $\{B_i\}$ are mutually exclusive and exhaustive. Then in Equation (1.4) we have

$$P(B_i|A_j) = \frac{P(B_i)P(A_j|B_i)}{\sum_{k=1}^{K} P(B_k)P(A_j|B_k)} = \frac{P(B_i)P(A_j|B_i)}{P(B_i)P(A_j|B_i)} = 1$$

because $P(B_k) = 1$ if $k = i$

$= 0$ otherwise.

4) For some distributions it is possible to include the classical procedures as a special case.

The principal disadvantages of Bayes analysis are:

1) The estimates are often biased

2) Sometimes it is hard to quantify our preconceptions in the form of a prior distribution.

1.3. Overview of the Problem

Dr. Doris Grosh in [1,2,3] follows the Bayesian Procedure to solve the stratified allocation problem where she assumed the following conditions:

1) We have a finite stratified population of size N consisting of K strata.

2) The K strata are considered independent of each other; consequently if they are priorly independent they also are posteriorly independent of each other.

3) The population in the i^{th} stratum is finite and has N_i elements. Of these, M_i are classed as "defective" and the remainder $(N_i - M_i)$ are classed as "good".

Let

$$P_i = \frac{M_i}{N_i}$$

be defined as the fraction defective in the i^{th} stratum.

We are interested in making inferences about the P_i and linear functions of the P_i. To this end, we proceed with a sampling procedure. We decide to take in each stratum a sample (without replacement) of size n_i; of these, x_i are classed as "defective". So we are dealing in each stratum with a hypergeometric distribution given by

$$f_H(x_i|M_i,N_i,n_i) = \frac{\binom{M_i}{x_i}\binom{N_i-M_i}{n_i-x_i}}{\binom{N_i}{n_i}}, \quad x_i = 0, 1, 2, \ldots, n_i \tag{1.5}$$

with a mean value

$$E(\tilde{x}_i) = n_iP_i = n_i \frac{M_i}{N_i} \tag{1.6}$$

and variance

$$V(\tilde{x}_i) = n_iP_i(1-P_i) \frac{N_i-n_i}{N_i-1} \tag{1.7}$$

4) When the Bayesian Procedure is applied, the prior distribution should be chosen in such a way that it meets the conditions for being the natural conjugate of the conditional distribution, (in the sense

explained by Raiffa and Schlaifer in [5]).

Grosh concluded that the natural conjugate of the hypergeometric distribution is the Beta-Binomial distribution (see [1,2]) with probability function of the form

$$f_{\beta B}(M|N,a,b) = \binom{N}{M} \frac{B(M+a,\ N-M+b)}{B(a,b)} \qquad M = 0,\ 1,\ \ldots,\ N. \tag{1.8}$$

It can be proved that this distribution has moments

$$E(\tilde{M}|N,a,b) = N \frac{a}{a+b} \tag{1.9}$$

and

$$Var(\tilde{M}|N,a,b) = \frac{N\ a\ b\ (a+b+N)}{(a+b)^2 (a+b+1)} \tag{1.10}$$

Fig. 1.1 and 1.2 show the wide variety of shapes of this rich family for selected values of parameters a and b.

5) Grosh applied the Bayes Procedure to the set of distributions formed by the hypergeometric as conditional distribution with the Beta-Binomial as prior distribution and found the following posterior distribution

$$f^*_{\beta B}(M|x;N,a,b,n) = \binom{N-n}{M-x} \frac{B(M+a,\ N-M+b)}{B(x+a,\ n-x+b)}, \qquad x \leq M \leq N-n+x \tag{1.11}$$

which also is a Beta-Binomial distribution.

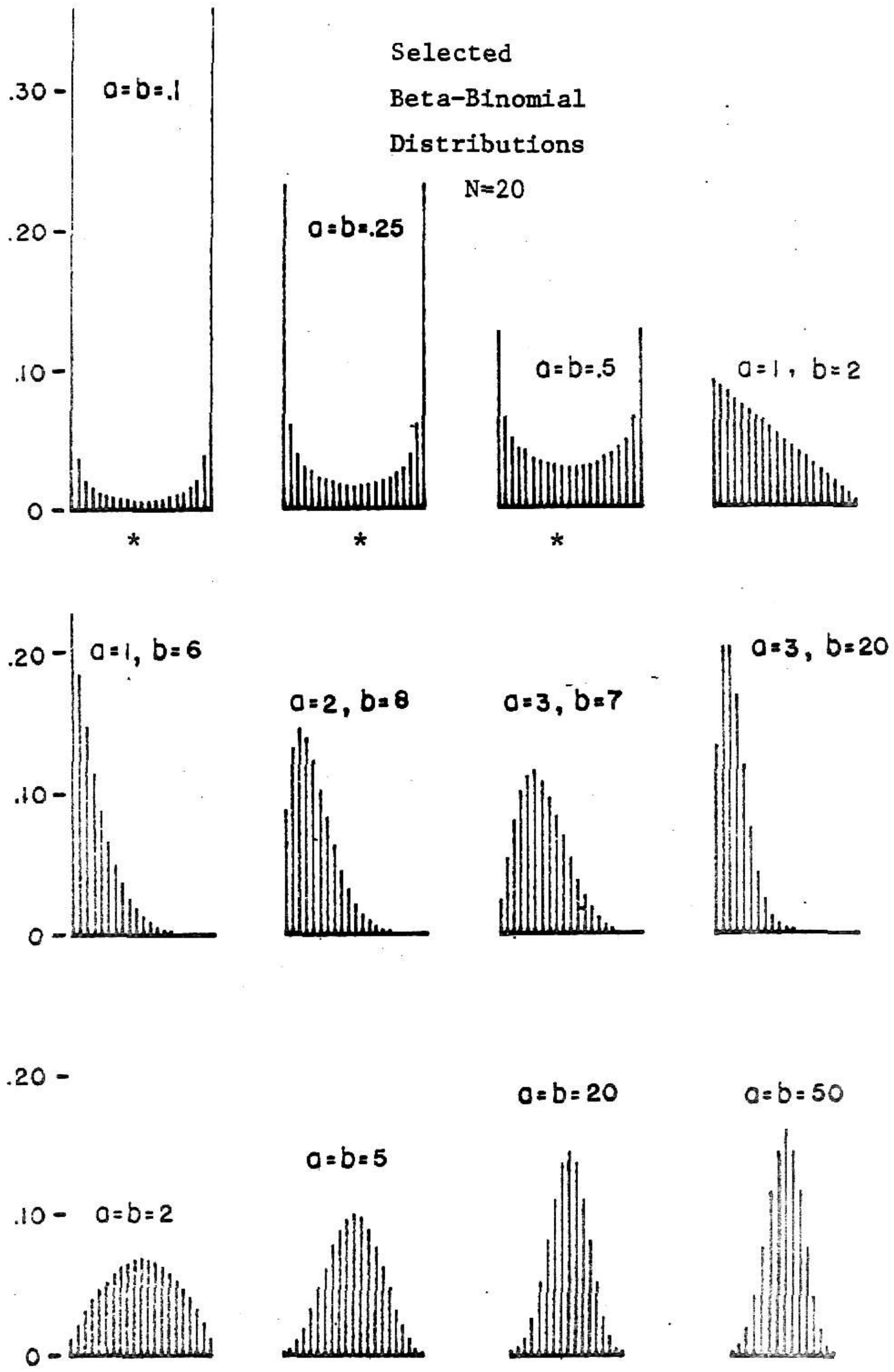


Figure 1.1

Reproduced from [2] with the Author's Permission.

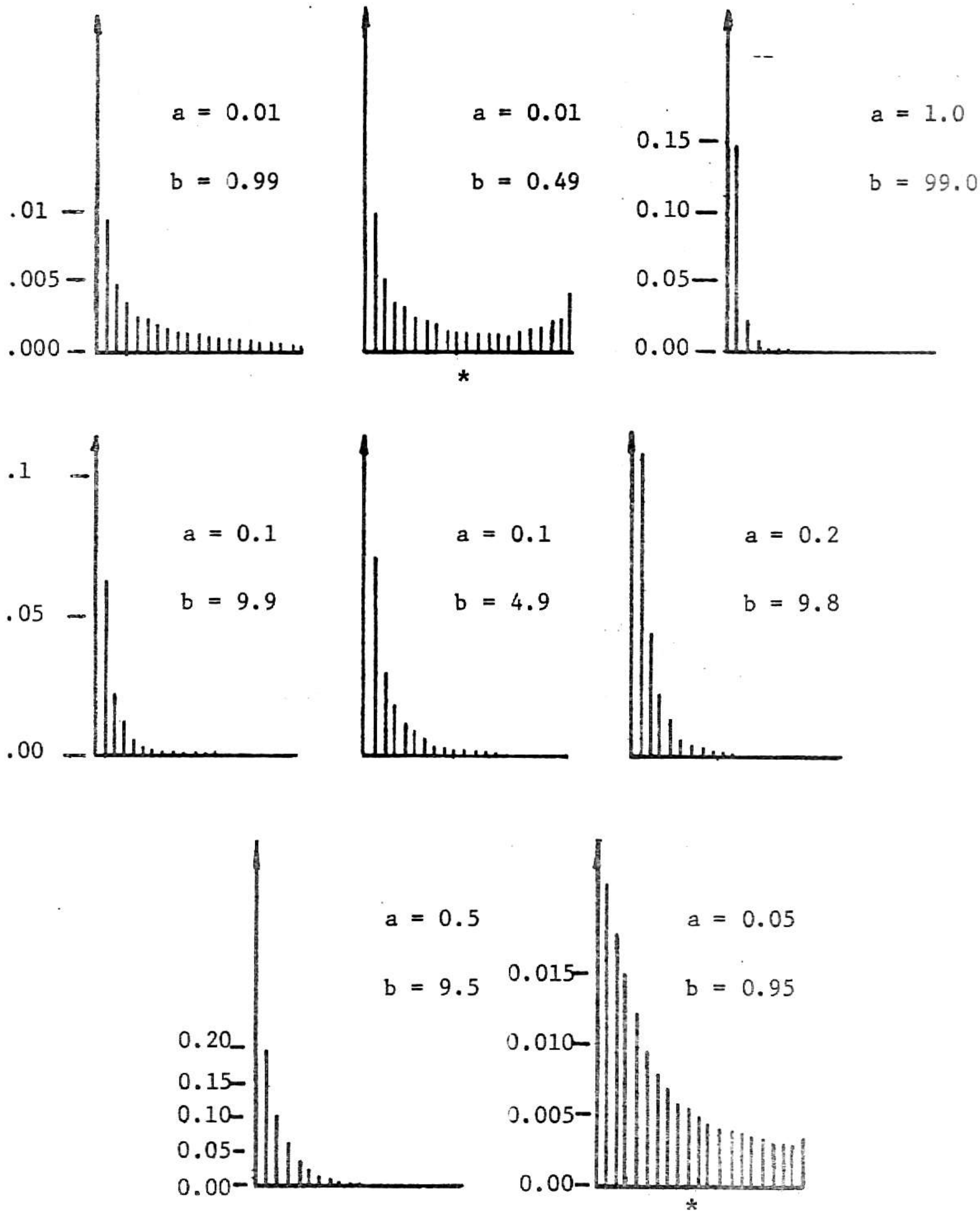


Figure 1.2

Selected Beta-Binomial Distributions (N = 20).

The two posterior moments are given by

$$E(\tilde{M}|x;N,a,b,n) = x + \frac{(a+x)(N-n)}{(a+b+n)} \quad (1.12)$$

$$Var(\tilde{M}|x;N,a,b,n) = Var(\tilde{M}-x|x,N,a,b,n) =$$

$$= \frac{(N-n)(a+b+N)(a+x)(n+b-x)}{(a+b+n)^2\ (a+b+n+1)} . \quad (1.13)$$

1.4. Theoretical Procedure

The theoretical procedure we followed in the present work was developed by Dr. Doris Grosh in [1,2,3] along the lines laid out by Zacks [6].

For the reader's benefit we summarize here their work.

Problem definition:

Given a finite, stratified and dichotomous population with

a) K strata priorly independent of each other,

b) P_i, the fraction defective of the i^{th} stratum, $P_i = \frac{M_i}{N_i}$

c) λ_i, a factor representing the weight or importance the experimenter assigns to the i^{th} stratum.

We are interested in finding the optimum allocation of stratum sample sizes for estimating

$$\theta = \sum_{i=1}^{K} \lambda_i P_i \quad (1.14)$$

subject to the budgetary restriction expressed by

$$\sum_{i=1}^{K} c_i n_i \leq C \tag{1.15}$$

where c_i is the cost of making one observation in the i^{th} stratum

n_i is the number of observations made in the i^{th} stratum (sample size)

C is the total budget for sampling alone. Set up costs are not included.

A time honored criterion for rating the desirability of an estimator for a statistical variable is its sampling variability, as measured by the mean square error. Smaller square error means higher desirability.

With this idea in mind they introduced a squared loss function of the form

$$L(\theta,\hat{\theta}) = (\theta-\hat{\theta})^2 \tag{1.16}$$

where θ is defined by Eq. (1.14) and $\hat{\theta}$ is an estimator of θ based on the sample outcomes. Its structure will be defined below.

In order to avoid a possible source of confusion due to the large number of parameters in the distributions under discussion we will use the following vector notation instead of using a full listing of all of them.

$\underline{N} = (N_1,N_2, \ldots, N_K)$ represents the stratum sizes.

$\underline{M} = (M_1,M_2, \ldots, M_K)$ represents the number of total defective units in the strata.

$\underline{c} = (c_1, c_2, \ldots, c_K)$ represents the unit cost of sampling in the different strata.

$\underline{n} = (n_1, n_2, \ldots, n_K)$ represents the sample size of the sampling procedure.

$\underline{X} = (x_1, x_2, \ldots, x_K)$ represents the number of defective units we got from an actual sampling procedure.

Note that in the present work, the letter $\tilde{x}$ stands for the random variable, whereas x stands for an observation or realization of $\tilde{x}$.

Single stage scheme:

In our assumptions we said that the prior distributions as well as the sampling procedures are independent by strata. Consequently the joint conditional probability of the sample outcomes $\underline{\tilde{X}} = (\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n)$ is given by the product over all K's of the hypergeometric distribution

$$f(\underline{X}|\underline{M},\underline{n}) = \prod_{i=1}^{K} f(x_i|N_i,M_i,n_i). \quad (1.17)$$

In the same way the joint prior probability of the stratum defective totals $\tilde{M} = (\tilde{M}_1, \tilde{M}_2, \ldots, \tilde{M}_K)$ is also given by the product over all K's of the Beta-Binomial distributions

$$f(\underline{M}) = \prod_{i=1}^{K} f_{\beta B}(M_i|N_i,a_i,b_i). \quad (1.18)$$

The factorability of the previous probabilities assures us of the same property of the marginal and the posterior probabilities.

Now with the loss function defined as in Equation (1.16)

$$L(\theta,\hat{\theta}) = (\theta-\hat{\theta})^2,$$

the estimator $\hat{\theta}$ based on the sample outcome $(\hat{\underline{X}})$ is given by

$$\hat{\theta} = \sum_{i=1}^{K} \lambda_i \hat{P}_i = \lambda_i \frac{\hat{M}_i}{N_i} . \tag{1.19}$$

where $\hat{M}_i = \hat{M}_i(\tilde{\underline{X}})$ is a function of the observations and is an estimate of M_i. From Equation (1.12)

$$\hat{M}_i = x_i + \frac{(N_i-n_i)(a_i+x_i)}{(a_i+b_i+n_i)} .$$

This is so because under squared error function the Bayes estimator is the posterior mean.

The posterior risk or Bayes risk associated with this estimator is given by its posterior variance

$$R(\theta,\hat{\theta}|\underline{n}) = \text{Var}\ (\hat{\theta}|\underline{X}) = \sum_{i=1}^{K} \frac{\lambda_i^2}{N_i^2} \text{Var}\ (\hat{M}_i|\underline{X}) \tag{1.20}$$

Substituting $\hat{M}_i$ in Equation (1.19) we have

$$\hat{\theta} = \hat{\theta}(\tilde{\underline{X}}) = \sum_{i=1}^{K} \frac{\lambda_i}{N_i} \left[\tilde{x}_i + \frac{(N_i-n_i)(a_i+x_i)}{(a_i+b_i+n_i)}\right] = \sum_{i=1}^{K} \frac{\lambda_i}{N_i} \left[\frac{(x_i+a_i)(a_i+b_i+N_i)}{(a_i+b_i+n_i)} - a_i\right] \tag{1.21}$$

and substituting Var $(M_i|\underline{X})$ in Equation (1.20) we have

$$R(\theta,\hat{\theta}|\underline{n}) = \sum_{i=1}^{K} \frac{\lambda_i^2}{N_i^2} \frac{(N_i-n_i)(a_i+x_i)(n_i+b_i-x_i)(a_i+b_i+N_i)}{(a_i+b_i+n_i)^2\ (a_i+b_i+n_i+1)} . \tag{1.22}$$

This is the quantity to be minimized by suitable choice of $\underline{n}$, but this equation can not be used to determine the allocation $\underline{n}$, since it is a function of $\underline{X}$ and by the time $\underline{X}$ is known, is too late to determine the optimal allocation. Thus, in order to solve this problem we proceed to average the risk over all possible future outcomes before taking any sample.

To do so we use the joint marginal probability function given by

$$f(\underline{X}) = \prod_{i=1}^{K} f_{\beta B}\ (x_i|a_i,b_i,n_i). \tag{1.23}$$

The resulting expected value is the prior risk

$$\rho(\underline{n}) = \sum_{i=1}^{K} \frac{\lambda_i^2}{N_i^2} \frac{(N_i-n_i)(a_i+b_i+N_i)\ a_i b_i}{(a_i+b_i+n_i)(a_i+b_i)(a_i+b_i+1)} \tag{1.24}$$

A more attractive form of Equation (1.24) is

$$\rho(\underline{n}) = \sum_{i=1}^{K} \gamma_i \left(\frac{1}{a_i+b_i+n_i} - \frac{1}{a_i+b_i+N}\right), \tag{1.25}$$

where

$$\gamma_i = \frac{\lambda_i^2\ (a_i+b_i+N_i)\ a_i b_i}{N_i^2\ (a_i+b_i)(a_i+b_i+1)} . \tag{1.26}$$

Now our problem is to minimize $\rho(\underline{n})$, subject to the budgetary constraint given by

$$\sum_{i=1}^{K} c_i n_i \leq C \tag{1.15}$$

To get the minimum value of Equation (1.25) Grosh used the technique of Lagrangian multipliers. Of course she obtained an approximate solution that must be rounded to obtain integer values for all the n_i.

That solution is given by the following equation

$$n_i = \frac{C + \sum_{j=1}^{K} (a_j+b_j)\ c_j}{\sum_{j=1}^{K} \sqrt{\gamma_j c_j}} \sqrt{\gamma_i c_i^{-1}} - (a_i+b_i) \tag{1.27}$$

It is possible that for some strata the above equation may give some negative allocations. In those cases we set the corresponding $n_i = 0$ and resolve Equation (1.27) leaving out those strata.

For convenience she defined an indicator function as follow:

$$\begin{cases} J_j = 1 & \text{if the } j^{th} \text{ stratum is to be sampled} \\ = 0 & \text{otherwise.} \end{cases} \tag{1.28}$$

Letting

$$C^* = C + \sum_{j=1}^{K} J_j\ (a_j+b_j) c_j$$

and substituting into (1.27) yields the optimum allocation in the i^{th} stratum

$$n_i^o = \max\left\{0, \left(\frac{C^* \sqrt{\gamma_i c_i^{-1}}}{\sum_{j=1}^{K} J_j \sqrt{\gamma_j c_j}} - (a_i+b_i)\right)\right\} \qquad i = 1,2,\ldots, K . \quad (1.29)$$

After substituting (1.29) into (1.25) we obtain the minimum risk for the optimal allocation

$$\rho(\underline{n}^o) = \frac{1}{C^*}\left(\sum_{j=1}^{K} J_j \sqrt{\gamma_j c_j}\right)^2 - \sum_{j=1}^{K} \frac{\gamma_j}{a_j + b_j} + \sum_{j=1}^{K} \frac{N_j \gamma_j}{(a_j+b_j)(a_j+b_j+N_j)} \quad (1.30)$$

Restricted sampling cases:

Sometimes an experimenter may find himself in the situation that for economic reasons (set up costs, say) he can not sample all the strata given by the solutions of Equation (1.29). Those economic reasons are assumed to be independent of the one already mentioned where we require only that $\sum_{i=1}^{K} c_i n_i \leq C$.

