OPTIMAL SEQUENTIAL PLANS

BASED ON PRIOR DISTRIBUTIONS AND COSTS

by

Mainak Mazumdar

TECHNICAL REPORT NO. 3

April 1966

Prepared under Contract Nonr-401(53)

(NR 042-244)

for the Office of Naval Research

and

Contract GP-3783

for the National Science Foundation

Reproduction in Whole or in Part is Permitted for

any Purpose of the United States Government

DEPARTMENT OF INDUSTRIAL ENGINEERING AND OPERATIONS RESEARCH

COLLEGE OF ENGINEERING

CORNELL UNIVERSITY

ITHACA, NEW YORK
Abstract

A. Hald in Technometrics, Vol. 2, No. 3 proposed a model of sampling inspection by attributes. He postulated certain definite forms for the losses which are associated with the acceptance or rejection of a lot, and for the cost of taking a single observation. He further assumed that there exists an a priori distribution of the number of defectives, the distribution being known to the experimenter. Under these assumptions he derived the associated single sampling plans which minimize the expected costs averaged with respect to the known a priori distribution. J. Pfanzagl in Technometrics, Vol. 5, No. 2 specialized Hald's model by considering a particular type of a priori distribution, and obtained the double sampling plans which are optimal in the same sense. In this thesis*, Pfanzagl's work has been generalized by extending his results to sequential plans.

The various elements entering into the formulation which underlies Hald's, Pfanzagl's and the present work, are as follows:

a) Model:

Let \( N \) be the number of items in a lot which contains \( \theta \) defective items. It is assumed that \( 0 \leq \theta \leq N \). Items are taken from the lot without replacement and the \( i \)-th observation \( X_i = 1 \) or \( 0 \), according as the \( i \)-th item sampled is defective or effective.

* A dissertation submitted in partial fulfillment of the requirements of the degree of Doctor of Philosophy in the Field of Industrial Engineering and Operations Research, Cornell University, Ithaca, New York, June 1966.
b) **Decision space:**

There are two actions which are denoted by \( a_1 \) and \( a_2 \). \( a_1 \) corresponds to acceptance of the lot and \( a_2 \) to rejection of the lot.

c) **Loss function:**

\[ \theta(j) \quad (j = 0, 1, 2, \ldots, N) \]

denotes the number of defectives in the remainder of the lot when \( j \) items have been inspected. \( L_j (\theta; a_i) \) denotes the loss due to taking action \( a_i \) \((i = 1, 2)\) when \( \theta \) is the true number of defectives after \( X_1, X_2, \ldots, X_j \) have been observed. It is assumed that \( L_j (\theta; a_1) = \theta(j) \) and \( L_j (\theta; a_2) = (N - j) k_r \) \((j = 1, 2, \ldots, N)\), where \( k_r \) is a given positive constant.

d) **Cost due to taking an observation:** \( k_s \), a given positive constant.

e) **A priori distribution:**

\[ \theta \] is assumed to have a Polya distribution given by

\[ \binom{N}{\theta} \frac{\Gamma(\alpha + \theta, \beta + N - \theta)}{\Gamma(\alpha, \beta)} \quad (\theta = 0, 1, 2, \ldots, N). \]

This family of distributions is reproducible to hypergeometric sampling.

In this thesis, the optimal (Bayes) sequential sampling plans are derived for the above model. It is shown that the usual characterization of the Bayes solution in the set-up of identically distributed random variables and stage-independent loss function also applies to the decision problem under consideration although the random variables here are dependent and not identically distributed, and the loss at each stage depends on the observed outcome. An upper bound to the point of truncation for the Bayes sequential procedure is given, and it is indicated that the region of continuation lies within a particular pentagon. The relations for obtaining
the operating characteristics of the optimal sequential sampling plans are derived. The Bayes risks are compared with those of the corresponding optimal single sampling plans for certain representative values of the parameters. It is found that very little is gained by taking observations sequentially when the average risk for the optimal sequential procedure is compared with the corresponding quantity for the optimal single and double sampling plans. However, the values of the ASN-function for the optimal plans are lower than the corresponding fixed sample size of the optimal single-stage procedure over a large range of values of \( \theta \).

The limiting behavior for large lots of the optimal boundaries for the sequential decision problem is considered. Extending Chernoff and Ray's [Annals of Mathematical Statistics, October 1965] approach, the problem of finding the normalized boundaries with appropriate normalization is reduced to that of solving