If we study Equation (1.29) a little deeper we note that $\sum_{j=1}^{K} J_j$, the number of strata to be sampled, is obviously a function of the total budget available C. Thus, larger values of C will result in fewer zero values for the J_j's. Now if we assume for the moment that C is a variable, there will be some least value of C, say $C^{(0)}$, for which all the strata could be sampled. (Note that the superscript is an indicator of how many strata are excluded for economic reasons). To find that $C^{(0)}$

value, set $n_i^{(0)} > 0$ for all $i = 1,2, \ldots, K$ and solve for C

$$\frac{C^{(0)} + \sum_{j=1}^{K} (a_j+b_j)\, c_j}{\sum_{j=1}^{K} \sqrt{\gamma_j c_j}} > \frac{(a_i+b_i)\sqrt{c_i}}{\sqrt{\gamma_i}} \qquad i = 1,2, \ldots, K. \tag{1.31}$$

It is convenient to let

$$D_i = (a_i+b_i)\sqrt{\frac{c_i}{\gamma_i}} \qquad i = 1,2, \ldots, K \tag{1.32}$$

and reorder the strata, if necessary, so that

$$D_K < D_{K-1} < \ldots < D_2 < D_1 < \infty$$

Then all the strata may be sampled if C is large enough that

$$g_{(1)}(C) = \frac{C + \sum_{j=1}^{K} c_j(a_j+b_j)}{\sum_{j=1}^{K} \sqrt{\gamma_j\, c_j}} > D_1. \tag{1.33}$$

More generally Grosh defined

$$g_m(C) = \frac{C + \sum_{j=m}^{K} c_j\, (a_j+b_j)}{\sum_{j=m}^{K} \sqrt{\gamma_j\, c_j}} \tag{1.34}$$

It is possible to obtain a sequence of "cut off" cost values for determining the number of strata to be sampled for every possible budget, by the relationship

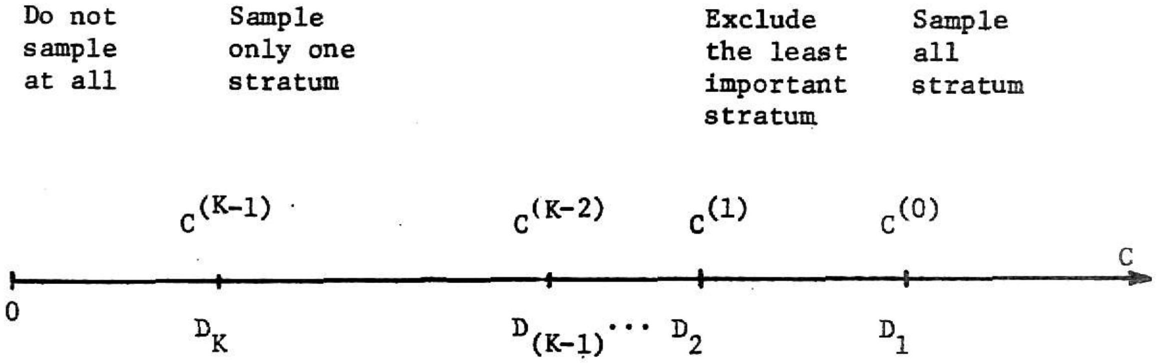
$$C^{(P)} = \min_{C} \left\{ C \middle| g_m(C) > D_m \right\} \qquad \begin{array}{l} P = m - 1 \\ m = 1, 2, \ldots, K, \end{array} \tag{1.35}$$

in other words, as the solution of

$$\frac{C + \sum_{j=m}^{K} c_j(a_j + b_j)}{\sum_{j=m}^{K} \sqrt{\gamma_j \; c_j}} = D_m. \tag{1.36}$$

The sequence $\{C^{(P)}\}$ is monotone non-increasing.

The values of D_i obtained from Equation (1.32) represent an importance relationship among the strata. It is a function of all the stratum parameters and is used to determine the order in which the strata are going to be arranged for our sampling procedure or excluded from it in the necessary cases.



Double stage scheme:

In general the use of multiple stage processes is more efficient than the single stage. The information obtained in the earlier stages is used by the experimenter in deciding how to proceed with the future stages in order to reduce the variance of his estimator.

Grosh approached and developed the solution for finding the optimal allocation in a two-stage procedure for the problem already stated.

As explained in [1,3] for easy programming there was no attempt to establish the optimal partitioning between the two stages, of the total budget of C dollars. Insteady, she assumed various values of C1 dollars for the first stage and C2 dollars for the second stage which are arbitrarily fixed and subject to C1 + C2 = C.

The solution of the problem of determining the optimal split of the total budget is given by the combination of C1 and C2 that provides the smallest risk for the optimal first stage allocation after a trial and error process.

With the same assumptions as before the following procedure was developed. In the i^{th} stratum draw a first stage sample of n_i; $\tilde{x}_i$ of which are classed as "defective". As a consequence the i^{th} stratum now consist of

$N_i - n_i$ individuals,

$M_i - \tilde{x}_i$ of them defectives, and

$N_i - n_i - M_i + \tilde{x}_i$ non defectives.

A second stage sample of size m_i is drawn of which $\tilde{y}_i$ are found to be defectives. The conditional probability of $\tilde{y}_i$ based on the stratum structure between the two samples processes is (suppressing subscripts) the hypergeometric distribution

$$f_H(y|(N-n),(M-x),m) = \frac{\binom{M-x}{y}\binom{N-n-M+x}{m-y}}{\binom{N-n}{m}} \qquad \begin{array}{r} x = 0,1,\ldots,n \\ y = 0,1,\ldots,m \\ x+y = 0,1,\ldots,M \end{array} \tag{1.37}$$

The joint probability of $\tilde{x}$ and $\tilde{y}$ is given by the TWO-STAGE HYPERGEOMETRIC probability defined by

$$f_{TSH}(x,y|N,M,n,m) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{m}} \cdot \frac{\binom{M-x}{y}\binom{N-n-M+x}{m-y}}{\binom{N-n}{m}} \qquad \begin{array}{l} x = 0,1,\ldots,n \\ y = 0,1,\ldots,m \\ x \leq M \\ y \leq M-x. \end{array} \tag{1.38}$$

After re-arrangement of the factors she obtained

$$f_{TSH}(x,y|N,M,n,m) = \frac{\binom{M}{y}\binom{N-M}{m-y}}{\binom{N}{m}} \cdot \frac{\binom{M-y}{x}\binom{N-M-m+y}{n-x}}{\binom{N-m}{n}} \qquad \begin{array}{l} x = 0,1,\ldots,n \\ y = 0,1,\ldots,m \\ x+y \leq M \end{array} \tag{1.39}$$

On the right hand side the first factor is the marginal probability of $\tilde{y}$ and the second factor is the conditional probability of $\tilde{x}$ given y.

Let $z = x+y$; then $y = z-x$ and substitute into Equation (1.39).

We get

$$f(z,y|N,M,n,m) = \frac{\binom{M}{z}\binom{N-M}{m+n-z}}{\binom{N}{m+n}} \frac{\binom{z}{x}\binom{M+n-z}{n-x}}{\binom{M+n}{n}} \quad \begin{array}{l} x = 0,1, \ldots, z \\ z = 0,1, \ldots, m+n. \end{array} \tag{1.40}$$

As before, a Beta-Binomial prior distribution with parameters a and b is assigned to $\tilde{M}$.

The joint probability of $\tilde{z}$ and $\tilde{M}$ is given by

$$(z,M|N,n,m,a,b) = \frac{\binom{M}{z}\binom{N-M}{m+n-z}}{\binom{N}{m+1}} \binom{N}{M} \frac{B(M+a,N-M+b)}{B(a,b)} \quad \begin{array}{l} z = 0,1, \ldots, m+n \\ M = 0,1, \ldots, N \end{array} \tag{1.41}$$

and leads to the following posterior distribution of M given z,N,n,m,a,b

$$f_{\beta B}(M|z,\ldots,b) = \binom{N-n-m}{M-z} \frac{B(M+a,N-M+b)}{B(a+x+y,\ b+n+m-x-y)} \tag{1.42}$$

$$M = z,\ z+1,\ \ldots,\ N-n-m+z\ ,$$

with the posterior mean of $\tilde{M}$ given z equal to

$$E(\tilde{M}|z,\ldots,b) = x+y + \frac{(N-m-n)(a+x+y)}{a+b+n+m} \tag{1.43}$$

Recalling that $P = M/N$, after some re-arrangement of term we have that the Bayes estimate of P is given by

$$\hat{P} = \hat{P}(z) = \frac{a+x+y}{N}\,\frac{a+b+N}{a+b+n+m} - \frac{a}{N} \tag{1.44}$$

and the posterior variance (or Bayes risk) after two stages of sampling in any stratum is

$$\mathrm{Var}\,(\hat{P}) = \frac{(N-m-n)(a+b+N)(a+\tilde{x}+\tilde{y})(b+m+n-\tilde{x}-\tilde{y})}{N^2(a+b+n+m)^2(a+b+n+m+1)} \tag{1.45}$$

Since $\hat{\theta} = \sum_{i=1}^{K} \lambda_i \hat{P}_i$ the posterior variance of $\hat{\theta}$, based on the sampling outcomes in all the strata, is

$$R(\underline{\tilde{X}},\underline{\tilde{Y}}|\underline{m},\underline{n}) = \sum_{i=1}^{K} \frac{\lambda_i^2(N_i-m_i-n_i)(a_i+b_i+N_i)(a_i+\tilde{x}_i+\tilde{y}_i)(b_i+m_i+n_i-\tilde{x}_i-\tilde{y}_i)}{N^2(a_i+b_i+m_i+n_i)^2\;(a_i+b_i+n_i+m_i+1)} \tag{1.46}$$

In order to obtain the optimal two-stage allocation we need to complete the following steps.

a) Find the expected value of Equation 1.46 with respect to the joint marginal distribution of the $\tilde{y}$'s given the x's.

b) Optimize the second stage allocation $\underline{m}$ as function of the $\tilde{x}$'s in the same form as in the single stage explained before.

c) Evaluate the expected value with respect to the joint marginal distribution of the $\tilde{x}$'s.

d) Optimize the first stage allocation $\underline{n}$.

The first step depends on the relationship

$$E(a+x+\tilde{y})(b+m+n-x-\tilde{y}) = \frac{(a+x)(b+n-x)(a+b+n+m)(a+b+n+m+1)}{(a+b+n)(a+b+n+1)} \qquad (1.47)$$

developed in Appendix E of [1].

Grosh defined the term INTERMEDIATE RISK to indicate the risk which is prior to the second stage but posterior to the first stage and designated it as $R(\underline{\tilde{X}}|\underline{n},\underline{m})$.

Combining (1.46) and (1.47) yields

$$R(\underline{\tilde{X}}|\underline{n},\underline{m}) = \sum_{i=1}^{K} \frac{\lambda_i^2 (N_i-m_i-n_i)(a_i+b_i+N_i)(a_i+b_i)(b_i+n_i-\tilde{x}_i)}{N_i^2(a_i+b_i+n_i+m_i)(a_i+b_i+n_i)(a_i+b_i+n_i+1)} =$$

$$(1.48)$$

$$= \sum_{i=1}^{K} \gamma_i^{(2)}(\tilde{x}_i)\left(\frac{1}{a_i+b_i+n_i+m_i} - \frac{1}{a_i+b_i+N_i}\right)$$

where $\gamma_i^{(2)}(\tilde{x}_i) = \delta_i^{(2)}(\tilde{x}_i)(a_i+b_i+N_i) \qquad i = 1,2,\ldots, K \quad (1.49)$

and $$\delta_i^{(2)}(\tilde{x}_i) = \frac{\lambda_i^2(a_i+b_i+N_i)(a_i+x_i)(b_i+n_i-\tilde{x}_i)}{N_i^2(a_i+b_i+n_i)(a_i+b_i+n_i+1)} \qquad i = 1,2,\ldots, K \quad (1.50)$$

Following the technique used in the single stage sampling, the optimal second stage allocation was obtained by the method of Lagrangian multipliers.

The result is

$$m_i^o = m_i^o(\underline{\tilde{X}};n) = \max\left\{0, \left(\frac{C2 + \sum_{j=1}^{K} c_j(a_j+b_j+n_j)L_j}{\sum_{j=1}^{K} L_j\sqrt{\gamma_j^{(2)}(\tilde{x}_j)\ c_j}}\sqrt{\gamma_i^{(2)}(\tilde{x}_i)c_i^{-1}} - (a_i+b_i+n_i)\right)\right\}$$

$$i = 1, 2, \ldots, K \qquad (1.51)$$

where the L_j is an indicator function defined as

$$\begin{cases} L_j = 1 \text{ if the } j^{th} \text{ stratum is sampled} \\ \quad\ = 0 \text{ otherwise} \end{cases} \qquad (1.51a)$$

Note that it is very important to keep in mind that L_j is an implicit function of $\underline{\tilde{X}}$ and $\underline{n}$ but for convenience we use L_j instead the full form $L_j(\underline{\tilde{X}};\underline{n})$.

Defining

$$C^* = C2 + \sum_{j=1}^{K} L_j(a_j+b_j+n_j)c_j. \qquad (1.52)$$

and substituting from (1.51) and (1.52) into (1.48), she found that the intermediate risk corresponding to the allocation $\underline{m}^o$ is

$$R(\underline{\tilde{X}};\underline{n},m^o(\underline{X};n)) = \frac{Q^2}{C^*} - \sum_{j=1}^{K} L_j\ \delta_j^{(2)}(\tilde{x}_j) + \sum_{j=1}^{K} (1-L_j)\ \delta_j^{(2)}(\tilde{x}_j)\ \frac{N_j-n_j}{a_j+b_j+n_j}\ ,$$

$$(1.53)$$

where

$$Q = \sum_{j=1}^{K} L_j \sqrt{\gamma_j^{(2)}(\tilde{x}_j)\ c^{-1}} \ . \qquad (1.54)$$

In order to continue with the minimization process it is necessary to average (1.53) with respect to the joint marginal probability function of the $\tilde{x}$'s.

Due to the radical in (1.54) which is part of Equation (1.53) the latter does not lend itself algebraically to the averaging process and numerical methods must be applied.

In order to begin the numerical process a trial first stage allocation ($\underline{n}^{(1)}$) is chosen and optimization is achieved through a searching procedure.

Poor choice of $\underline{n}^{(1)}$ only increases the number of iterations needed to obtain that optimum.

Grosh sugests some methods for choosing the first trial first stage allocation:

a) Classical: stratum sample size proportional to stratum size;

b) The single stage optimum allocation given in equation (1.29) based on budget C1;

c) Using the approximation

$$n_i^{(1)} = C1 \frac{\sqrt{\gamma_i\ c_i^{-1}}}{\sum_{i=1}^{K} \sqrt{\gamma_i c_1^{-1}}} \ ,$$

where γ_i is defined in Equation (1.26); and

d) Arbitrary $\underline{n}^{(1)}$.

For the chosen initial first stage allocation an average must be obtained over every possible resulting vector $\tilde{\underline{X}}$ using the joint marginal probability function of the $\tilde{x}$'s shown in Equation (1.48).

The obtained result is the prior risk

$$\rho(\underline{n}^{(1)}, \underline{m}^{o}(\underline{n}^{(1)})) = E\left\{R(\tilde{X}, n^{(1)}, m^{o}(n^{(1)}))\right\}, \tag{1.55}$$

where the expectation is taken using the multivariate probability function

$$f(\underline{X}) = \prod_{i=1}^{K} \binom{n_i}{x_i} \frac{B(a_i+x_i, b_i+n_i-x_i)}{B(a_i, b_i)} \tag{1.56}$$

1.5 The Searching Procedure

<u>Definition</u>: $i_{(m)}$ is the index of the stratum when the strata have been ranked according to the size of the importance numbers D_m defined in Equation (1.32).

The procedure is:

a) Determine $\underline{n}^{(1)}$ as you prefer.

b) Compute $\rho[\underline{n}^{(1)}, m^{o}(\underline{n}^{(1)})]$ and save this value.

c) Construct a second first stage allocation $\underline{n}^{(2)}$ as follows:

Using the D_i values as stratum importance indices (remember smaller D_i values means more important stratum), keep fixed all the components of $\underline{n}^{(1)}$ except the most important and the least important.

Increase by one unit the sample size in the most important stratum $i_{(K)}$ and modify the least important stratum $i_{(1)}$ (if $n_{i_{(1)}} \neq 0$, otherwise go to strata $i_{(2)}$, etc) reducing its sample size until the cost constraint is satisfied.

d) Compute $\rho^* = \rho[\underline{n}^{(2)}, m^o(\underline{n}^{(2)})]$. If this ρ^* value is smaller than the previous ρ value save it and $\underline{n}^{(2)}$. Repeat the process from c) but using these new values as comparison base. If this ρ^* value is larger, save the original ρ and $\underline{n}^{(1)}$ values and repeat the process from c), but now trying to increase the sample size of the next most important stratum $i_{(K-1)}$.

e) Once the point is reached where no improvement can be achieved by the process explained in d), we proceed as follows:

Beginning in the next most important stratum, call it $i_{(K-q)}$, that stratum where the sample size was last increased, decrease the sample size of this stratum by one and make the proper adjustment increasing the sample size of the next most important stratum $i_{(K-q-1)}$ say, in order to meet the cost constraint.

f) Evaluate $\rho^* = \rho[\underline{n}^{(2)}, m^o(\underline{n}^{(2)})]$. If this ρ^* value is smaller than the previous saved ρ value, save the new ρ^* and $\underline{n}^{(2)}$. Repeat the process from b) using the new values as comparison base and using as the most important stratum, that one given by $i_{(K)}$.

If this ρ^* value is larger than the saved one, keep the saved ρ value as well as the $\underline{n}^{(1)}$ and repeat the process from e) but now trying to decrease by one unit the sample size of the stratum next in importance, say $i_{(K-q-1)}$ to that in which we just intended to reduce its sample size.

g) The searching procedure terminates when the optimal allocation is reached, that is when $\rho\ (\underline{n}, \underline{m}^{o}\ (\underline{n}))$ increases for any change in any component n_i $(i = 1,2, \ldots, K)$.

CHAPTER 2

ANALYSIS OF THE PROBLEM

2.1. Introduction

The motivation for the present work was founded upon some work by Grosh [3]. A number of unanswered questions arose with this work and are summarized below:

1) The prior variance of P_i in the two cases showed is small;
2) In the two cases used as examples the ratio $c_2/c_1 = 1$;
3) The author tried the very special cases in which both strata are identical except for the importance rating λ_i;
4) The author did not mention any kind of practical application to justify the prior expectations and the strata sizes used in her studies (see [1], [3]).

A propos the last observation, it should be pointed out here that the present work is oriented basically to some kind of quality control situation, hence the range of values given to the parameter a_i and b_i in the present work.

With them are obtained expected percent defective values between 1 and 10, common values in that kind of work.

However, the same procedure can be used for other purposes, as those mentioned by Grosh in [1]; for instance to establish the viewpoint (accept or reject) of several communities (strata) on some social or policial issue. In this example the expected percent "defective" can be as large as necessary for the particular problem.

At the same time, there is a question of the large value of N_i compared with (a_i+b_i). At this point we want to say that in general when we decide to try some kind of multiple stage sampling scheme it is because the population in hand is large; otherwise we would probably prefer a simple sample procedure; see [4].

2.2. The Prior Variance of P_i

The suggested question of examining some cases when the prior variance for the fraction defective of each stratum is large, was attempted.

This prior variance is given by

$$V_i = \frac{\left(1 + \frac{a_i+b_i}{N_i}\right)\left(\frac{a_i}{a_i+b_i}\right)\left(\frac{b_i}{a_i+b_i}\right)}{(a_i+b_i+1)} = f(a_i,b_i), \text{ say.}$$

Since $f(a_i,b_i)$ is continuous and has continuous partial derivatives with respect to a_i and b_i, we could find a value for a_i and b_i, which would maximize $f(a_i,b_i)$ by solving the simultaneous equations

$$\frac{\partial f(a_i,b_i)}{\partial a_i} = 0$$

$$\frac{\partial f(a_i,b_i)}{\partial b_i} = 0. \qquad (2.2)$$

An equivalent representation of V_i which is easier to work at this time is (suppressing the subscripts for convenience),

$$V = \frac{1}{N}\,\frac{(a+b+N)ab}{(a+b)^2\;(a+b+1)} \tag{2.3}$$

then

$$\frac{\partial V}{\partial a} = \frac{1}{N}\,\frac{[(a+b+N)b+ab](a+b)^2(a+b+1)\;-\;(a+b+N)ab[(a+b)^2+2(a+b+1)(a+b)]}{(a+b)^4\;(a+b+1)^2} \tag{2.4}$$

$$\frac{\partial V}{\partial b} = \frac{1}{N}\,\frac{[(a+b+N)a+ab](a+b)^2(a+b+1)-(a+b+N)ab[(a+b)^2+2(a+b+1)(a+b)]}{(a+b)^4\;(a+b+1)^2} \tag{2.5}$$

As can be seen, the obtained equations, when set equal to zero, are equations of high degrees in a_i and b_i which are not easily solved.

Rather than use the analytical method we will appeal to the following intuitive procedure.

The term $(\frac{a}{a+b})(\frac{b}{a+b})$ can be proved to have a maximum equal to 0.25 when a = b.

If we substitute this part of the V equation for its maximum possible value of 0.25 we get

$$V \leq V^* = \frac{1}{4}\,\frac{\left(1 + \frac{a+b}{N}\right)}{(a+b+1)} = \frac{1}{4N}\,\frac{(a+b+N)}{(a+b+1)} \tag{2.6}$$

$$\frac{\partial V^*}{\partial a} = \frac{1}{4N}\,\frac{(a+b+1)\;-\;(a+b+N)}{(a+b+1)^2} = \frac{1-N}{4N(a+b+1)^2} \tag{2.7}$$

$$\frac{\partial V^*}{\partial b} = \frac{1}{4N} \frac{(a+b+1) - (a+b+N)}{(a+b+1)^2} = \frac{1-N}{4N(a+b+1)^2} \qquad (2.8)$$

If we solve this system of equations (when the derivatives are set equal to zero), we obtain the result that the maximum V* is obtained when N = 1, no matter what the values of a and b are.

Now, putting this two facts together, N = 1 and a = b, we get

$$\text{Max } V = \frac{(1+2a)}{4(2a+1)} = \frac{1}{4} .$$

That means that the question of having very large prior variances for P_i is not possible.

The other observations will be studied in the next chapters of the present work.

2.3 Computing program problems

The original program that was used for Grosh's work (see [1], appendix F) was found to have some problems connected with it. The most important are:

1) The program is very slow. Consequently it consumed too much computer time and was too expensive.
2) The program only worked for integer a_i and b_i and the special case when the sampling cost c_i is the same for all strata.

We effected a revision of the program and made the following changes in order to speed up the program and make it more general.

A) The subroutine RISK was changed as follows:

1) Instead of evaluating the probability of all the possibles outcomes of a given sample size, as well as the corresponding $\gamma_i^{(2)}$ value, the present program evaluates those values up to the point in which the remaining (right hand tail) probabilities add to less than or equal to 1×10^{-6}.

2) When the allocation of the second stage occurs, the strata that had been eliminated because their initial allocation was negative, are no longer reactivated, as was erroneously done in the original program, according to equation (1.27) and the following explanation.

3) A modification was introduced that allows us to use the same program for single stage sampling schemes, skipping the unnecessary part of the program for this particular case in which C2 = 0.0

B) In the ALLOCN subroutine a modification was introduced similar to the second one done in the RISK subroutine but now for the first stage allocation. In addition, a modification was made in the last part of ALLOCN in order to allow us to work properly when the sample costs in the strata are different.

C) The VALUE subroutine was left unmodified.

D) The INDEX subroutine was completely re-done in order to improve the search procedure as will be explained below.

We also include in the new INDEX subroutine the necessary steps for allowing us to use the program when the sampling costs c_i among the strata are not equal.

The search procedure was modified by introducing a directional device for proceding in the risk-minimizing direction once it was found.

In the original program it was true only if the improvement was achieved by increasing the allocation in the most important stratum. Otherwise, the searching procedure was done in such a way that it proceded forward and backward, repeating the full evaluation of possible solutions which had already been improved.

E) The MAIN PROGRAM was modified as follows:

1) We introduce a device that allows us to choose the type of initial first stage allocation instead of the usual one (which is obtained by default and given by the ALLOCN subroutine). These "forced starting values" are read in as data.

2) The search procedure part of this section of the program was modified in such a way to complement the new INDEX subroutine.

In general the program as in its present form, is capable of working with up to K = 9 strata. However, it will be easy to modify, for larger K, by changing the proper "INTEGER" and "REAL" statements at the beginning of each subroutine and main program, plus making the proper modifications in the set of nested "do loops" and statements to set up the "POINTER VECTOR" in the RISK subroutine.

Finally, the whole program was modified to allow the use of non-integer values for the parameters a_i and b_i and to work with different stratum sampling costs c_i. The entire revised program is listed in Appendix A. In Appendix B we give some selected examples of the program output.

2.4. Conclusions About the Computer Problem

1) It was found that with all these modifications, the program runs satisfactorily in all the two strata cases tested, as well as some isolated cases using up to four strata with different values for all the parameters.

2) It was noted that prior expected fraction defectives are very closely related to the computer time needed to solve a particular case.

Each case is defined by the strata constants a_i, b_i, c_i, λ_i, N_i and the values we arbitrarily choose for C1 and C2.

When the prior expected fraction defective $\frac{a_i}{a_i+b_i}$ of a stratum is increased the computer time needed to solve the case is also increased.

3) A comparison was made using the strata from Table 8.3 of [1] with the original and revised program. The computer time to complete the same table was reduced by 50%.

In page 139 of [1] a trial case with three strata is reported. Dr. Grosh said that after 10.19 minutes working, the computer completed only three output lines.

With the revised program we ran ten cases of a four strata problem in only 7.14 minutes. That is, we solved the total problem determining how we should split our total budget in the two stage sampling in order to get the minimum risk. (The results of this problem are given in Table 4.6). This gives us an idea about the time improvement obtained with the revised program

4) When Tables 8.3 page 146 of [1] (shown here as Table 2.1) and Table 2.2, which is the same problem but solved by the revised

program, are compared, we found that we obtained the same optimal first stage allocation and the difference in the risk function is not larger than four units in the fourth significant digit, which for practical purposes is assumed acceptable.

2.5 Notation used in the risk tables

C1 or C_1 = Total budget to be used in the first stage.

$\underline{n}^{(o)}$ = Allocation vector. The vector gives us the sample size of each stratum. Thus $\underline{n}^{o}$ = (25,20) means that the experimenter is to make 25 observations in the first stratum and 20 observations in the second stratum.

Risk = Risk value we obtained for the given allocation.

N_i or NS_i = Stratum size.

Table 8.3 Prior Bayes Risk x 10^3 for Various Sampling Schemes for Populations with Polya Prior Distributions ($a=3, b=7$). In All Cases $\lambda_1=1$, $c_1=c_2=1$, and $C_1+C_2=100$.

	$C_1 = 20$		$C_1 = 30$		$C_1 = 40$		$C_1 = 50$		$C_1 = 60$		$C_1 = 70$		$C_1 = 75$		$C_1 = 80$		$C_1 = 85$		$C_1 = 90$		$C_1 = 100$	
	WITH OPTIMAL FIRST STAGE ALLOCATION																					
λ_2	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk
1	(10,10)	5.8129	(15,15)	5.8051	(20,20)	5.7994	(25,25)	5.7949	(30,30)	5.7913	(35,35)	5.7884	(37,38)	5.7878	(40,40)	5.7881	(42,43)	5.7908	(45,45)	5.7974	(50,50)	5.8418
1.5	(10,10)	9.0359	(15,15)	9.0242	(20,20)	9.0156	(24,26)	9.0091	(26,34)	9.0041	(27,43)	9.0008	(28.47)	8.9997	(29,51)	9.0001	(31,54)	9.0033	(33,57)	9.0127	(38,62)	9.0791
2	(10,10)	12.091	(15,15)	12.070	(19,21)	12.864	(21,29)	12.857	(21,39)	12.852	(21,49)	12.848	(22,53)	12.846	(22,58)	12.847	(24,61)	12.850	(25,65)	12.861	(30,70)	12.949
3	(10,10)	22.502	(13,17)	22.479	(14,26)	22.465	(14,36)	22.457	(14,46)	22.451	(14,56)	22.447	(14,61)	22.445	(14,66)	22.444	(15,70)	22.445	(16,74)	22.455	(20,80)	22.588
5	(7,13)	49.324	(7,23)	49.296	(7,33)	49.279	(7,43)	49.269	(7,53)	49.261	(7,63)	49.255	(7,68)	49.252	(7,73)	49.250	(7,78)	49.248	(7,83)	49.252	(10,90)	49.461
9	(1,19)	133.39	(1,29)	133.35	(1,39)	133.33	(1,49)	133.31	(1,59)	133.30	(1,69)	133.29	(1,74)	133.29	(1,79)	133.28	(1,84)	133.28	(1,89)	133.28	(2,98)	133.58
	WITH PSEUDO-OPTIMAL FIRST STAGE ALLOCATION																					
λ_2	$^{(1)}\underline{n}^o$	Risk	$^{(1)}\underline{n}^o$	Risk	$^{(1)}\underline{n}^o$	Risk	$^{(1)}\underline{n}^o$	Risk	$^{(1)}\underline{n}^o$	Risk	$^{(1)}\underline{n}^o$	Risk	$^{(1)}\underline{n}^o$	Risk	$^{(1)}\underline{n}^o$	Risk	$^{(1)}\underline{n}^o$	Risk	$^{(1)}\underline{n}^o$	Risk		
1	(10,10)	5.8129	(15,15)	5.8051	(20,20)	5.7994	(25,25)	5.7949	(30,30)	5.7913	(35,35)	5.7884	(37,38)	5.7878✓	(40,40)	5.7881	(42,43)	5.7908	(45,45)	5.7974		
1.5	(6,14)	9.0377	(10,20)	9.0258	(14,26)	9.0171	(18,32)	9.0104	(22,38)	9.0050	(26,44)	9.0009	(28,47)	8.9997✓	(30,50)	9.0001	(32,52)	9.0037	(34,56)	9.0131		
2	(3,17)	12.899	(7,23)	12.882	(10,30)	12.871	(13,37)	12.862	(17,43)	12.854	(20,50)	12.849	(22,53)	12.847	(23,57)	12.847✓	(25,60)	12.851	(27,63)	12.864		
3	(0,20)	22.526	(2,26)	22.504	(5,35)	22.484	(7,43)	22.471	(10,50	22.458	(12,58)	22.449	(14,61)	22.445	(15,65)	22.444✓	(16,69)	22.447	(17,73)	22.457		
5	(0,20)	49.357	(0,30)	49.338	(0,40)	49.326	(2,48)	49.299	(3,57)	49.283	(5,65)	49.264	(6,69)	49.255	(7,73)	49.250	(7,78)	49.248✓	(8,82)	49.256		
9	(0,20)	133.40	(0,30)	133.37	(0,40)	133.34	(0,50)	133.33	(0,60)	133.33	(0.60)	133.31	(0,75)	133.30	(0,80)	133.30	(1,84)	133.28	(1,89)	133.28✓		

——— Best (lowest) optimal risk for any cost partitioning.
✓ Best (lowest) pseudo-optimal risk for any cost partitioning.
- - - Optimum allocation is the same as pseudo-optimum allocation.

TABLE 2.1

Reproduced from [1] with the Author's Permission.

Table 2.2 Prior Bayes Risk x 10^3 for Various Sampling Schemes for Populations with Beta-Binomial Prior Distributions. In all cases $\lambda_1 = 1$ and $NS_1 = 500$

$a_1 = 3.$ $b_1 = 7.$ $c_1 = 1.$

$a_2 = 3.$ $b_2 = 7.$ $c_2 = 1.$

	C1 = 20		C1 = 30		C1 = 40		C1 = 50		C1 = 60		C1 = 70		C1 = 75		C1 = 80		C1 = 85		C1 = 90		C1 = 100	
WITH OPTIMAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)																						
λ_2	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk
1	(10,10)	5.8131	(15,15)	5.8055	(20,20)	5.8000	(25,25)	5.7958	(30,30)	5.7925	(35,35)	5.7900	(37,38)	5.7895	(40,40)	5.7900	(42,43)	5.7929	(45,45)	5.7997	(50,50)	5.8418
1.5	(10,10)	9.0361	(15,15)	9.0247	(20,20)	9.0164	(25,25)	9.0101	(26,34)	9.0056	(27,43)	9.0026	(28,47)	9.0017	(29,51)	9.0022	(31,54)	9.0057	(33,57)	9.0154	(38,62)	9.0792
2	(10,10)	12.6921	(15,15)	12.8767	(19,21)	12.8658	(21,29)	12.8584	(21,39)	12.8534	(21,49)	12.8500	(22,53)	12.8489	(22,58)	12.8488	(24,61)	12.8522	(25,65)	12.8634	(30,70)	12.9194
3	(10,10)	22.5021	(13,17)	22.4802	(14,26)	22.4670	(14,36)	22.4586	(14,46)	22.4528	(14,56)	22.4489	(14,61)	22.4471	(14,66)	22.4462	(15,70)	22.4476	(16,74)	22.4579	(20,80)	22.5583
5	(7,13)	49.3243	(7,23)	49.2072	(7,33)	49.2811	(7,43)	49.2704	(7,53)	49.2629	(7,63)	49.2573	(7,68)	49.2549	(7,73)	49.2530	(7,78)	49.2517	(7,83)	49.2550	(10,90)	49.4607
9	(1,19)	133.3907	(1,29)	133.3552	(1,39)	133.3340	(1,49)	133.3188	(1,59)	133.3073	(1,69)	133.2988	(1,74)	133.2951	(1,79)	133.2920	(1,84)	133.2893	(1,89)	133.2870	(2,98)	133.5829
WITH INITIAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)																						
1	(10,10)	5.8131	(15,15)	5.8055	(20,20)	5.8000	(25,25)	5.7958	(30,30)	5.7925	(35,35)	5.7900	(37,38)	5.7895	(40,40)	5.7900	(42,43)	5.7929	(45,45)	5.7997	(50,50)	5.8418
1.5	(6,14)	9.0379	(10,20)	9.0262	(14,26)	9.0178	(18,32)	9.0115	(22,38)	9.0064	(26,44)	9.0026	(28,47)	9.0017	(30,50)	9.0024	(32,53)	9.0061	(34,56)	9.0159	(38,62)	9.0732
2	(3,17)	12.8992	(7,23)	12.8825	(10,30)	12.8714	(13,37)	12.8630	(17,43)	12.8554	(20,50)	12.8504	(22,53)	12.8489	(23,57)	12.8490	(25,60)	12.8534	(27,63)	12.8663	(30,70)	12.9494
3	(0,20)	22.5260	(2,28)	22.5044	(5,35)	22.4845	(7,43)	22.4720	(10,50)	22.4596	(12,58)	22.4513	(14,61)	22.4471	(15,65)	22.4467	(16,69)	22.4493	(17,73)	22.4603	(20,80)	22.5683
5	(0,20)	49.3575	(0,30)	49.3388	(0,40)	49.3269	(2,48)	49.3003	(3,57)	49.2852	(5,65)	49.2662	(6,69)	49.2581	(7,73)	49.2530	(7,78)	49.2517	(8,82)	49.2597	(10,90)	49.4607
9	(0,20)	133.4012	(0,30)	133.3682	(0,40)	133.3460	(0,50)	133.3317	(0,60)	133.3202	(0,70)	133.3119	(0,75)	133.3084	(0,80)	133.3052	(0,85)	133.3026	(1,89)	133.2870	(2,98)	133.5829

CHAPTER 3

MAIN RESULTS

3.1. Introduction

In the present chapter we are going to answer the remaining questions. What is the behavior of the two stage sampling procedure when the sampling costs are not necessarily the same for all the strata?

In order to study this part of the problem we chose five different strata which, in addition to the two strata studied by Grosh in [1], are summarized in Table 3.1.

As can be seen, the stratum size was maintained constant and equal to 200; this was done only for simplicity; in the general case they need not be the same.

In the present work it is assumed that the stratum sizes do not influence the final results, but only the total number of possible outcomes to be taken into account in each stratum in order to reach the pre-established value for the cumulative probability of defective units.

This assumption is based on the relatively small value of the ratios $\frac{a}{a+b}$ in each stratum. Besides when we reduce the stratum sizes we accelerate the process of building up the cumulative probability of defective units and at the same time we answer one of the questions which were raised.

At this time it is suitable to reflect upon the reasons for choosing the stratum constants a_i and b_i.

Strata Constants Table

Identification	a	b	Stratum size	Expected mean	Expected variance	Expected fraction defective P_i	Expected variance of P_i
H	1	1	500	250	20916.66	0.5	0.08366
L	3	7	500	150	4868.18	0.3	0.01947
M	0.1	9.9	200	2	37.8	.01	0.000945
N	0.1	4.9	200	4	133.93	.02	0.003348
P	10	90	200	20	53.46	.1	0.001336
Q	2	98	200	4	11.64	.02	0.000291
S	.2	9.8	200	4	74.83	.02	0.00187

Table 3.1

The combination of a_i, b_i and N_i giving us a Beta-Binomial probability function for which the probability does not decrease monotonically after reaching some maximum value as we move to the right (see examples marked with * in Figures 1.1 and 1.2) were ruled out. They have no practical application in the present problem.

Our concept about the stratum fraction defective is given by the ratio $\frac{a_i}{a_i+b_i} = E(P_i)$.

This $E(P_i)$ value along with the stratum size give us the expected mean value for the stratum

$$E(\tilde{M}_i | N_i, a_i, b_i) = \frac{a_i}{a_i+b_i} N_i \qquad \text{(See Equation 1.9)}$$

In the same way we can introduce our idea about the dispersion around the expected stratum mean, which is determined by the values of a and b the by the relationship

$$Var(\tilde{M}_i | N_i, a_i, b_i) = \frac{N_i \; a_i \; b_i \; (a_i+b_i+N_i)}{(a_i+b_i)^2 \; (a_i+b_i+1)} \qquad \text{(See Equation 1.10)}$$

For instance, see Table 3.1. Strata N, Q and S have the same $E(P_i)$ value, but the variance of each is different.

For a expected given fraction defective $\frac{a}{a+b}$ higher values of a and b are indication of smaller stratum variance, and indication of stronger prior belief.

3.2 The studied cases

Once the modifications in the computer program were done and tested with the results obtained by Grosh (as it was shown in Chapter 2) we proceeded as follow:

1) It was decided to continue the study with only two strata, for economic reasons.

2) It was considered that the number of cases studied for different values of the ratio λ_2/λ_1 when all the other constant values were maintained, should be reduced from 6 to 4 in order to save money and computer time.

3) It also was decided to study the cases of three different values for the ratio c_2/c_1 when all the remaining constant values are maintained.

4) In order to maintain a comparable set of solutions in the preliminary studies we decided to keep the same C1 set of values studied by Grosh in [1], and the total budget C1 + C2 = 100.

5) Finally we studied several cases with more than two strata. In these cases we were interested in knowing how to split our full budget in a two stage sampling scheme in order to obtain the minimum risk. Note that in these cases all the stratum constants were already fixed. Our decision variable was the C1 value with the proper combination of C2 in order to maintain our budget limitation. The obtained results of one of these problem is shown in Table 4.6.

In the two strata cases used as preliminary studies, the chosen values for the ratio λ_2/λ_1 were 1.,1.5,2. and 3. In all the cases for simplicity, λ_1 was maintained equal to 1.

The chosen values for the ratio c_2/c_1 were 1.,1.5 and 2.5. Again for simplicity $c_1 = 1$.

These values for λ_i and c_i, as well as those of a_i, b_i and N_i were chosen arbitrarily assuming that in general they represent some of the practical real life situations. At the same time, they could show us what is the tendency when the λ's and c's ratio varies. The obtained results are summarized in Tables 3.2 to 3.10.

3.3 Explanation of the tables

Each table has three principal parts. At the top is shown the obtained optimal first stage allocations and the corresponding risk.

In the middle of the table is shown the initial first stage allocations and their corresponding risk. This allocation is obtained from a direct application of the Equation (1.29) followed by a rounding off in detriment of the least important strata or stratum, when it was necessary in order to meet the budgetary constraint.

At the bottom is shown the "pseudo-optimal" allocation and corresponding risk.

Definition: Pseudo-optimal allocation is the allocation which would be optimal if there were a single stage problem with a total budget equal to C1. These values were obtained when we set C2 = 0.0 in each case. This is the type of solution that could be used by the investigator who does not have computing facilities, and must procede only on the basis of single stage formulae.

In these tables is interesting to note the following facts.

A) In general there is a difference between the risk values of the initial and the optimal first stage allocation. Optimal risk is less than or equal to initial risk. Consequently a person who does not have access to a computer may use for the first stage sampling vector that initial first stage allocation, knowing that the resulting risk value represents an upper bound.

B) In general the pseudo-optimal allocation is very close to the initial first stage allocation. The former starts with Equation (1.29) and then averages over all possible first stage outcomes. The latter also starts with Equation (1.29) and merely makes whatever adjustments need to be made to give integer value for the n_i without violating the cost constraint.

C) The risk values of the lowest and the upper two parts of the table are not comparable. The initial as well as the optimal allocations take in account for their risk evaluation the expected outcomes of the second stage. The pseudo-optimal by definition does not have a second stage. This conceptual difference is the explanation for the unequal behavior of the risk values between the bottom part of the table (where the risk decreases monotonically) and the other two parts (where the risk function is convex). It should be pointed out at this time in order to avoid misunderstanding that Grosh in her work at the bottom of her Table 8.3 (reproduced here as Table 2.1), used the term pseudo-optimal to designate the initial allocation, on the assumption that they were the same.

The current investigation has shown that this is not true in the more general cases treated here.

3.4 Ratio comparison studies

There we have two different problems to study. How do the risk and the optimal allocation change for different values of the ratio λ_2/λ_1? And how do they change when the ratio c_2/c_1 changes? In both cases the remaining parameter values were maintained constant.

The influence of different λ_2/λ_1 ratio values on the risk can be seen in Tables 3.2 to 3.10.

On the other hand we concluded that it was going to be more informative to the reader to show the influence of different c_2/c_1 ratio values on the risk in the same graph. A selection of them is shown in Figures 3.1 to 3.5.

3.5 Special Case Study

Due to the unexpected result of very small or no improvement at all when the single and double sampling schemes were compared when put together strata P and Q as they are defined in table 3.1 we decided to do a very brief comparison study when we combine other selected strata. The study was done only for the case when $c_2/c_1 = 1$ and for $\lambda_2/\lambda_1 = 1$ and 3.

We noted that in all the three cases in which we obtained such curious results, the (a+b) value of each stratum for each of the strata were equal. Besides, we noted that in the two examples worked by Grosh in [1] the expected fraction defective were equal for each stratum of the set.

In order to look a little bit more for the possible causes of those results we studied the pair of strata L and S, with the same (a_i+b_i) value but different $E(P_i)$ value. On the other hand we studied the pair of strata Q and S which represent the case of equal $E(P_i)$ value with different (a_i+b_i) value. The obtained results are summarized in Tables 3.11 and 3.12.

ILLEGIBLE DOCUMENT

THE FOLLOWING DOCUMENT(S) IS OF POOR LEGIBILITY IN THE ORIGINAL

THIS IS THE BEST COPY AVAILABLE

Table 3.2 Prior Bayes Risk x 10^3 for Various Sampling Schemes for Populations with Beta-binomial Prior Distributions. In all cases $\lambda_1 = 1$ and $NS_1 = 200$

$a_1 = 0.1 \quad b_1 = 9.9 \quad c_1 = 1.$

$a_2 = 0.1 \quad b_2 = 4.9 \quad c_2 = 1.$

	C1 = 20		C1 = 30		C1 = 40		C1 = 50	
λ_2	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk
WITH OPTIMAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)								
1	(9,11)	.2632	(12,18)	.2510	(15,25)	.2465	(17,33)	.2479
1.5	(7,13)	.4266	(8,22)	.4106	(9,31)	.4033	(10,40)	.4030
2	(5,15)	.6260	(6,24)	.6069	(6,34)	.5966	(6,44)	.5936
3	(2,18)	1.1322	(3,27)	1.1072	(3,37)	1.0915	(2,48)	1.0825
WITH INITIAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)								
1	(5,15)	.2651	(9,21)	.2520	(14,26)	.2465	(18,32)	.2479
1.5	(2,18)	.4338	(5,25)	.4134	(9,31)	.4033	(12,38)	.4037
2	(0,20)	.6404	(2,28)	.6140	(5,35)	.5972	(8,42)	.5947
3	(0,20)	1.1393	(0,30	1.1152	(1,39)	1.0944	(3,47)	1.0827
WITH PSEUDO-OPTIMAL ALLOCATION (C2 = 0.0)								
1	(5,15)	1.3885	(9,21)	1.0513	(14,26)	.8360	(18,32)	.6872
1.5	(2,18)	2.2700	(5,25)	1.7129	(9,31)	1.3591	(12,38)	1.1133
2	(0,20)	3.3558	(2,28)	2.5248	(5,35)	1.9954	(8,42)	1.6296
3	(0,20)	6.3693	(0,30)	4.6042	(1,39)	3.6115	(3,47)	2.9327

	C1 = 60		C1 = 70		C1 = 75		C1 = 80		C1 = 85		C1 = 90		C1 = 100	
λ_2	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk
WITH OPTIMAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)														
1	(19,41)	.2540	(23,47)	.2647	(26,49)	.2721	(28,52)	.2810	(31,54)	.2913	(34,56)	.3031	(40,60)	.3315
1.5	(10,50)	.4078	(13,57)	.4212	(15,60)	.4320	(18,62)	.4454	(20,65)	.4614	(23,67)	.4803	(29,71)	.5269
2	(6,54)	.5955	(8,62)	.6089	(9,66)	.6218	(11,69)	.6389	(13,72)	.6606	(16,74)	.6870	(22,78)	.7550
3	(2,58)	1.0779	(3,67)	1.0849	(3.72)	1.0983	(4,76)	1.1194	(6,79)	1.1499	(7,83)	1.1911	(13,87)	1.3045
WITH INITIAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)														
1	(22,38)	.2545	(27,43)	.2661	(29,46)	.2734	(31,49)	.2819	(33,52)	.2918	(35,55)	.3032	(40,60)	.3315
1.5	(15,45)	.4115	(19,51)	.4276	(20,55)	.4364	(22,58)	.4494	(24,61)	.4648	(25,65)	.4615	(29,71)	.5269
2	(11,49)	.6035	(13,57)	.6183	(15,60)	.6327	(16,64)	.6471	(18,67)	.6685	(19,71)	.6910	(22,78)	.7550
3	(5,55)	1.0864	(7,63)	1.1036	(8,67)	1.1180	(9,71)	1.1383	(10,75)	1.1654	(11,79)	1.2005	(13,87)	1.3045
WITH PSEUDO-OPTIMAL ALLOCATION (C2 = 0.0)														
1	(22,38)	.5782	(27,43)	.4947	(29,46)	.4599	(31,49)	.4288	(33,52)	.4008	(35,55)	.3755		
1.5	(15,45)	.9335	(19,51)	.7960	(20,55)	.7387	(22,58	.6873	(24,61)	.6413	(25,65)	.5995		
2	(11,49)	1.3615	(13,57)	1.1564	(15,60)	1.0708	(16,64)	.9943	(18,67)	.9256	(19,71)	.8632		
3	(5,55)	2.4349	(7,63)	2.0543	(8,67)	1.8956	(9,71)	1.7537	(10,75)	1.6260	(11,79)	1.5105		

Table 3.3 Prior Bayes Risk x 10^3 for Various Sampling Schemes for Populations with Beta-Binomial Prior Distributions. In all cases $\lambda_1 = 1$ and $SS_1 = 200$

$a_1 = 0.1$ $b_1 = 9.9$ $c_1 = 1.$

$a_2 = 0.1$ $b_2 = 4.9$ $c_2 = 1.5.$

λ_2	C1 = 20 $\underline{n}^o$	C1 = 20 Risk	C1 = 30 $\underline{n}^o$	C1 = 30 Risk	C1 = 40 $\underline{n}^o$	C1 = 40 Risk	C1 = 50 $\underline{n}^o$	C1 = 50 Risk	C1 = 60 $\underline{n}^o$	C1 = 60 Risk
WITH OPTIMAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)										
1	(3,10)	.3721	(6,[illegible])	.3627	(10,20)	.3521	(14,24)	.3521	(18,28)	.3588
1.5	(2,12)	.5842	([illegible],15)	[illegible]	(7,22)	.6178	(8,28)	.6158	(12,32)	.6243
2	(2,12)	.9389	(0,[illegible])	.9891	(4,24)	.9563	(5,30)	.9537	(6,36)	.9530
3	(2,12)	1.9059	(0,26)	1.8775	(1,26)	1.8313	(2,32)	1.8435	(3,38)	1.8449
WITH INITIAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)										
1	(3,10)	.3751	(7,18)	.3642	(11,19)	.3547	(15,23)	.3551	(19,27)	.3617
1.5	(0,13)	.6696	(4,17)	.6362	(7,22)	.6178	(9,27)	.6206	(12,32)	.6243
2	([illegible],13)	1.0164	(1,19)	.9895	(4,24)	.9563	(6,29)	.9602	(9,34)	.9729
3	([illegible],13)	1.9256	(0,20)	1.8775	(1,26)	1.8513	(2,32)	1.8435	(3,38)	1.8[illegible]
WITH PSEUDO-OPTIMAL ALLOCATION (C2 = 0.0)										
1	(3,10)	1.5745	(6,16)	1.3063	(10,20)	1.0516	(14,24)	.6742	(18,28)	.7434
1.5	(2,12)	2.8525	(3,16)	2.3064	(7,22)	1.7781	(8,28)	1.4857	(12,32)	1.2589
2	(0,13)	4.4235	(0,20)	3.3558	(4,24)	2.6935	(5,30)	2.2406	(9,34)	1.9002
3	(0,13)	8.7715	(0,20)	6.3593	(1,26)	5.0834	(2,32)	4.2004	(3,38)	3.5543

λ_2	C1 = 70 $\underline{n}^o$	C1 = 70 Risk	C1 = 75 $\underline{n}^o$	C1 = 75 Risk	C1 = 80 $\underline{n}^o$	C1 = 80 Risk	C1 = 85 $\underline{n}^o$	C1 = 85 Risk	C1 = 90 $\underline{n}^o$	C1 = 90 Risk	C1 = 100 $\underline{n}^o$	C1 = 100 Risk
WITH OPTIMAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)												
1	(22,32)	.3690	(24,34)	.3796	(36,26)	.3883	(28,38)	.3980	(30,40)	.4145	(34,44)	.4448
1.5	(13,38)	.6306	(12,42)	.6416	(17,42)	.6594	(19,44)	.6742	(21,46)	.7012	(25,50)	[illegible]
2	(7,42)	.9566	(9,44)	.9762	([illegible],46)	.9938	(13,48)	1.0136	(15,50)	1.0327	(16,56)	[illegible]
3	(1,46)	1.8195	(3,48)	1.8435	(2,52)	1.8534	(4,54)	1.8775	(6,56)	1.9403	(10,60)	2.0715
WITH INITIAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)												
1	(23,31)	.3721	(25,33)	.3829	(27,35)	.3916	(29,37)	.4011	(31,39)	.4175	(35,43)	[illegible]
1.5	(16,36)	.6364	(16,39)	.6537	(16,41)	.6661	(20,43)	.6806	(21,46)	.7012	([illegible],50)	.7499
2	(10,40)	.9683	(12,42)	.9912	(12,45)	1.0043	(14,47)	1.0240	(15,50)	1.0527	(17,55)	1.1294
3	(5,43)	1.8586	(6,46)	1.8754	(6,49)	1.8907	(7,52)	1.8973	([illegible],54)	[illegible]	(10,60)	2.0715
WITH PSEUDO-OPTIMAL ALLOCATION (C2 = 0.0)												
1	(22,32)	.6429	(24,34)	.6009	(26,26)	.5632	(28,38)	.5292	(30,40)	[illegible]		
1.5	(16,36)	1.0878	(15,40)	1.0193	(17,42)	.9534	(19,44)	.8945	(21,46)	.8416		
2	(10,40)	1.6394	(12,42)	1.5204	(11,46)	1.4363	(13,48)	1.3444	(15,50)	1.2628		
3	(4,44)	3.0600	(6,46)	2.6478	(5,50)	2.6689	(7,52)	2.4925	(9,54)	2.3303		

Table 3.4 Prior Bayes Risk x 10^3 for *Various Sampling* Schemes for Populations with Beta-Binomial Prior Distributions. In all cases $\lambda_1 = 1$ and $NS_1 = 200$

$a_1 = 0.1 \cdot b_1 = 9.9 \quad c_1 = 1.$

$a_2 = 0.1 \quad b_2 = 4.9 \quad c_2 = 2.5$

	C1 = 20		C1 = 30		C1 = 40		C1 = 50		C1 = 60		C1 = 70		C1 = 75		C1 = 80		C1 = 85		C1 = 90		C1 = 100	
λ_2	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk
WITH OPTIMAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)																						
1	(5,6)	.5659	(5,10)	.5524	(10,12)	.5366	(10,16)	.5344	(15,18)	.5396	(15,22)	.5479	(15,24)	.5553	(20,24)	.5688	(20,26)	.5800	(25,26)	.5974	(30,28)	0.6371
1.5	(5,6)	1.0383	(5,10)	1.0167	(5,14)	1.0011	(5,18)	.9958	(5,22)	.9928	(5,26)	.9964	(5,28)	1.0060	(10,28)	1.0204	(10,30)	1.0397	(15,30)	1.0686	(20,32)	1.1367
2	(5,6)	1.6545	(5,10)	1.6250	(5,14)	1.6099	(5,18)	1.5949	(5,22)	1.5917	(0,28)	1.5900	(5,28)	1.5952	(5,30)	1.6050	(5,32)	1.6285	(10,32)	1.6704	(15,34)	1.7748
3	(0,8)	3.2932	(0,12)	3.2630	(0,16)	3.2375	(0,20)	3.2144	(0,24)	3.2021	(0,28)	3.1894	(0,30)	3.1798	(0,32)	3.1819	(0,34)	3.2060	(0,36)	3.2589	(5,33)	3.4525
WITH INITIAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)																						
1	(5,6)	.5659	(7,9)	.5529	(10,12)	.5366	(12,15)	.5370	(17,17)	.5447	(20,20)	.5531	(22,21)	.5660	(22,23)	.5735	(25,24)	.5848	(27,25)	.6035	(30,28)	.6371
1.5	(0,8)	1.0555	(2,11)	1.0317	(5,14)	1.0011	(7,17)	1.0003	(10,20)	.9978	(12,23)	1.0173	(12,25)	1.0236	(15,26)	1.0373	(15,28)	1.0491	(17,29)	1.0702	(20,32)	1.1367
2	(0,8)	1.6685	(0,12)	1.6374	(2,15)	1.6213	(5,18)	1.5949	(5,22)	1.5917	(7,25)	1.6097	(10,26)	1.6267	(10,28)	1.6346	(10,30)	1.6466	(12,31)	1.6904	(12,35)	1.7850
3	(0,8)	3.2932	(0,12)	3.2630	(0,16)	3.2375	(0,20)	3.2144	(2,23)	3.2250	(2,27)	3.2146	(2,29)	3.2129	(5,30)	3.2447	(5,32)	3.2530	(5,34)	3.2774	(7,37)	3.4602
WITH PSEUDO-OPTIMAL ALLOCATION (C2 = 0.0)																						
1	(5,6)	2.0906	(5,10)	1.6745	(10,12)	1.3746	(15,14)	1.1691	(15,18)	1.0120	(20,20)	.8862	(20,22)	.8354	(20,24)	.7915	(25,24)	.7443	(25,26)	.7061		
1.5	(0,8)	3.7267	(5,10)	2.9999	(5,14)	2.4560	(10,16)	2.0391	(10,20)	1.8049	(10,24)	1.5919	(15,24)	1.4927	(15,26)	1.4068	(15,28)	1.3313	(15,30)	1.2645		
2	(0,8)	5.8902	(0,12)	4.6479	(2,15)	3.8768	(5,18)	3.2638	(5,22)	2.8217	(10,24)	2.4810	(10,26)	2.3283	(10,28)	2.1941	(10,30)	2.0752	(10,32)	1.9692		
3	(0,8)	12.0718	(0,12)	9.2764	(0,16)	7.5460	(0,20)	6.3093	(0,24)	5.5172	(5,26)	4.8429	(5,28)	4.5409	(5,30)	4.2735	(5,32)	4.0350	(5,34)	3.8209		

Table 3.5 Prior Bayes Risk x 10^3 for Various Sampling Schemes for Populations with Beta-Binomial Prior Distributions. In all cases $\lambda_1 = 1$ and $NS_1 = 200$

$a_1 = 10.$ $b_1 = 90.$ $c_1 = 1.$

$a_2 = 2.$ $b_2 = 98.$ $c_2 = 1.$

	C1 = 20		C1 = 30		C1 = 40		C1 = 50		C1 = 60		C1 = 70		C1 = 75		C1 = 80		C1 = 85		C1 = 90		C1 = 100	
λ_2	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk
WITH OPTIMAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)																						
1	(20,0)	.6252	(30,0)	.6252	(40,0)	.6252	(50,0)	.6251	(60,0)	.6250	(70,0)	.6250	(75,0)	.6250	(80,0)	.6250	(85,0)	.6249	(90,0)	.6249	(100,0)	.6252
1.5	(1,19)	.9236	(11,19)	.9283	(22,18)	.9280	(32,18)	.9275	(42,18)	.9274	(52,18)	.9272	(57,18)	.9272	(62,18)	.9276	(66,19)	.9283	(70,20)	.9207	(76,24)	.9357
2	(0,20)	1.2376	(0,30)	1.2336	(4,36)	1.2316	(15,35)	1.2310	(25,35)	1.2306	(34,36)	1.2306	(39,36)	1.2312	(42,38)	1.2324	(46,39)	1.2344	(49,41)	1.2374	(55,45)	1.2475
3	(0,20)	1.8564	(0,30)	1.8505	(0,40)	1.8454	(0,50)	1.8409	(0,60)	1.8371	(8,62)	1.8380	(11,64)	1.8398	(14,66)	1.8427	(17,68)	1.8471	(19,71)	1.8532	(25,75)	1.8713
WITH INITIAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)																						
1	(20,0)	.6252	(30,0)	.6252	(40,0)	.6252	(50,0)	.6251	(60,0)	.6250	(70,0)	.6250	(75,0)	.6250	(80,0)	.6250	(85,0)	.6249	(90,0)	.6249	(100,0)	.6252
1.5	(20,0)	.9343	(30,0)	.9338	(40,0)	.9334	(47,3)	.9318	(53,7)	.9300	(59,11)	.9286	(62,13)	.9281	(65,15)	.9280	(68,17)	.9285	(71,19)	.9298	(76,24)	.9357
2	(14,6)	1.2429	(19,11)	1.2400	(24,16)	1.2373	(29,21)	1.2349	(34,26)	1.2328	(40,30)	1.2320	(42,33)	1.2318	(45,35)	1.2328	(47,38)	1.2344	(50,40)	1.2375	(55,45)	1.2475
3	(0,20)	1.8564	(0,30)	1.8505	(0,40)	1.8454	(4,46)	1.8425	(8,52)	1.8403	(12,58)	1.8397	(15,60)	1.8414	(17,63)	1.8437	(19,66)	1.8477	(21,69)	1.8535	(25,75)	1.8713
WITH PSEUDO-OPTIMAL ALLOCATION (C2 = 0.0)																						
1	(20,0)	1.2936	(30,0)	1.1650	(40,0)	1.0549	(50,0)	.9594	(60,0)	.8759	(70,0)	.8022	(75,0)	.7685	(80,0)	.7366	(85,0)	.7065	(90,0)	.6780		
1.5	(20,0)	1.6574	(30,0)	1.5289	(40,0)	1.4187	(47,3)	1.3210	(53,7)	1.2328	(59,11)	1.1503	(62,13)	1.1112	(65,15)	1.0736	(68,17)	1.0373	(71,19)	1.0023		
2	(14,6)	2.1559	(19,11)	2.0078	(24,16)	1.8720	(29,21)	1.7471	(34,26)	1.6319	(40,30)	1.5251	(42,33)	1.4746	(45,35)	1.4260	(47,38)	1.3790	(50,40)	1.3337		
3	(0,20)	3.3015	(0,30)	3.0496	(0,40)	2.8336	(4,46)	2.6412	(8,52)	2.4635	(13,57)	2.2991	(15,60)	2.2213	(17,63)	2.1463	(19,66)	2.0739	(21,69)	2.0040		

Table 3.6 Prior Bayes Risk x 10^5 for Various Sampling Schemes for Populations with Beta-Binomial Prior Distributions. In all cases $\lambda_1 = 1$ and $NS_1 = 200$

$a_1 = 10. \quad b_1 = 90. \quad c_1 = 1.$

$a_2 = 2. \quad b_2 = 98. \quad c_2 = 1.5.$

λ_2	C1 = 20 $\underline{n}^o$	C1 = 20 Risk	C1 = 30 $\underline{n}^o$	C1 = 30 Risk	C1 = 40 $\underline{n}^o$	C1 = 40 Risk	C1 = 50 $\underline{n}^o$	C1 = 50 Risk
WITH OPTIMAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)								
1	(20,0)	.6252	(30,0)	.6252	(40,0)	.6252	(50,0)	.6252
1.5	(20,0)	.9809	(30,0)	.9798	(40,0)	.9800	(50,0)	.9802
2	(20,0)	1.3856	(30,0)	1.3848	(34,4)	1.3783	(41,6)	1.3789
3	(2,12)	2.2387	(3,18)	2.2300	(7,22)	2.2276	(11,26)	2.2262
WITH INITIAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)								
1	(20,0)	.6252	(30,0)	.6252	(40,0)	.6252	(50,0)	.6252
1.5	(20,0)	.9809	(30,0)	.9798	(40,0)	.9800	(50,0)	.9802
2	(20,0)	1.3856	(30,0)	1.3848	(35,3)	1.3824	(40,6)	1.3865
3	(0,13)	2.2433	(3,18)	2.2300	(7,22)	2.2276	(11,26)	2.2262
WITH PSEUDO-OPTIMAL ALLOCATION (C2 = 0.0)								
1	(20,0)	1.2936	(30,0)	1.1650	(40,0)	1.0549	(50,0)	.9694
1.5	(20,0)	1.6574	(30,0)	1.5289	(40,0)	1.4187	(50,0)	1.3253
2	(20,0)	2.1668	(30,0)	2.0383	(34,4)	1.9251	(41,6)	1.8191
3	(2,12)	3.4961	(3,18)	3.2956	(7,22)	3.1166	(11,26)	2.9468

λ_2	C1 = 60 $\underline{n}^o$	C1 = 60 Risk	C1 = 70 $\underline{n}^o$	C1 = 70 Risk	C1 = 75 $\underline{n}^o$	C1 = 75 Risk	C1 = 80 $\underline{n}^o$	C1 = 80 Risk	C1 = 85 $\underline{n}^o$	C1 = 85 Risk	C1 = 90 $\underline{n}^o$	C1 = 90 Risk	C1 = 100 $\underline{n}^o$	C1 = 100 Risk
WITH OPTIMAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)														
1	(60,0)	.6252	(70,0)	.6252	(75,0)	.6252	(80,0)	.6252	(85,0)	.6252	(90,0)	.6252	(100,0)	.6252
1.5	(60,0)	.9798	(70,0)	.9801	(75,0)	.9791	(77,2)	.9786	(79,4)	.9776	(84,4)	.9798	(88,8)	.9803
2	(45,10)	1.3774	(49,14)	1.3727	(51,16)	1.3736	(53,18)	1.3723	(55,20)	1.3716	(60,20)	1.3788	(64,24)	1.3805
3	(12,32)	2.2231	(16,36)	2.2211	(18,38)	2.2220	(20,40)	2.2245	(22,42)	2.2253	(24,44)	2.2305	(28,48)	2.2433
WITH INITIAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)														
1	(60,0)	.6252	(70,0)	.6252	(75,0)	.6252	(80,0)	.6252	(85,0)	.6252	(90,0)	.6252	(100,0)	.6252
1.5	(60,0)	.9798	(70,0)	.9801	(75,0)	.9791	(78,1)	.9625	(80,3)	.9813	(83,4)	.9854	(88,8)	.9803
2	(45,10)	1.3774	(49,14)	1.3727	(52,15)	1.3785	(54,17)	1.3769	(55,19)	1.3767	(59,20)	1.3841	(63,24)	1.3860
3	(13,31)	2.2286	(17,35)	2.2277	(19,37)	2.2281	(21,39)	2.2300	(23,41)	2.2311	(25,43)	2.2369	(29,47)	2.2493
WITH PSEUDO-OPTIMAL ALLOCATION (C2 = 0.0)														
1	(60,0)	.8759	(70,0)	.8021	(75,0)	.7685	(80,0)	.7366	(85,0)	.7065	(90,0)	.6780		
1.5	(60,0)	1.2397	(70,0)	1.1660	(75,0)	1.1323	(77,2)	1.1001	(79,4)	1.0689	(84,4)	1.0385		
2	(45,10)	1.7200	(49,14)	1.6271	(51,16)	1.5829	(53,18)	1.5400	(55,20)	1.4985	(60,20)	1.4580		
3	(12,32)	2.7890	(16,36)	2.6397	(18,38)	2.5685	(20,40)	2.4995	(22,42)	2.4326	(24,44)	2.3676		

Table 3.7 Prior Bayes Risk x 10^3 for Various Sampling Schemes for Populations with Beta-Binomial Prior Distributions. In all cases $\lambda_1 = 1$ and $NS_1 = 200$

$a_1 = 10.$ $b_1 = 90.$ $c_1 = 1.$

$a_2 = 2.$ $b_2 = 98.$ $c_2 = 2.5$

	C1 = 20		C1 = 30		C1 = 40		C1 = 50		C1 = 60		C1 = 70		C1 = 75		C1 = 80		C1 = 85		C1 = 90		C1 = 100	
λ_2	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk
WITH OPTIMAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)																						
1	(20,0)	.6252	(30,0)	.6252	(40,0)	.6252	(50,0)	.6252	(60,0)	.6252	(70,0)	.6252	(75,0)	.6252	(80,0)	.6252	(85,0)	.6252	(90,0)	.6252	(100,0)	.6252
1.5	(20,0)	.9891	(30,0)	.9891	(40,0)	.9891	(50,0)	.9896	(60,0)	.9891	(70,0)	.9893	(75,0)	.9890	(80,0)	.9893	(85,0)	.9893	(90,0)	.9895	(100,0)	.9891
2	(20,0)	1.4871	(30,0)	1.4831	(40,0)	1.4833	(50,0)	1.4820	(60,0)	1.4815	(70,0)	1.4803	(72,1)	1.4810	(75,2)	1.4792	(77,3)	1.4813	(80,4)	1.4789	(82,7)	1.4634
3	(15,2)	2.6301	(20,4)	2.6260	(22,7)	2.6257	(25,10)	2.6221	(30,12)	2.6200	(32,15)	2.6212	(32,17)	2.6222	(35,18)	2.6189	(35,20)	2.6139	(37,21)	2.6212	(40,24)	2.6230
WITH INITIAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)																						
1	(20,0)	.6252	(30,0)	.6252	(40,0)	.6252	(50,0)	.6252	(60,0)	.6252	(70,0)	.6252	(75,0)	.6252	(80,0)	.6252	(85,0)	.6252	(90,0)	.6252	(100,0)	.6252
1.5	(20,0)	.9891	(30,0)	.9891	(40,0)	.9891	(50,0)	.9896	(60,0)	.9891	(70,0)	.9893	(75,0)	.9890	(80,0)	.9893	(85,0)	.9893	(90,0)	.9895	(100,0)	.9891
2	(20,0)	1.4871	(30,0)	1.4831	(40,0)	1.4833	(50,0)	1.4820	(60,0)	1.4815	(70,0)	1.4803	(72,1)	1.4810	(74,2)	1.4858	(76,3)	1.4885	(78,4)	1.4925	(82,7)	1.4834
3	(15,2)	2.6351	(18,4)	2.6453	(21,7)	2.6425	(24,10)	2.6296	(28,12)	2.6407	(31,15)	2.6349	(32,17)	2.6222	(34,18)	2.6264	(35,20)	2.6139	(37,21)	2.6212	(40,24)	2.6730
WITH PSEUDO-OPTIMAL ALLOCATION (C2 = 0.0)																						
1	(20,0)	1.2935	(30,0)	1.1650	(40,0)	1.0549	(50,0)	.9594	(60,0)	.8759	(70,0)	.8021	(75,0)	.7684	(80,0)	.7366	(85,0)	.7065	(90,0)	.6760		
1.5	(20,0)	1.6374	(30,0)	1.5289	(40,0)	1.4187	(50,0)	1.3233	(60,0)	1.2397	(70,0)	1.1660	(75,0)	1.1373	(80,0)	1.1005	(85,0)	1.0704	(90,0)	1.0419		
2	(20,0)	2.1668	(30,0)	2.0383	(40,0)	1.9281	(50,0)	1.8327	(60,0)	1.7491	(70,0)	1.6754	(72,1)	1.6444	(75,2)	1.6075	(77,3)	1.5779	(80,4)	1.5427		
3	(15,2)	3.6178	(20,4)	3.4711	(22,7)	3.3378	(25,10)	3.1982	(30,12)	3.0727	(32,15)	2.9578	(32,17)	2.8994	(35,18)	2.8372	(35,20)	2.7817	(37,21)	2.7329		

Table 3.8 Prior Bayes Risk x 10^3 for Various Sampling Schemes for Populations with Beta-Binomial Prior Distributions. In all cases $\lambda_1 = 1$ and $NS_1 = 200$

$a_1 = 10 \quad b_1 = 90 \quad c_1 = 1.$

$a_2 = .2 \quad b_2 = 9.8 \quad c_2 = 1.$

	C1 = 20		C1 = 30		C1 = 40		C1 = 50		C1 = 60		C1 = 70		C1 = 75		C1 = 80		C1 = 85		C1 = 90		C1 = 100	
	WITH OPTIMAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)																					
λ_2	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk
1	(0,20)	.7923	(5,25)	.7850	(15,25)	.7851	(24,26)	.7865	(34,26)	.7907	(42,28)	.8002	(47,28)	.8076	(50,30)	.8174	(53,32)	.8299	(56,34)	.8452	(60,40)	.8841
1.5	(0,20)	1.0465	(0,30)	1.0206	(5,35)	1.0171	(13,37)	1.0240	(22,38)	1.0372	(29,41)	1.0588	(32,43)	1.0732	(35,45)	1.0902	(37,48)	1.1097	(39,51)	1.1317	(43,57)	1.1830
2	(0,20)	1.3104	(0,30)	1.2774	(0,40)	1.2555	(3,47)	1.2646	(10,50)	1.2877	(16,54)	1.3208	(19,56)	1.3413	(21,59)	1.3642	(23,62)	1.3897	(25,65)	1.4177	(29,71)	1.4818
3	(0,20)	1.9029	(0,30)	1.8596	(0,40)	1.8303	(0,50)	1.8081	(0,60)	1.8112	(0,70)	1.8481	(0,75)	1.8756	(0,80)	1.9082	(1,84)	1.9453	(4,86)	1.9560	(6,92)	2.0795
	WITH INITIAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)																					
1	(0,20)	.7923	(7,23)	.7863	(14,26)	.7853	(22,28)	.7880	(29,31)	.7959	(37,33)	.8063	(41,34)	.8136	(45,35)	.8225	(49,36)	.8335	(52,38)	.8483	(60,40)	.8841
1.5	(0,20)	1.0465	(0,30)	1.0206	(2,38)	1.0187	(9,41)	1.0282	(16,44)	1.0433	(22,48)	1.0669	(26,49)	1.0792	(29,51)	1.0956	(33,52)	1.1126	(36,54)	1.1355	(43,57)	1.1850
2	(0,20)	1.3104	(0,30)	1.2774	(0,40)	1.2555	(0,50)	1.2667	(5,55)	1.2928	(11,59)	1.3256	(14,61)	1.3452	(17,63)	1.3670	(20,65)	1.3914	(23,67)	1.4185	(29,71)	1.4818
3	(0,20)	1.9029	(0,30)	1.8596	(0,40)	1.8303	(0,50)	1.8081	(0,60)	1.8112	(0,70)	1.8481	(0,75)	1.8756	(0,80)	1.9082	(1,84)	1.9453	(3,87)	1.9863	(6,92)	2.0795
	WITH PSEUDO-OPTIMAL ALLOCATION (C2 = 0.0)																					
1	(0,20)	1.8979	(7,23)	1.7072	(14,26)	1.5425	(22,28)	1.3985	(29,31)	1.2715	(37,33	1.1584	(41,34)	1.1065	(45,35)	1.0574	(49,36)	1.0108	(52,38)	.9064		
1.5	(0,20)	2.5995	(0,30)	2.2312	(2,38)	2.0077	(9,41)	1.8273	(16,44)	1.6681	(22,48)	1.5267	(26,49)	1.4616	(29,51)	1.4000	(33,52)	1.3416	(36,54)	1.2861		
2	(0,20)	3.5817	(0,30)	2.9269	(0,40)	2.5340	(0,50)	2.2721	(5,55)	2.0759	(11,59)	1.9026	(14,61)	1.8230	(17,63)	1.7475	(20,65)	1.6760	(23,67)	1.6080		
3	(0,20)	6.3881	(0,30)	4.9147	(0,40)	4.0367	(0,50)	3.4414	(0,60)	3.0205	(0,70)	2.7047	(0,75)	2.5747	(0,80)	2.4592	(1,84)	2.3557	(3,87)	2.2590		

Table 3.9 Prior Bayes Risk x 10^3 for Various Sampling Schemes for Populations with Beta-Binomial Prior Distributions. In all cases $\lambda_1 = 1$ and $NS_1 = 200$

$a_1 = 10$ $b_1 = 90$ $c_1 = 1.$

$a_2 = .2$ $b_2 = 9.8$ $c_2 = 1.5.$

	C1 = 20		C1 = 30		C1 = 40		C1 = 50		C1 = 60		C1 = 70		C1 = 75		C1 = 80		C1 = 85		C1 = 90		C1 = 100	
	WITH OPTIMAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)																					
λ_2	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk
1	(2,12)	.9396	(3,18)	.9192	(10,20)	.9172	(20,20)	.9196	(27,22)	.9258	(34,24)	.9376	(36,26)	.9502	(41,26)	.9586	(43,28)	.9718	(48,28)	.9874	(55,30)	1.0228
1.5	(0,13)	1.3165	(0,20)	1.2800	(1,26)	1.2568	(5,30)	1.2645	(9,3[illegible])	1.2849	(16,36)	1.3071	(21,36)	1.3226	(23,38)	1.3417	(25,40)	1.3612	(30,40)	1.3854	(37,42)	1.4347
2	(0,13)	1.7283	(0,20)	1.6869	(1,26)	1.6615	(2,32)	1.6457	(0,40)	1.6587	(4,44)	1.6936	(6,46)	1.7211	(11,46)	1.7435	(13,48)	1.7682	(15,50)	1.8052	(22,52)	1.8683
3	(0,13)	2.7272	(0,20)	2.6796	(1,26)	2.6446	(0,33)	2.6301	(0,40)	2.5946	(1,46)	2.5967	(0,50)	2.6163	(0,53)	2.6497	(1,56)	2.6671	(0,60)	2.7145	(1,66)	2.8012
	WITH INITIAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)																					
1	(0,13)	.9420	(4,17)	.9272	(11,19)	.9217	(20,20)	.9196	(27,22)	.9258	(34,24)	.9376	(37,25)	.9519	(41,26)	.9586	(44,27)	.9736	(48,28)	.9874	(55,30)	1.0228
1.5	(0,13)	1.3165	(0,20)	1.2800	(1,26)	1.2568	(5,30)	1.2645	(10,33)	1.2875	(17,35)	1.3095	(21,36)	1.3226	(24,37)	1.3450	(26,39)	1.3641	(30,40)	1.3654	(37,42)	1.4347
2	(0,13)	1.7283	(0,20)	1.6869	(1,26)	1.6615	(0,33)	1.6460	(0,40)	1.6587	(4,44)	1.6936	(7,45)	1.7263	(11,46)	1.7435	(13,48)	1.7682	(16,49)	1.8106	(22,52)	1.8683
3	(0,13)	2.7272	(0,20)	2.6796	(1,26)	2.6446	(0,33)	2.6301	(0,40)	2.5946	(1,46)	2.5967	(0,50)	2.6163	(0,53)	2.6497	(1,56)	2.6671	(0,60)	2.7145	(1,66)	2.8012
	WITH PSEUDO-OPTIMAL ALLOCATION (C2 = 0.0)																					
1	(2,12)	2.0967	(3,18)	1.8863	(10,20)	1.7156	(20,20)	1.5637	(27,22)	1.4307	(34,24)	1.3121	(36,26)	1.2580	(41,26)	1.2058	(43,28)	1.1572	(48,28)	1.1098		
1.5	(0,13)	3.0479	(0,20)	2.5995	(1,26)	2.3341	(5,30)	2.1357	(9,34)	1.9652	(16,36)	1.8105	(21,36)	1.7391	(23,38)	1.6721	(25,40)	1.6092	(30,40)	1.5475		
2	(0,13)	4.3789	(0,20)	3.5817	(1,26)	3.1253	(0,33)	2.7808	(0,40)	2.5340	(4,44)	2.3405	(6,46)	2.2522	(11,46)	2.1669	(13,48)	2.0866	(15,50)	2.0106		
3	(0,13)	8.1817	(0,20)	6.3881	(1,26)	5.3860	(0,33)	4.6064	(0,40)	4.0307	(1,46)	3.6320	(0,50)	3.4414	(0,53)	3.3011	(1,56)	3.1537	(0,60)	3.0205		

Table 3.10 Prior Bayes Risk x 10^3 for Various Sampling Schemes for Populations with Beta-Binomial Prior Distributions. In all cases $\lambda_1 = 1$ and $NS_1 = 200$

$a_1 = 10 \quad b_1 = 90 \quad c_1 = 1.$

$a_2 = .2 \quad b_2 = 9.8 \quad c_2 = 2.5.$

	C1 = 20		C1 = 30		C1 = 40		C1 = 50		C1 = 60		C1 = 70		C1 = 75		C1 = 80		C1 = 85		C1 = 90		C1 = 100	
λ_2	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk
	WITH OPTIMAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)																					
1	(0,8)	1.1539	(0,12)	1.1305	(5,14)	1.1204	(15,14)	1.1235	(25,14)	1.1308	(30,16)	1.1410	(35,16)	1.1497	(35,18)	1.1625	(40,18)	1.1745	(45,18)	1.1899	(50,20)	1.2296
1.5	(0,8)	1.7210	(0,12)	1.6886	(0,16)	1.6635	(0,20)	1.6453	(5,22)	1.6503	(10,24)	1.6856	(15,24)	1.7007	(15,26)	1.7278	(20,46)	1.7426	(20,28)	1.7709	(30,28)	1.6266
2	(0,8)	2.3861	(0,12)	2.3482	(0,16)	2.3179	(0,20)	2.2897	(0,24)	2.2724	(0,26)	2.2908	(0,30)	2.3133	(0,32)	2.3445	(5,32)	2.3717	(5,34)	2.4090	(10,36)	2.4884
3	(0,8)	4.0224	(0,12)	4.0023	(0,16)	3.9794	(0,20)	3.9576	(0,24)	3.9422	(0,28)	3.9185	(0,30)	3.9104	(0,32)	3.9124	(0,34)	3.9263	(0,36)	3.9511	(0,40)	4.0307
	WITH INITIAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)																					
1	(0,8)	1.1539	(2,11)	1.1389	(10,12)	1.1291	(17,13)	1.1315	(25,14)	1.1308	(30,16)	1.1410	(35,16)	1.1497	(37,17)	1.1651	(40,18)	1.1745	(45,18)	1.1399	(50,20)	1.2296
1.5	(0,8)	1.7210	(0,12)	1.6886	(0,16)	1.6635	(0,20)	1.6453	(5,22)	1.6503	(12,23)	1.6898	(15,24)	1.7007	(17,25)	1.7260	(20,46)	1.7426	(22,27)	1.7738	(30,28)	1.8266
2	(0,8)	2.3861	(0,12)	2.3482	(0,16)	2.3179	(0,20)	2.2897	(0,24)	2.2724	(0,28)	2.2908	(0,30)	2.3133	(2,31)	2.3482	(5,32)	2.3717	(7,33)	2.4139	(12,35)	2.4938
3	(0,8)	4.0224	(0,12)	4.0023	(0,16)	3.9794	(0,20)	3.9576	(0,24)	3.9422	(0,28)	3.9185	(0,30)	3.9104	(0,32)	3.9124	(0,34)	3.9263	(0,36)	3.9511	(0,40)	4.0307
	WITH PSEUDO-OPTIMAL ALLOCATION (C2 = 0.0)																					
1	(0,8)	2.3344	(0,12)	2.1360	(10,12)	1.9537	(15,14)	1.8001	(25,14)	1.6606	(30,16)	1.5360	(35,16)	1.4788	(35,18)	1.4249	(40,18)	1.3718	(45,18)	1.3224		
1.5	(0,8)	3.5817	(0,12)	3.1353	(0,16)	2.8261	(0,20)	2.5995	(5,22)	2.4119	(10,24)	2.2439	(15,24)	2.1646	(15,26)	2.0924	(20,26)	2.0198	(20,28)	1.9552		
2	(0,8)	5.3279	(0,12)	4.5342	(0,16)	3.9847	(0,20)	3.5817	(0,24)	3.2736	(0,28)	3.0303	(0,30)	2.9269	(0,32)	2.8334	(5,32)	2.7379	(5,34)	2.6528		
3	(0,8)	10.3170	(0,12)	8.5311	(0,16)	7.2947	(0,20)	6.3881	(0,24)	5.6947	(0,28)	5.1474	(0,30)	4.9147	(0,32)	4.7043	(0,34)	4.5129	(0,36)	4.3382		

Table 3.11 Prior Bayes Risk x 10^3 for Various Sampling Schemes for Populations with Beta-Binomial Prior Distributions. In all cases $\lambda_1 = 1$ and $NS_1 = 200$

$a_1 = 3.$ $b_1 = 7.$ $c_1 = 1.$

$a_2 = .2$ $b_2 = 9.8$ $c_2 = 1.$

	C1 = 20		C1 = 30		C1 = 40		C1 = 50		C1 = 60		C1 = 70		C1 = 75		C1 = 80		C1 = 85		C1 = 90		C1 = 100	
									WITH OPTIMAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)													
λ_2	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk
1	(7,13)	1.7660	(17,13)	1.7648	(27,13)	1.7640	(37,13)	1.7635	(47,13)	1.7632	(57,13)	1.7639	(62,13)	1.7661	(67,13)	1.7718	(71,14)	1.7337	(75,15)	1.8053	(82,18)	1.8936
3	(0,20)	4.1046	(0,30)	4.0116	(9,31)	4.0021	(19,31)	3.9982	(28,32)	4.0048	(36,34)	4.0493	(39,36)	4.0937	(42,38)	4.1557	(45,40)	4.2363	(48,42)	4.3364	(53,47)	4.5985

Table 3.12 Prior Bayes Risk x 10^3 for Various Sampling Schemes for Populations with Beta-Binomial Prior Distributions. In all cases $\lambda_1 = 1$ and $NS_1 = 200$

$a_1 = 2.$ $b_1 = 98.$ $c_1 = 1.$

$a_2 = .2$ $b_2 = 9.8$ $c_2 = 1.$

	C1 = 20		C1 = 30		C1 = 40		C1 = 50		C1 = 60		C1 = 70		C1 = 75		C1 = 80		C1 = 85		C1 = 90		C1 = 100	
									WITH OPTIMAL FIRST STAGE ALLOCATION (C1 + C2 = 100.)													
λ_2	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk	$\underline{n}^o$	Risk
1	(0,20)	.3025	(0,30)	.2949	(0,40)	.2897	(1,49)	.2904	(7,53)	.2950	(13,57)	.3037	(15,60)	.3085	(17,63)	.3139	(20,65)	.3198	(22,68)	.3263	(26,74)	.3413
3	(0,20)	1.0566	(0,30)	1.0561	(0,40)	1.0531	(0,50)	1.0490	(0,60)	1.0450	(0,70)	1.0412	(0,75)	1.0394	(0,80)	1.0381	(0,85)	1.0392	(0,90)	1.0428	(0,100)	1.0564

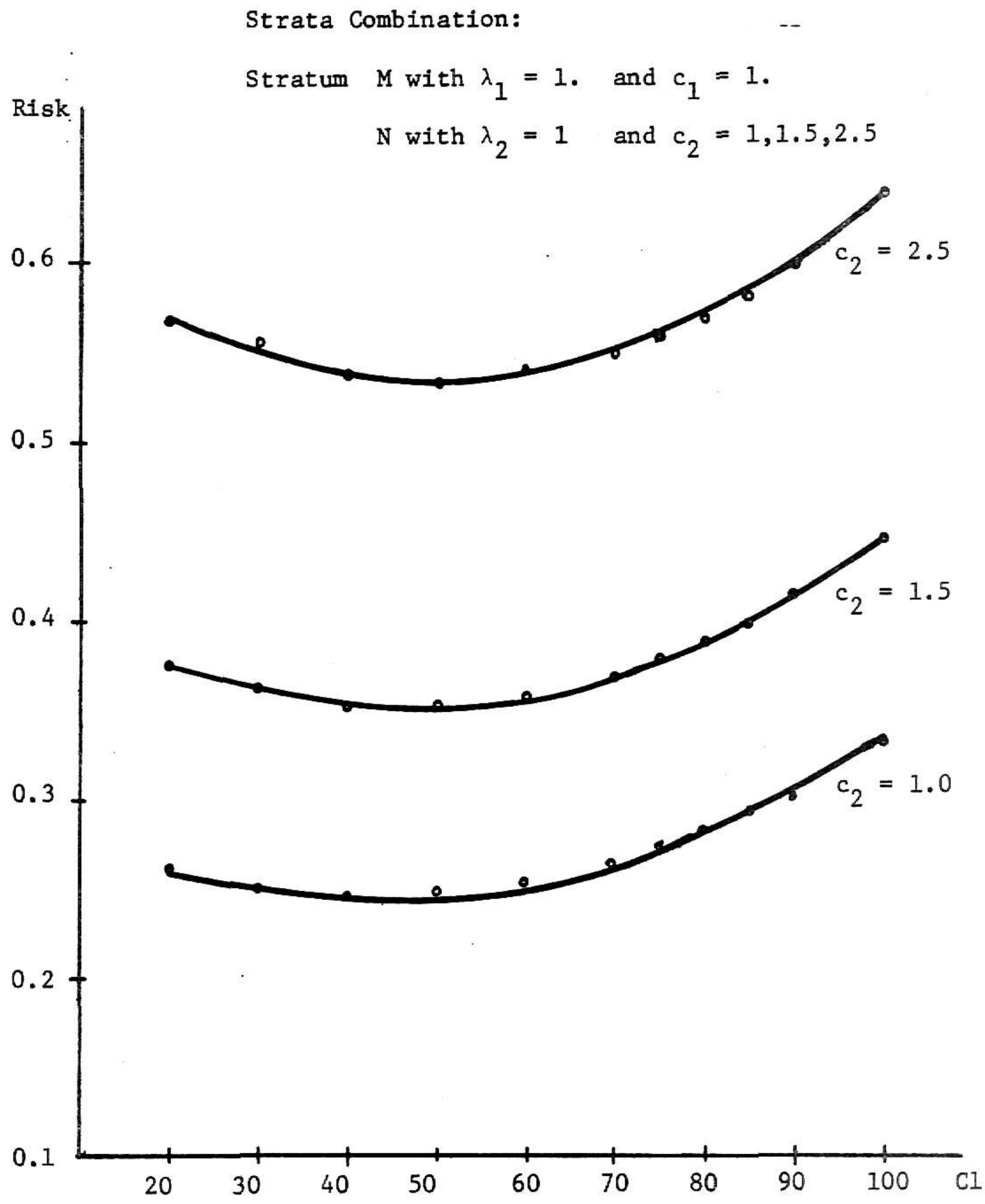


Figure 3.1 Risk x 10^3 as function of the budget partitioning for a fixed budget. Data from Tables 3.2,3.3 and 3.4.

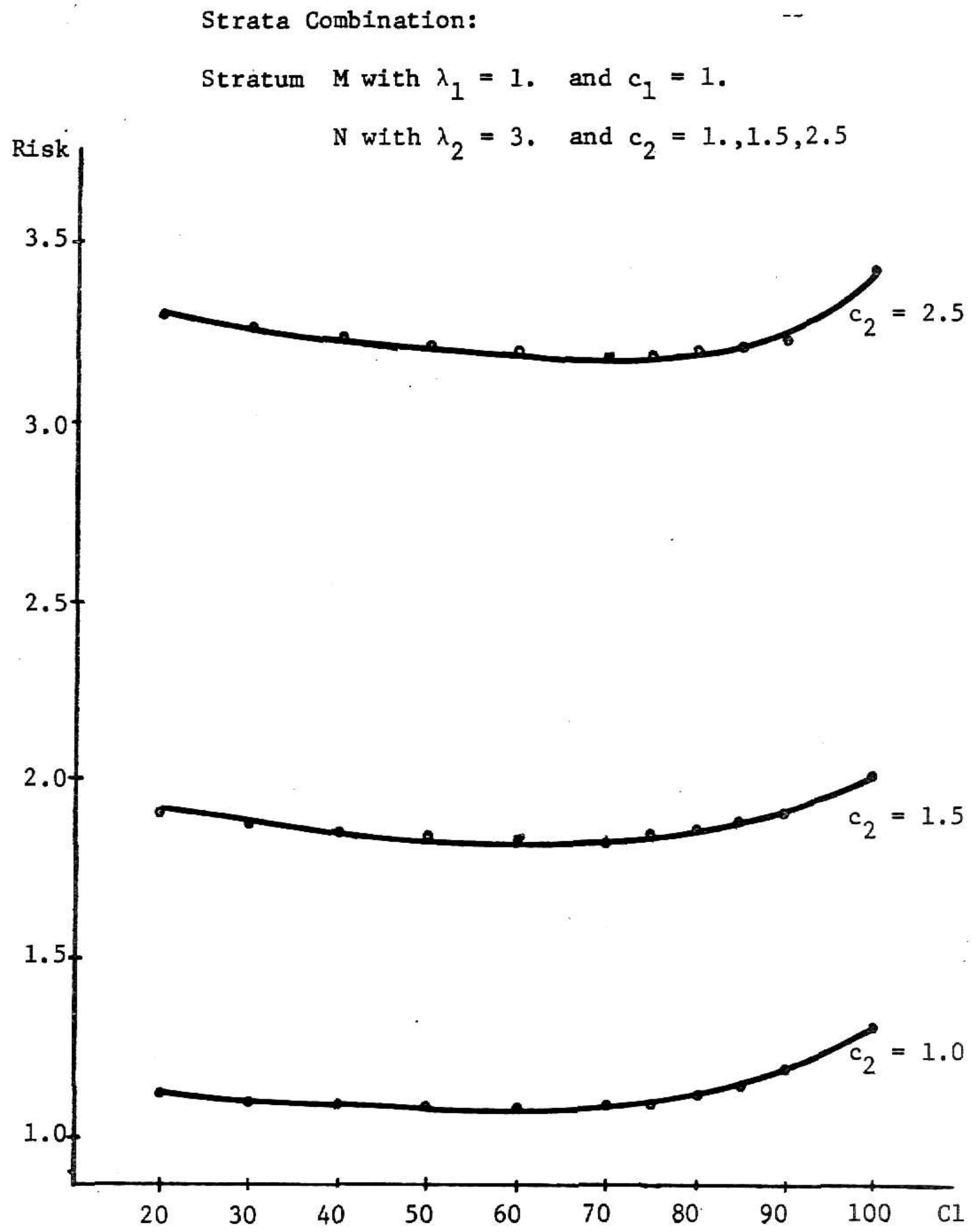


Figure 3.2. Risk x 10^3 as function of the budget partitioning for a fixed budget. Data from Tables 3.2,3.3,3.4.

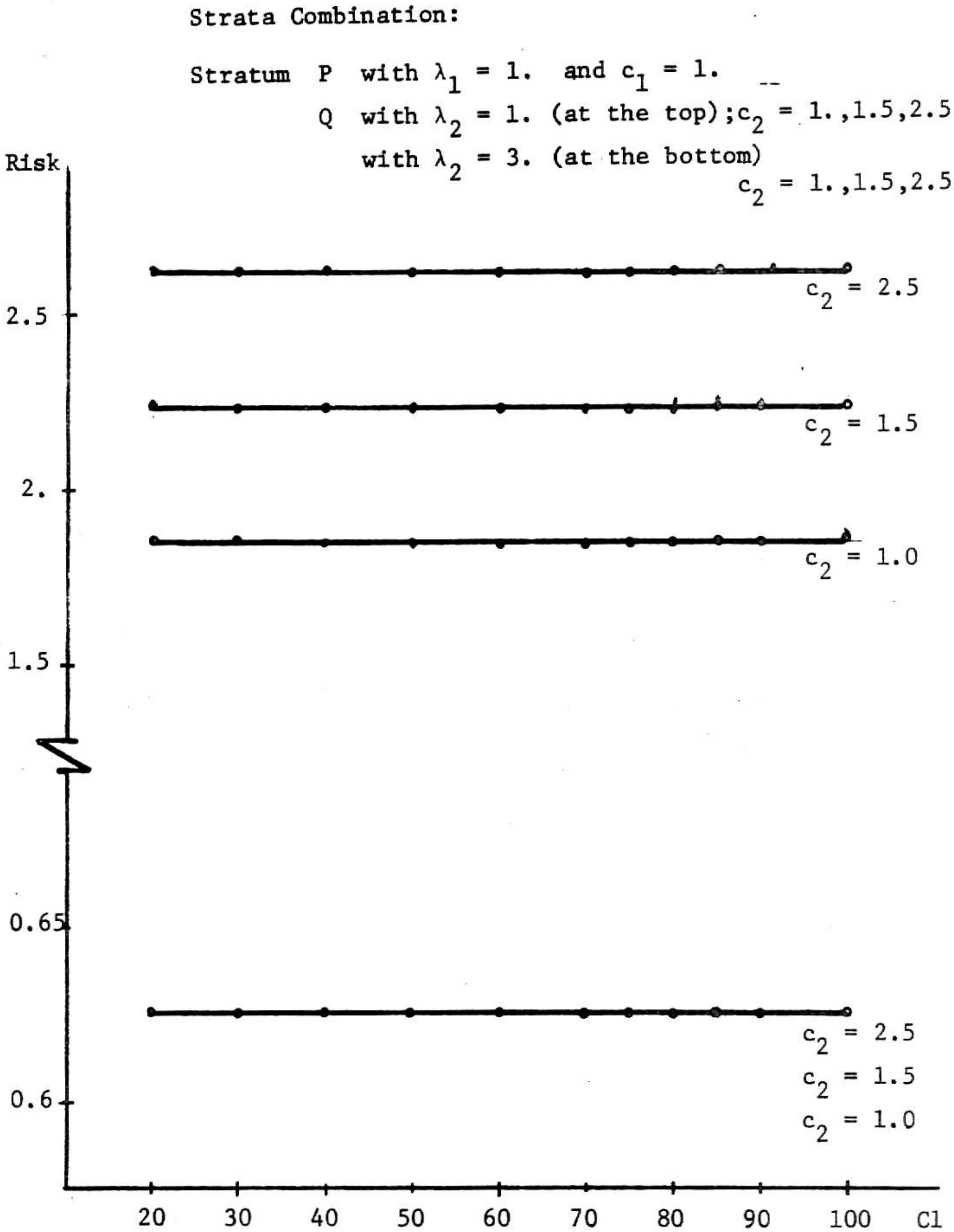


Figure 3.3. Risk x 10^3 as function of the budget partitioning for a fixed budget. Data from Tables 3.5,3.6,3.7.

Strata Combination:

Stratum P with λ_1 = 1. and c_1 = 1.

S with λ_2 = 1. and c_2 = 1.,1.5,2.5

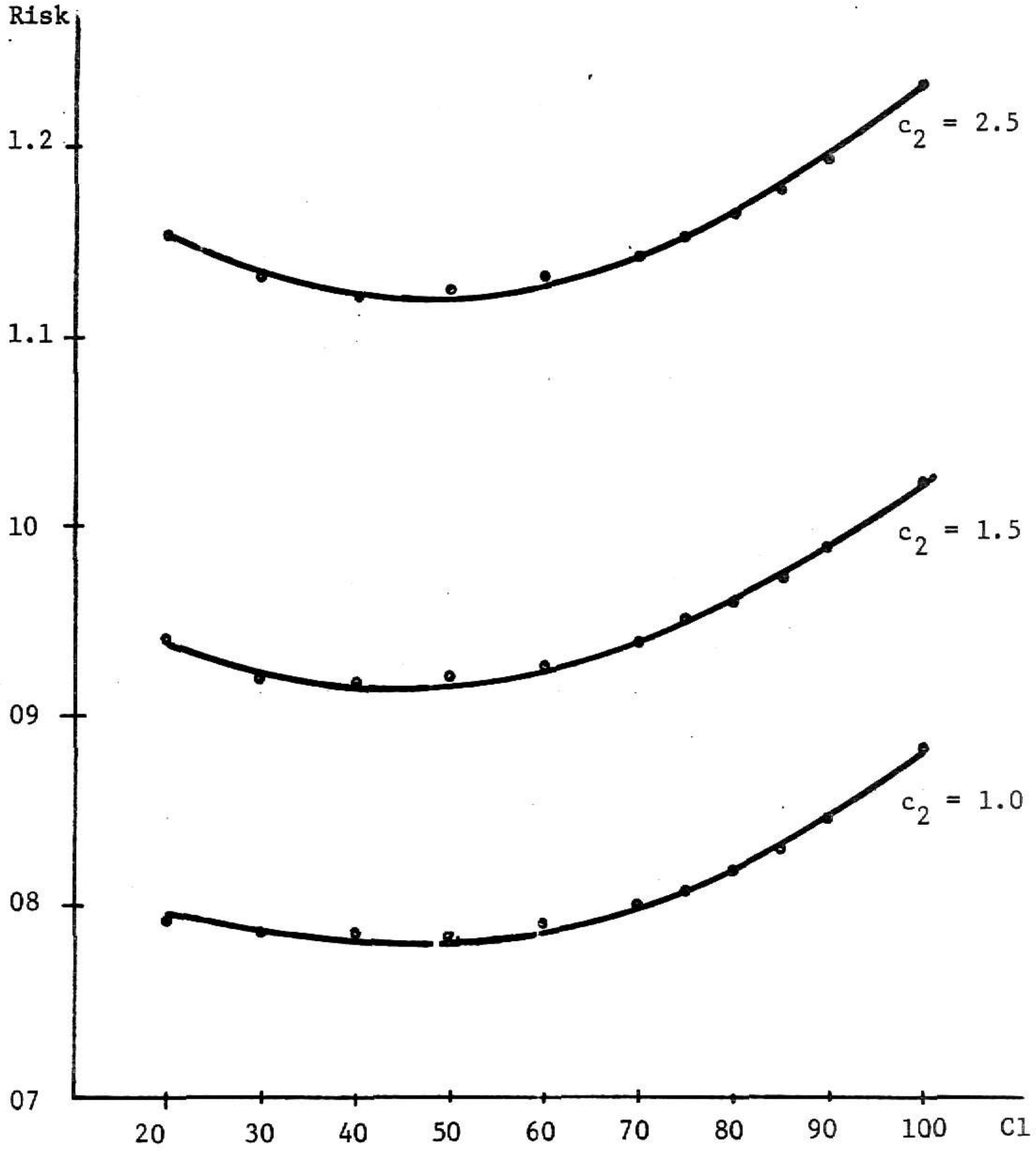


Figure 3.4. Risk x 10^3 as function of the budget partitioning for a fixed total budget. Data from Tables 3.8, 3.9,3.10

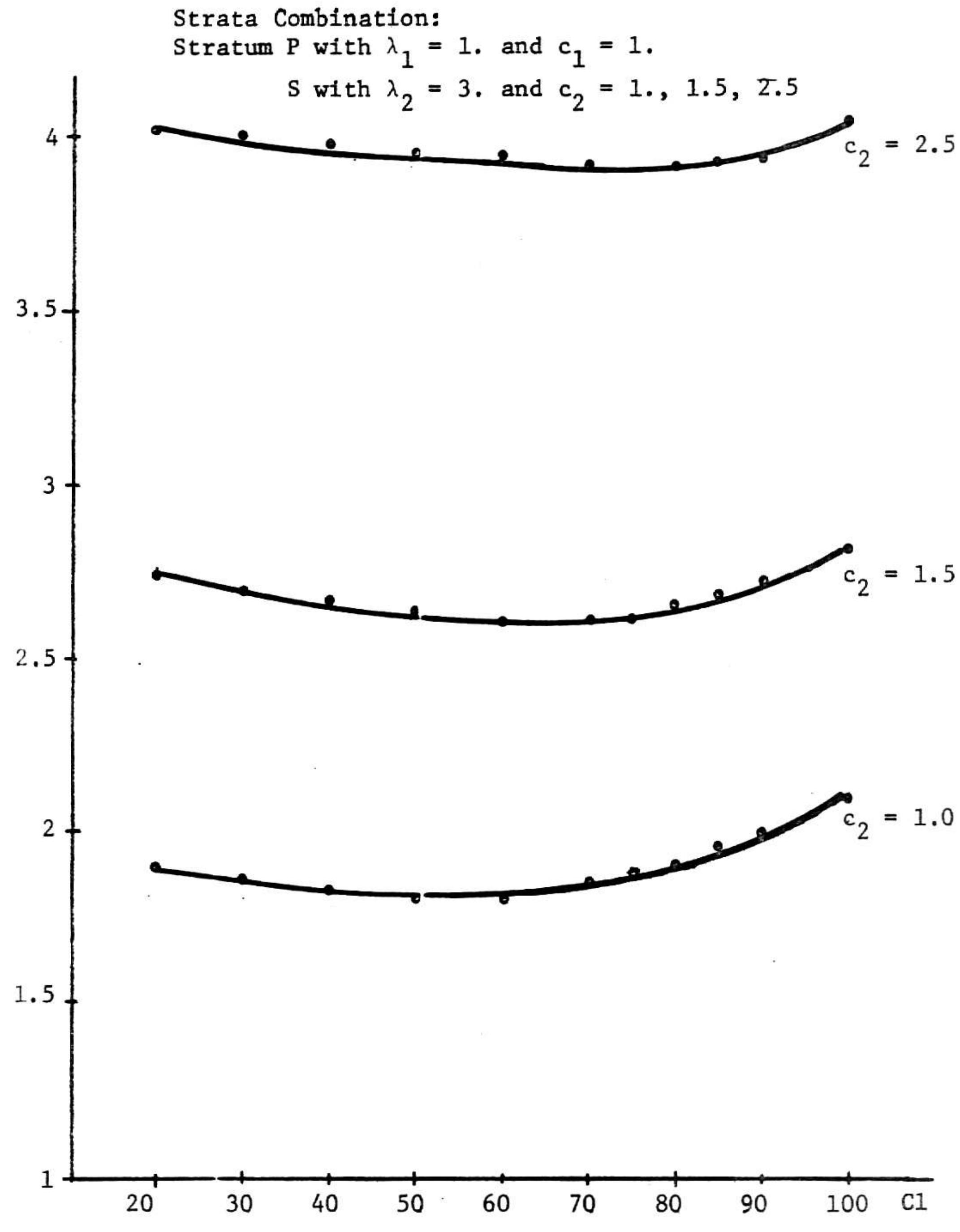


Figure 3.5. Risk x 10^3 as function of the budget partitioning for a fixed total budget. Data from Table 3.8, 3.9, 3.10

CHAPTER 4

CONCLUSIONS

1) Clearly it is impossible to proceed with an exhaustive study of the present problem, since the number of possible stratum definition is infinite, as is the set of possible lamda's, stratum sizes, and stratum unit costs. Furthermore, we can have problems with a wide range in the number of strat involved.

Consequently our results are based only on the cases we studied, assuming that they can be projected to a generalization. But the real behavior of a given problem is very closely related with the structure of the strata in the set and an answer can not always be anticipated.

2) Due to the form of the risk function (See Equations 1.25 and 1.48) the risk curve is convex (except for round-off "ripples") so the program yields a solution in which the risk is minimum. The degree of the convexity depend on the strata combination for a particular problem.

3) Working with two-strata problems and holding all the stratum constants fixed except the value of the ratio λ_2/λ_1 it was noted that in general:

A) The value of C1, the budget for the optimal first stage sampling, increases when λ_2/λ_1 increases. That is, the minimum point in the risk curve moves to the right when the ratio λ_2/λ_1 increases (see Fig. 3.1 to 3.5).

B) When the ratio λ_2/λ_1 increases the risk also increases (see Tables 3.2 to 3.10).

4) Working with two-strata problems and holding all the strata constants fixed except the value of the ratio c_2/c_1, it was found that the risk curves are very similar in shape to each other but the risk curve as a whole is shifted to a higher value for larger values of c_2/c_1 (see Figures 3.1 to 3.5).

This result was expected and is explained by the fact that the total allocation should decrease when the individual sample cost increases, since we are subject to the same total cost constraint.

5) It was found in some special two strata cases that the risk curves do not follow the general behavior explained above. (See Figure 3.3). In these cases the convexity radius is very large so that for practical purposes the curves are flat. It was believed that the following factors cause this behavior:

a) Our preconceptions about both strata (as reflected in the choice of a_i and b_i) are similar. Those are the two cases studied by Grosh in [1,3].

b) Our preconceptions about both strata are strong, regardless the difference in the expected fraction defective value. This fact is represented by high values of $(a_i + b_i)$. This case is typified by the combination we did of strata P and Q.

c) It was thought at first that one reason for that flat risk curve might be the fact that the term $(a_i + b_i)$ was the same for both strata, and the fact that the two of them had the same expected fraction defective value, $\frac{a_i}{a_i+b_i}$. The last two mentioned causes were withdrawn in light of the results obtained in the brief study summarized in Tables 3.11 and 3.12. There we see the usual convex behavior of the risk function.

6) It was recognized that the search procedure used in each case is not unique and/or exhaustive, due to the fact that we are dealing with a combinatorial problem in which the number of possible first stage allocations may increase to some very large value. However we believe that our results are very close to the right ones. It is possible that they may be improved in the future with the use of some different and more sophisticated searching technique.

7) Defining improvement as follows:

$$\frac{\text{Risk of Optimal Single Plan - Risk of Best Optimal Double Plan}}{\text{Risk of Optimal Single Plan}},$$

we conclude that, in general, two stage sampling is better than single stage sampling (see Tables 4.1 to 4.6). Of course there exist some special cases in which this is not true. For instance, the cases worked by Grosh (see Tables 8.8 and 8.9 in [1]) and the case worked here with strata P and Q (see Table 4.2). Again, we repeat, the behavior of a given problem, in general can not be forecasted, due to the influence of many factors in the final results.

8) In general, as it was intuitivaly expected, the inclusion of a particular stratum as well as its participation in the allocation are influenced by the different stratum constants as follows:

Larger values of λ_i tend to increase the stratum participation.

Larger values of c_i tend to prevent that participation.

Weak prior knowledge about the stratum, which is represented by smaller values of $(a_i + b_i)$(for a given $E(P_i)$ value) tend to increase the number of units to be sampled in that particular stratum.

9) We leave for future investigation to study the cases in which the set up costs are included in the total disposable budget C.

Table 4.1

Minimal Prior Bayes Risk x 10^3 for Single Sampling vs. Best Double Sampling for the Set Formed by Stratum M and Stratum N.

From Table 3.2 $\lambda_1 = 1.$ $c_1 = 1.$ $c_2 = 1.$

λ_2	OPTIMAL SINGLE PLAN	OPTIMAL C1(*)	BEST OPTIMAL DOUBLE PLAN	DIFFERENCE	% IMPROVEMENT
1	.3315	50	.2479	0.0836	25.21
1.5	.5269	50	.4030	0.1239	23.51
2	.7550	50	.5936	0.1614	21.37
3	1.3095	60	1.0779	0.2316	22.27

From Table 3.3 $\lambda_1 = 1.$ $c_1 = 1.$ $c_2 = 1.5.$

λ_2	OPTIMAL SINGLE PLAN	OPTIMAL C1(*)	BEST OPTIMAL DOUBLE PLAN	DIFFERENCE	% IMPROVEMENT
1	.4448	40	.3521	0.0927	20.84
1.5	.7499	50	.6158	0.1341	17.88
2	1.1248	60	.9530	0.1718	15.27
3	2.0715	70	1.8195	0.2520	12.16

From Table 3.4 $\lambda_1 = 1.$ $c_1 = 1.$ $c_2 = 2.5$

λ_2	OPTIMAL SINGLE PLAN	OPTIMAL C1(*)	BEST OPTIMAL DOUBLE PLAN	DIFFERENCE	% IMPROVEMENT
1	.6371	50	.5344	0.1027	16.11
1.5	1.1387	50	.9923	0.1459	12.81
2	1.7748	70	1.5900	0.1848	10.41
3	3.4525	75	3.1793	0.2727	7.89

*IN CASE OF A TIE THE SMALLER ONE WAS CHOSEN AS OPTIMAL.

Table 4.2

Minimal Prior Bayes Risk x 10^3 for Single Sampling vs. Best Double Sampling for the Set Formed by Stratum P and Stratum Q.

From Table 3.5 $\lambda_1 = 1.$ $c_1 = 1.$ $c_2 = 1.$

λ_2	OPTIMAL SINGLE PLAN	OPTIMAL C1(*)	BEST OPTIMAL DOUBLE PLAN	DIFFERENCE	% IMPROVEMENT
1	.6252	85	.6249	0.0003	0.04
1.5	.9357	70	.9272	0.0085	0.90
2	1.2475	60	1.2306	0.0169	1.35
3	1.8713	60	1.8371	0.0342	1.82

From Table 3.6 $\lambda_1 = 1.$ $c_1 = 1.$ $c_2 = 1.5$

1	.6252	20	.6252	0.0	0.0
1.5	.9803	75	.9791	0.0012	0.12
2	1.3805	70	1.3727	0.0078	0.56
3	2.2433	70	2.2211	0.0222	0.98

From Table 3.7 $\lambda_1 = 1.$ $c_1 = 1.$ $c_2 = 2.5$

1	.6252	20	.6252	0.0	0.0
1.5	.9891	75	.9890	0.0001	0.01
2	1.4834	90	1.4789	0.0045	0.30
3	2.6230	85	2.6139	0.0091	0.34

*IN CASE OF A TIE THE SMALLER ONE WAS CHOSEN AS OPTIMAL.

Table 4.3

Minimal Prior Risk x 10^3 for Single Sampling vs. Best Double Sampling for the Set Formed by Stratum P and Stratum S

From Table 3.8 $\lambda_1 = 1.$ $c_1 = 1.$ $c_2 = 1.$

λ_2	OPTIMAL SINGLE PLAN	OPTIMAL C1(*)	BEST OPTIMAL DOUBLE PLAN	DIFFERENCE	% IMPROVEMENT
1	.8841	30	.7850	0.0991	11.20
1.5	1.1830	40	1.0171	0.1659	14.02
2	1.4818	40	1.2555	0.2263	15.27
3	2.0795	50	1.8081	0.2714	13.05

From Table 3.9 $\lambda_1 = 1.$ $c_1 = 1.$ $c_2 = 1.5.$

1	1.0228	40	.9172	0.1056	10.32
1.5	1.4347	40	1.2568	0.1779	12.39
2	1.8683	50	1.6457	0.2226	11.91
3	2.8012	60	2.5946	0.2066	7.37

From Table 3.10 $\lambda_1 = 1.$ $c_1 = 1.$ $c_2 = 2.5$

1	1.2296	40	1.1204	0.1092	8.88
1.5	1.8266	50	1.6453	0.1813	9.92
2	2.4884	60	2.2724	0.2160	8.68
3	4.0307	75	3.9104	0.1203	2.98

*IN CASE OF A TIE THE SMALLER WAS CHOSEN AS OPTIMAL.

Table 4.4

Minimal Prior Bayes Risk x 10^3 for Single Sampling vs. Best Double Sampling for the Set Formed by Stratum L and Stratum S.

From Table 3.11 $\lambda_1 = 1.$ $c_1 = c_2 = 1.$

λ_2	OPTIMAL SINGLE PLAN	OPTIMAL C1	BEST OPTIMAL DOUBLE PLAN	DIFFERENCE	% IMPROVEMENT
1	1.8936	60	1.7632	.1304	6.88
3	4.5985	50	3.9982	.6003	13.05

Table 4.5

Minimal Prior Bayes Risk x 10^3 for Single Sampling vs. Best Double Sampling for the Set Formed by Stratum Q and Stratum S

From Table 3.12 $\lambda_1 = 1.$ $c_1 = c_2 = 1.$

λ_2	OPTIMAL SINGLE PLAN	OPTIMAL C1	BEST OPTIMAL DOUBLE PLAN	DIFFERENCE	% IMPROVEMENT
1	.3413	40	.2897	0.0516	15.11
3	1.0564	80	1.0381	0.0183	1.73

Table 4.6

Prior Bayes Risk x 10^2 for Double Sampling Allocation of the Following Four Strata Combination:

Stratum	λ_i	c_i
P	1.35	1.10
S	1.00	0.95
N	3.00	2.00
M	1.20	1.55

Cl	$n^{(0)}$	Risk
10	(0,0,5,0)	.6490
20	(0,2,9,0)	.6285
30	(0,2,14,0)	.6146
40	(0,2,19,0)	.6047
50	(0,4,23,0)	.5977
60	(0,5,26,2)	.5974*
70	(0,8,31,0)	.6095
80	(0,11,33,2)	.6210
90	(0,12,36,4)	.6436
100	(0,14,40,4)	.6792**

*Optimal double sampling scheme

**Optimal single sampling scheme

$$\text{IMPROVEMENT} = \frac{.6792 - .5974}{.6792} = 0.1204 = 12.04\%$$

Strata Combination:

Stratum	λ_i	c_i
P	1.35	1.10
S	1.00	0.95
N	3.00	2.00
M	1.20	1.55

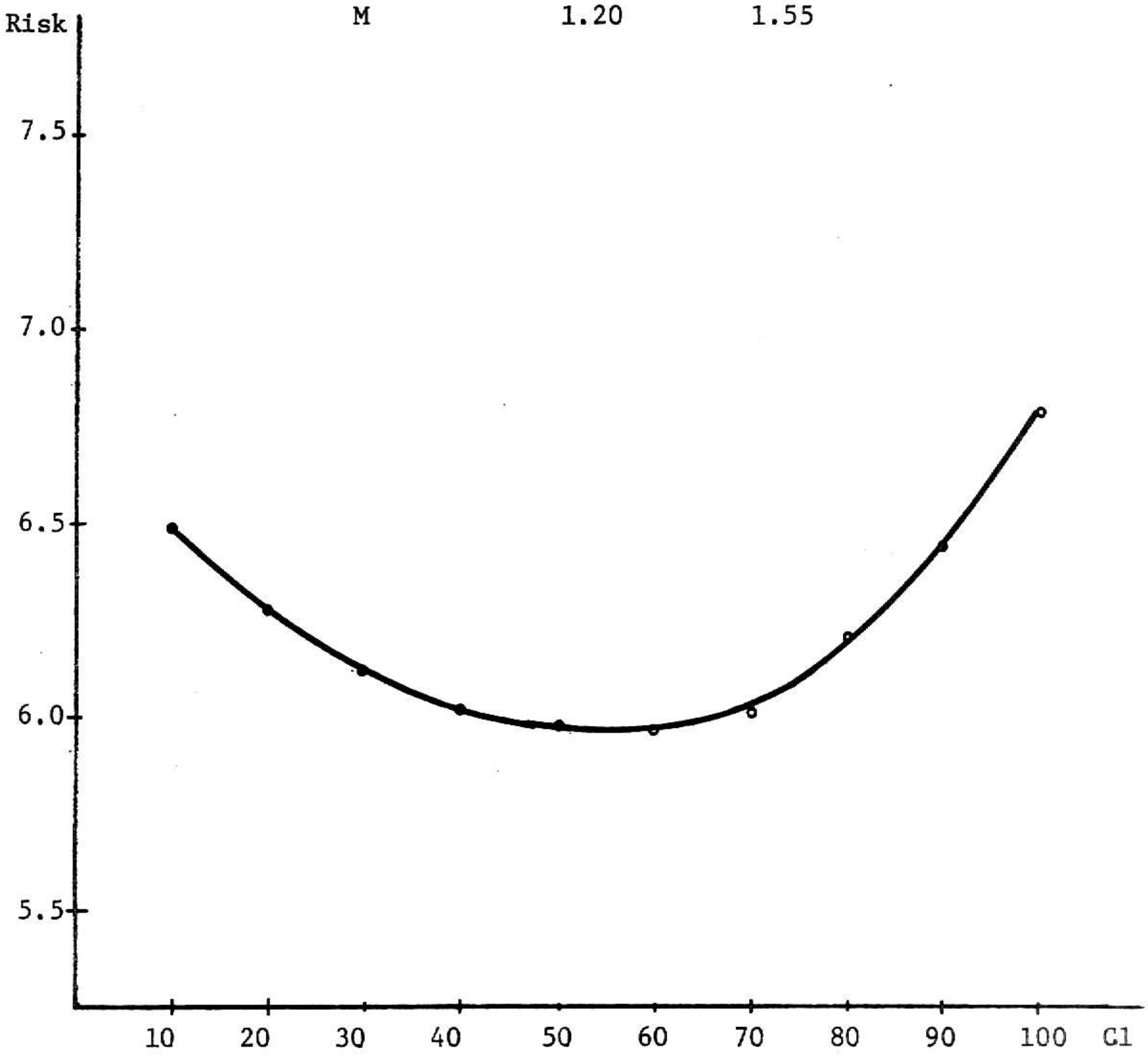


Figure 4.1 Risk x 10^3 as function of the budget partitioning for a fixed total budget. Data from Table 4.6.

REFERENCES

[1] Grosh, Doris Lloyd (1969). Bayes One- and Two- Stage Stratified Sampling Schemes for Finite Population with Beta-Binomial Prior Distributions. Ph.D. Dissertation. Department of Statistics and Computer Science, Kansas State University.

[2] Grosh, Doris Lloyd (1972) "A Bayes Sampling Allocation Scheme for Stratified Finite Population with Hyperbinomial Prior Distribution". TECHNOMETRICS, Vol 14 No. 3, 599-612.

[3] Grosh, Doris L. and L. E. Grosh (1970) "A Bayes Two-Stage Sampling Scheme for Stratified Finite Populations with Hyporbinomial Prior Distributions". Unpublished manuscript.

[4] Military Standard "Sampling Procedures and Tables for Inspection by Attributes" (MIL-STD-105D). 1963 Department of Defence U.S.A.

[5] Raiffa, Howard, and Robert Schaifer (1961). Applied Statistical Decision Theory, Division of Research, Harvard Business School, Boston.

[6] Zacks, S. (1970) "Bayesian Design of Single and Double Stratified Sampling for Estimating Proportion in Finite Population" Technometrics Vol 12, No. 1 119-130.

APPENDIX A

THE COMPUTER PROGRAM

Definition of notation to be used in the present program.

A(I) and R(I)	= a_i
B(I) and S(I)	= b_i
C(I)	= c_i, Unit sampling cost in the i^{th} stratum
C1	= Total budget to be used in the first stage allocation
C2	= Total budget to be used in the second stage allocation
C1M and C1N	= Tentatives total allocation costs
C1A	= Total cost of a particular allocation
CUMF	= Cumulative probability of F(JPOS)
DELTA	= Introduced modification in the allocation of some particular stratum during the searching procedure.
DNUM	= Numerator for the evaluation of Equation (1.29)
DNOM	= Denominator for the evaluation of Equation (1.29)
ERN	= Expected Risk
F(JPOS)	= Probability associated with the number of defective
G(I)	= γ_i value, as defined by Equation (1.26)
H(JPOS)	= $\gamma_i^{(2)}$ value, as defined by Equation (1.49)
I	= Index in which the strata were feeded-in
IACTL	= Actual value of index I
ICON	= Signal used in the searching procedure
INALOC	= Signal used to indicate the type of initial allocation desired INALOC = 0 Normal procedure using ALLOCN subroutine $\neq$ 0 Forced in values to read as data
ITEST	= Signal used to indicate whether or not you want some intermediate results ITEST = 0 NOT DESIRED > 0 DESIRED

IY	= Number of defective unit in the sample
IX(1)	= Index for stratum order after they had been reorder according with the D_i index defined in Equation (1.36)
J(I)	= J_j value defined by Equation (1.28)
JPOS	= Position index on the P matrix
JTEST	= Signal to indicate when is necessary to recompute the initial first allocation
JUG	= Allocation adjustment during the searching procedure
K	= Number of strata studied in the present case
KEY	= Signal used in the searching procedure
L(I)	= L_j value defined by (1.51a)
LTEST	= Signal to indicate when is necessary to recompute the initial second stage allocation
LAMDA(I)	= λ_i, Weight factor for the i^{th} stratum
(LEAST)	= Index to indicate the least important stratum
M(I)	= Second stage sample size for the i^{th} stratum
N(I)	= First stage sample size for the i^{th} stratum
NCASE	= Number of cases to be run
NMAX(I)	= Maximum possible sample size in the i^{th} stratum
NOP(I)	= Optimal sample size in the i^{th} stratum
(NLS)	= Index for next least important stratum
NS(I)	= N_i, stratum size
NTRY	= Tentative new allocation for a particular stratum
NXTRA	= Next important stratum
P(I,J)	= P matrix, used to save F(JPOS) and H(JPOS)
TRY	= Tentative budget for the new allocation in study.
RMIN	= Minimum risk value
Z(I)	= D_i, Importance ranking index as defined by Equation (1.36)

```
      SUBROUTINE RISK(A,B,C,N,NS,LAMDA,K,C2,ERN,IX,ITEST)
      INTEGER P(2,9),M(9),L(9),IJ(9),N(9),NS(9),IX(9)
      REAL LAMDA(K),C(K),A(K),B(K)
      DOUBLE PRECISION R(9),S(9),F(500),H(500),RSK,PROB,ERN,WA,WB
      DO 40 I=1,K
      R(I)=A(I)
40    S(I)=B(I)
C
C          INITIALIZE F(JPOS)  AND  H(JPOS)
C
      DO 1 KQ=1,500
      F(KQ)=1.
1     H(KQ)=0.
C
C          CALCULATE AND SAVE GAMMA VALUE FOR SAMPLE SIZE AND NUMBER OF
C          DEFECTIVE UNITS EQUAL TO ZERO
C
      P(1,1)=1
      DO 3 I=1,K
      JPOS=P(1,I)
      WA  =(LAMDA(I)**2)*((R(I)+S(I)+NS(I))**2)*R(I)*(S(I)+N(I))
      WB=  NS(I)**2*(R(I)+S(I)+N(I))*(R(I)+S(I)+N(I)+1)
      H(JPOS)=WA/WB
      ISIZE=N(I)
      IF(ISIZE.EQ.0)GO TO 300
C
C          CALCULATE AND SAVE THE PROBABILITY ASSOCIATED WITH
C          NO DEFECTIVE UNITS AND SAMPLE SIZE NOT EQUAL TO ZERO
C
      DO 4 IK=1,ISIZE
    4 F(JPOS)=F(JPOS)*(S(I)+IK-1)/(R(I)+S(I)+IK-1)
      CUMF=F(JPOS)
      KOUNT=0
C
C          CALCULATE AND SAVE THE PROBABILITY AND GAMMA VECTORS ASSOCIA-
C          TED WITH  IY  DEFECTIVE UNITS AND SAMPLE SIZE NOT ECUAL TO
C          ZERO
C
      DO 5 IY=1,ISIZE
      F(JPOS+1)=F(JPOS)*(A(I)+IY-1)*(N(I)-IY+1)/(IY*(B(I)+N(I)-IY))
      H(JPOS+1)=H(JPOS)*(A(I)+IY)*(B(I)+N(I)-IY)/((A(I)+IY-1)*(B(I)+N(I)
     X-IY+1))
      JPOS=JPOS+1
      CUMF=CUMF+F(JPOS)
      KOUNT=KOUNT+1
C
C          CHECK FOR CUMULATIVE PROBABILITY OF DEFECTIVE UNITS
C
      IF(CUMF.GE.0.999999)GO TO 301
    5 CONTINUE
      GO TO 301
300   KOUNT=0
301   P(2,I)=P(1,I)+KOUNT
      P(1,I+1)=P(2,I)+1
    3 CONTINUE
      KBAL=K+1
```

```
C
C         FILL OUT THE BALANCE OF THE POSITIONS IN THE P MATRIX
C
      DO 6 J=KPAL,9
      P(1,J)=P(2,J-1)+1
    6 P(2,J)=P(1,J)
      IF(ITEST.EQ.0) GO TO 32
      IP=P(2,K)
      WRITE(3,99) (F(KQ),KQ=1,IP)
      WRITE(3,99) (H(KQ),KQ=1,IP)
99    FORMAT(' ',6G20.13)
C
C         SET UP ITERATION LIMITS
C
   32 I1B=P(1,1)
      I2B=P(1,2)
      I3B=P(1,3)
      I4B=P(1,4)
      I5B=P(1,5)
      I6B=P(1,6)
      I7B=P(1,7)
      I8B=P(1,8)
      I9B=P(1,9)
      I1E=P(2,1)
      I2E=P(2,2)
      I3E=P(2,3)
      I4E=P(2,4)
      I5E=P(2,5)
      I6E=P(2,6)
      I7E=P(2,7)
      I8E=P(2,8)
      I9E=P(2,9)
      ERN=0.
      DO 20 IK=1,K
   20 M(IK)=0
C
C         SET UP POINTER VECTOR
C
      DO 11 I9=I9B,I9E
      IJ(9)=I9
      DO 11 I8=I8B,I8E
      IJ(8)=I8
      DO 11 I7=I7B,I7E
      IJ(7)=I7
      DO 11 I6=I6B,I6E
      IJ(6)=I6
      DO 11 I5=I5B,I5E
      IJ(5)=I5
      DO 11 I4=I4B,I4E
      IJ(4)=I4
      DO 11 I3=I3B,I3E
      IJ(3)=I3
      DO 11 I2=I2B,I2E
      IJ(2)=I2
      DO 11 I1=I1B,I1E
      IJ(1)=I1
```

```
C
C         ALLOCATE TO SECOND STAGE
C
      IF(C2.EQ.0.0)GO TO 29
      IC=0
      DO 21 I=1,K
21    L(I)=1
27    IC=IC+1
      IF(IC.GT.K) GO TO 29
22    CSTAR=C2
      GCL=0.
      DO 23 I=1,K
      IF(L(I).EQ.0)GO TO 23
      CSTAR=CSTAR+C(I)*L(I)*(A(I)+B(I)+N(I))
      Y=H(IJ(I))*C(I)
      GCL=GCL+L(I)*SQRT(Y)
   23 CONTINUE
      LTEST=0
      DO 24 I=1,K
      IF(L(I).EQ.0) GO TO 24
      AY=H(IJ(I))/C(I)
      AM=CSTAR*SQRT(AY)/GCL-(A(I)+B(I)+N(I))
      IF(AM.GT.0) AM=AM+.5000
      M(I)=AM
      IF(M(I).LE.0) GO TO 25
      IF(M(I).GT.(C2/C(I))) M(I)=C2/C(I)
      IF(M(I).GT.(NS(I)-N(I))) M(I)=NS(I)-N(I)
      GO TO 24
25    M(I)=0
      L(I)=0
      LTEST=1
24    CONTINUE
C
C         CHECK FOR THE LTEST VALUE,WHEN EQUAL TO ONE RECOMPUTATION OF
C         SECOND STAGE ALLOCATION IS NECESSARY
C
      IF(LTEST.EQ.1) GO TO 27
C
C         ADJUST LEAST IMPORTANT STRATUM TO MAINTAIN C2
C
      KM=K
33    KK=KM-1
      TRY=0.
      DO 34 I=1,KK
34    TRY=TRY+M(IX(I))*C(IX(I))
 101  M(IX(KM))=(C2-TRY)/C(IX(KM))
      IF(M(IX(KM)).GE.0) GO TO 29
      M(IX(KM))=0
      KM=KM-1
      IF(KM.GT.1) GO TO 33
C
C         CALCULATE THE RISK GIVEN AN ALLOCATION VECTOR
C
29    RSK=0.
 100  DO 30 I=1,K
30    RSK=RSK+H(IJ(I))*(1.0/(R(I)+S(I)+N(I)+M(I))-1.0/(R(I)+S(I)+NS(I)))
```

```
C
C          CALCULATE THE PROBABILITY OF THE RISK
C
      PROB=1.
      DO 31 I=1,K
31    PROB=PROB*F(IJ(I))
C
C          CALCULATE EXPECTATION
C
      ERN=ERN+ RSK*PROB
      IF(ITEST.EQ.0) GO TO 11
      WRITE(3,98) (IJ(I),I=1,3),RSK,PROB,ERN,(M(I),I=1,K)
98    FORMAT(' ',3I4,3G20.13,9I4)
11    CONTINUE
      RETURN
      END



      SUBROUTINE ALLOCN(A,B,C,G,C1,N,K,J,IX,NS)
      INTEGER N(5),J(K),IX(K),NS(K)
      REAL C(K),G(K),A(K),B(K)
C
C          PROCEED WITH INITIAL FIRST ALLOCATION ACCORDING TO EQU.  1.29
C
      DO 5 I=1,K
    5 J(I)=1
   10 DNUM=C1
      DNOM=0.
      DO 1 I=1,K
      IF(J(I).EQ.0)GO TO 1
      DNUM=DNUM+J(I)*C(I)*(A(I)+B(I))
      DNOM=DNOM+J(I)*SQRT(G(I)*C(I))
    1 CONTINUE
      JTEST=0
      DO 2 I=1,K
      IF(J(I).EQ.0) GO TO 2
      AM   =DNUM*SQRT(G(I)/C(I))/DNOM-(A(I)+B(I))
      IF(AM.GT.0) AM=AM+.5000000
      N(I)=AM
      IF(N(I).LE.0) GO TO 3
      IF(N(I).GT.(C1/C(I))) N(I)=C1/C(I)
      IF(N(I).GT.NS(I)) N(I)=NS(I)
      GO TO 2
    3 N(I)=0
      J(I)=0
      JTEST=1
    2 CONTINUE
C
C          CHECK FOR THE JTEST VALUE,WHEN EQUAL TO ONE RECOMPUTATION OF
C          INITIAL FIRST STAGE ALLOCATION IS NECESSARY
C
      IF(JTEST.EQ.1) GO TO 10
      KM=K
33    KK=KM-1
```

```
C
C         PROCEED WITH THE NECESSARY MODIFICATIONS IN THE ALLOCATION IN
C         ORDER TO MEET THE COST CONSTRAINT
C
      TRY=0.
      DO 32 I=1,KK
32    TRY=TRY+N(IX(I))*C(IX(I))
      IF(TRY.GT.C1)GO TO 34
      N(IX(KM))=(C1-TRY)/C(IX(KM))
      IF(N(IX(KM)).GE.0) RETURN
   34 N(IX(KM))=0
      KM=KM-1
      IF(KM.GT.1) GO TO 33
      RETURN
      END



      SUBROUTINE VALUE(A,B,C,K,M,N,NS,COST,ERN,LAMDA,G,E,C1,C2,L,IX,ITES
     XT)
      INTEGER M(K),N(K),NS(K),L(K),IX(K)
      REAL C(K),LAMDA(K),G(K),E(K),A(K),B(K)
      DOUBLE PRECISION ERN
C
C         EVALUATE THE ACTUAL ALLOCATION COST
C
      C1A=0.
      DO 7 I=1,K
7     C1A=C1A+N(I)*C(I)
      CALL RISK(A,B,C,N,NS,LAMDA,K,C2,ERN,IX,ITEST)
   87 FORMAT ('0',G13.7,'   1   ',F8.2,9I5)
      WRITE(3,87) ERN,C1A,(N(I),I=1,K)
      RETURN
      END



      SUBROUTINE INDEX(IX,N,NOP,NMAX,C,C1M,K,DELTA,I,ICON,C1)
      INTEGER IX(K),N(K),NOP(K),NMAX(K),DELTA
      REAL C(K)
      IF(I.LT.K) GO TO 100
      ICON=2
      RETURN
  100 ICON=0
      DO 101 IK=1,K
  101 N(IK)=NOP(IK)
      IF(DELTA.LE.0) GO TO 500
C
C         FOR POSITIVE DELTA VALUE INCREASE THE MOST IMPORTANT STRATUM
C         ALLOCATION
C
      J=K
      MOST=IX(I)
      N(MOST)=NOP(MOST)+DELTA
      C1N=C1M+C(MOST)*DELTA
```

```
C
C         MODIFY THE LEAST IMPORTANT STRATUM IN ORDER TO MEET THE FIRST
C         STAGE BUDGET CONSTRAINT
C
  150 DIFER=C1-C1N
      LEAST=IX(J)
      IF (DIFER.GE.0)GO TO 200
      JUG=-.99+DIFER/C(LEAST)
      GO TO 210
  200 JUG=DIFER/C(LEAST)
  210 NTRY=N(LEAST)+JUG
C
C         CHECK FOR THE VALIDITY OF THE NEW ALLOCATION SIZES
C
      IF(NTRY.LT.0)GO TO 230
      IF(NTRY.LE.NMAX(LEAST)) GO TO 220
      IF(N(LEAST).EQ.NMAX(LEAST))GO TO 235
      N(LEAST)=NMAX(LEAST)
      GO TO 235
  220 N(LEAST)=NTRY
      RETURN
  230 IF(N(LEAST).EQ.0)GO TO 245
      N(LEAST)=0
  235 C1N=0
      DO 240 L=1,K
  240 C1N=C1N+C(L)*N(L)
  245 J=J-1
C
C         IF WE TRY TO MODIFY IN TWO DIFFERENT WAYS THE SAME STRATUM AT
C         THE SAME TIME SET ICON GREATER THAN ZERO AND LEAVE THE ACTUAL
C         ALLOCATION
C
      IF(J.GT.I) GO TO 150
      J=J+1
      DO 250 KJ=1,K
  250 N(KJ)=NOP(KJ)
      ICON=2
      RETURN
C
C         WHEN DELTA HAS A NEGATIVE VALUE REDUCE THE MOST IMPORTANT
C         STRATUM ALLOCATION
C
  500 J=I+1
      MOST=IX(I)
      N(MOST)=NOP(MOST)+DELTA
      C1N=C1M+C(MOST)*DELTA
C
C         MODIFY THE NEXT LEAST IMPORTANT STRATUM ALLOCATION IN ORDER TO
C         MEET THE COST CONSTRAINT
C
  510 DIFER=C1-C1N
      LEAST=IX(J)
      JUG=DIFER/C(LEAST)
      IF(JUG.EQ.0)GO TO 550
```

```
C
C          CHECK FOR THE VALIDITY OF THE NEW ALLOCATION SIZES
C
  520 NTRY=N(LEAST)+JUG
  521 IF(NTRY.GT.NMAX(LEAST))GO TO 560
      N(LEAST)=NTRY
      RETURN
C
C          IF SOME MONEY IS STILL AVAILABLE TRY TO SPEND IT
C
  550 IF(J.EQ.K)GO TO 520
      DO 700 NJ=J,K
      NLS=IX(NJ)
      IF(DIFER.LT.C(NLS))GO TO 700
      NTRY=N(NLS)+DIFER/C(NLS)
      IF(NTRY.GT.NMAX(NLS))GO TO 710
      N(NLS)=NTRY
      RETURN
  710 N(NLS)=NMAX(NLS)
      GO TO 565
  700 CONTINUE
      GO TO 520
  560 N(LEAST)=NMAX(LEAST)
  565 C1N=0.
      DO 570 L=1,K
  570 C1N=C1N+C(L)*N(L)
      J=J+1
      IF(J.LE.K) GO TO 510
      RETURN
      END
```

```
      INTEGER M(9),NS(9),M(9),L(9),J(9),NOP(9),MOP(9),NMAX(9),IX(9)
      REAL C(9),LAMDA(9),G(9),Z(10),E(9),A(9),B(9)
      DOUBLE PRECISION ERN,RMIN
   79 FORMAT(2I5)
      READ(1,79) NCASE,ITEST
      DO 94 ICOUNT=1,NCASE
   80 FORMAT(I5,2F6.2,I2)
   81 FORMAT(2F6.2,I5,F5.2,F6.4)
   83 FORMAT ('1K=',I3,5X,'  C1=',F6.2,'  C2=',F6.2,'  INITIAL ALLOCATIO
     XN TYPE',I2)
   84 FORMAT('0      I    A(I)    B(I) NS(I)    C(I) LAMDA(I)       GAMMA(I)
     X  D(I)')
   85 FORMAT(' ',I6,2F7.2,I6,F7.2,F9.2,F13.6,F10.4)
C
C          READ DATA FOR THE CASE TO BE WORKED OUT
C
      READ(1,80) K,C1,C2,INALOC
      READ(1,81) (A(I),B(I),NS(I),C(I),LAMDA(I),I=1,K)
      DO 3 I=1,K
C
C          EVALUATE GAMMA(I),  G(I)  , ACCORDING TO EQUATION  1.26  AND
C          THE IMPORTANCE FACTOR D(I),  Z(I)  ,ACCORDING TO EQUATION 1.36
C
      G(I)=((LAMDA(I)*(A(I)+B(I)+NS(I))/NS(I))**2)*A(I)*B(I)/((A(I)+B(I)
     X)*(A(I)+B(I)+1))
    3 Z(I)=(A(I)+B(I))*SQRT(C(I)/G(I))
```

```
C
C         WRITE THE PRELIMINARY DATA OF THE CASE
C
      WRITE(3,83) K,C1,C2,INALOC
      WRITE(3,84)
      WRITE(3,85) (I,A(I),B(I),NS(I),C(I),LAMDA(I),G(I),Z(I),I=1,K)
      WRITE(3,86)
   86 FORMAT ('0  RISK',09X,'STAGE',2X,'COST',4X,'ALLOCATION')
C
C         RANK STRATA ACCORDING WITH THE Z(I) INDEX
C
      Z(10)=9999.
      DO 10 I=1,K
   10 IX(I)=10
      DO 11 I=1,K
      DO 12 IJ=1,K
      IF(Z(I).GE.Z(IX(IJ))) GO TO 12
      KK=K-IJ
      IF(K.EQ.IJ) GO TO 15
      DO 14 JJ=1,KK
   14 IX(K+1-JJ)=IX(K-JJ)
   15 IX(IJ)=I
      GO TO 11
   12 CONTINUE
   11 CONTINUE
C
C         CHECK FOR DESIRED INITIAL FIRST ALLOCATION TYPE
C
      IF(INALOC.EQ.0)GO TO 17
      READ(1,87)(N(I),I=1,K)
   87 FORMAT(9I5)
      GO TO 18
   17 CALL ALLOCN(A,B,C,G,C1,N,K,J,IX,NS)
   18 CALL VALUE(A,B,C,K,M,N,NS,COST,ERN,LAMDA,G,E,C1,C2,L,IX,ITEST)
C
C         SET UPPER LIMIT FOR N(I)
C
      DO 20 I=1,K
      NMAX(I)=  C1/C(I)
      IF(NS(I).LT.NMAX(I)) NMAX(I)=NS(I)
   20 CONTINUE
C
C         BEGIN THE SEARCHING PROCEDURE INITIALIZE ALL THE SIGNALS
C
      KEY=1
      NXTRA=1
      IACTL=K
C
C         SAVE THE LOWEST RISK VALUE AND THE CORRESPONDING ALLOCATION
C
   22 C1M=0.
      DO 21 I=1,K
      C1M=C1M+C(I)*N(I)
  21  NOP(I)=N(I)
      RMIN=ERN
      ICON=0
      DO30 I=1,K
      IF(NOP(IX(I)).GT.NMAX(IX(I))) GO TO 31
      IF(I.GE.IACTL)GO TO 35
```

```
C
C          PROCEED WITH SEARCHING PROCEDURE INCREMENT THE PARTICIPATION
C          OF THE MOST IMPORTANT STRATUM,DELTA=+1
C
      CALL INDEX(IX,N,NCP,NMAX,C,C1M,K,+1,I,ICON,C1)
      IF(ICON.GT.0)GO TO 35
      CALL VALUE(A,B,C,K,M,N,NS,CCST,ERN,LAMDA,G,E,C1,C2,L,IX,ITEST)
      IF(ERN.LT.RMIN)GO TO 302
      DO 350 KJ=1,K
  350 N(KJ)=NCP(KJ)
      GO TO 30
  302 NXTRA=I+1
      KEY=1
      GO TO 22
   31 KJ=IX(I)
      C1M=C1M-(NOP(KJ)-NMAX(KJ))*C(KJ)
      NOP(KJ)=NMAX(KJ)
   30 CONTINUE
C
C          PROCEED WITH SEARCHING PROCEDURE REDUCING THE PARTICIPATION OF
C          THE MOST IMPORTANT STRATUM, DELTA=-1
C
   35 IF(NXTRA.GE.K) GO TO 400
      IF(KEY.EQ.2)NXTRA=IACTL
      DO 45 I=NXTRA,K
      IF(NOP(IX(I)).EQ.0) GO TO 45
      CALL INDEX(IX,N,NCP,NMAX,C,C1M,K,-1,I,ICON,C1)
      IF(ICON.GT.0)GO TO 400
      CALL VALUE(A,B,C,K,M,N,NS,CCST,ERN,LAMDA,G,E,C1,C2,L,IX,ITEST)
C
C          CHECK FOR THE LOWEST RISK VALUE BETWEEN THE SAVED AND THE NEW
C          ONE
C
      IF(ERN.LT.RMIN)GO TO 48
      DO 450 KJ=1,K
  450 N(KJ)=NCP(KJ)
      GO TO 45
   48 KEY=2
      IACTL=I
      GO TO 22
   45 CONTINUE
  400 CONTINUE
C
C          PRINT ANSWER
C
      WRITE(3,95) RMIN
   95 FORMAT ('0MIN RISK=',G13.7)
      WRITE(3,92) C1M,(NOP(I),I=1,K)
   92 FORMAT ('0  COST',4X,'ALLOCATION'/' ',F6.2,9I5)
   94 CONTINUE
      STOP
      END
```

APPENDIX B

SAMPLE PRINT-OUTS

K= 2 C1= 30.00 C2= 70.00 INITIAL ALLOCATION TYPE 0

I	A(I)	B(I)	NS(I)	C(I)	LAMDA(I)	GAMMA(I)	D(I)
1	3.00	7.00	500	1.00	1.00	0.198621	22.4381
2	3.00	7.00	500	1.00	3.00	1.787594	7.4794

RISK	STAGE	COST	ALLOCATION	
0.2250441D-01	1	30.00	2	28
0.2250935D-01	1	30.00	1	29
0.2250012D-01	1	30.00	3	27
0.2249675D-01	1	30.00	4	26
0.2249334D-01	1	30.00	5	25
0.2249055D-01	1	30.00	6	24
0.2248821D-01	1	30.00	7	23
0.2248613D-01	1	30.00	8	22
0.2248434D-01	1	30.00	9	21
0.2248293D-01	1	30.00	10	20
0.2248166D-01	1	30.00	11	19
0.2248104D-01	1	30.00	12	18
0.2248023D-01	1	30.00	13	17
0.2248076D-01	1	30.00	14	16

MIN RISK=0.2248023D-01

COST	ALLOCATION	
30.00	13	17

K= 2 C1= 75.00 C2= 25.00 INITIAL ALLOCATION TYPE 0

I	A(I)	B(I)	NS(I)	C(I)	LAMDA(I)	GAMMA(I)	D(I)
1	0.10	9.90	200	1.00	1.00	0.009922	100.3899
2	0.10	4.90	200	2.50	1.00	0.017160	60.3502

RISK	STAGE	COST	ALLOCATION	
0.5660523D-03	1	74.50	22	21
0.5601955D-03	1	75.00	20	22
0.5600337D-03	1	74.50	17	23
0.5552642D-03	1	75.00	15	24
0.5589566D-03	1	74.50	12	25

MIN RISK=0.5552642D-03

COST	ALLOCATION	
75.00	15	24

K= 2 C1= 20.00 C2= 80.00 INITIAL ALLOCATION TYPE 1

I	A(I)	B(I)	NS(I)	C(I)	LAMDA(I)	GAMMA(I)	D(I)
1	1.00	1.00	500	1.00	1.00	0.168002	4.8795
2	1.00	1.00	500	1.00	9.00	13.608212	0.5422

RISK	STAGE	COST	ALLOCATION	
0.1342808	1	20.00	10	10
0.1337902	1	20.00	9	11
0.1334680	1	20.00	8	12
0.1332794	1	20.00	7	13
0.1331909	1	20.00	6	14
0.1332061	1	20.00	5	15

MIN RISK=0.1331909

COST	ALLOCATION	
20.00	6	14

K= 2 C1= 50.00 C2= 50.00 INITIAL ALLOCATION TYPE 0

I	A(I)	B(I)	NS(I)	C(I)	LAMDA(I)	GAMMA(I)	D(I)
1	5.00	95.00	500	1.00	1.00	0.067723	384.2668
2	5.00	95.00	500	1.25	1.00	0.067723	429.6233

RISK	STAGE	COST	ALLOCATION	
0.7072826D-03	1	50.00	30	16
0.7077512D-03	1	49.75	31	15
0.7098379D-03	1	49.00	29	16

MIN RISK=0.7072826D-03

COST	ALLOCATION	
50.00	30	16

K= 4 C1= 60.00 C2= 40.00 INITIAL ALLOCATION TYPE 0

I	A(I)	B(I)	NS(I)	C(I)	LAMDA(I)	GAMMA(I)	D(I)
1	10.00	90.00	200	1.10	1.35	0.365402	173.5046
2	0.20	9.80	200	0.95	1.00	0.019645	69.5410
3	0.10	4.90	200	2.00	3.00	0.154442	17.9930
4	0.10	9.90	200	1.55	1.20	0.014288	104.1536

RISK	STAGE	COST	ALLOCATION			
0.60040730D-02	1	59.25	0	6	26	1
0.60117290D-02	1	59.70	0	6	27	0
0.60442410D-02	1	58.65	0	7	26	0
0.60240960D-02	1	59.15	0	8	25	1
0.59741140D-02	1	59.85	0	5	26	2
0.60608280D-02	1	58.75	0	5	27	0
0.59935260D-02	1	60.00	1	4	26	2
0.60223690D-02	1	59.40	1	5	26	1

MIN RISK=0.59741140D-02

COST	ALLOCATION			
59.85	0	5	26	2

INFLUENCE OF STRATUM CONSTANTS ON A TWO STAGE BAYESIAN SAMPLING SCHEME FOR A FINITE, STRATIFIED AND DICHOTOMOUS POPULATION

by

ALFREDO GABRIEL PARDO

Ingeniero Mecánico, Universidad Central de Venezuela

Caracas-Venezuela, 1966

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Industrial Engineering

KANSAS STATE UNIVERSITY

Manhattan, Kansas

1972

INFLUENCE OF STRATUM CONSTANTS ON A TWO STAGE BAYESIAN SAMPLING SCHEME FOR A FINITE, STRATIFIED AND DICHOTOMOUS POPULATION

by

ALFREDO GABRIEL PARDO

Ingeniero Mecánico, Universidad Central de Venezuela

Caracas-Venezuela, 1966

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Industrial Engineering

KANSAS STATE UNIVERSITY

Manhattan, Kansas

1972

Abstract

Given a finite, stratified and dichotomous population with K strata independents of each other, a fraction defective P_i and a importance factor λ_i for each stratum. Sampling is carried out independently in each of the K strata for estimating the linear function

$$\theta = \sum_{i=1}^{K} \lambda_i P_i$$

A Bayesian procedure is used to determine the optimal two-stage allocation subject to a budgetary constraint that does not include the set up cost.

It was found that the final solution as well as behavior of the prior Bayes risk function versus the first-stage total cost is very closely related with the stratum composition of a particular set. As expected, the participation of a particular stratum is also a function of its proper set of constants and the relative value of them with respect to the constants of other stratum.

It also was determined that the two-stage scheme is better than single-stage scheme in the general cases. However in some cases no improvement was achieved.

In order to obtain our results we must appeal to the numerical solution rather than the analytical procedure, using to that end a computer program written in FORTRAN IV level H to be run primarily in a IBM 360/50 computer machine.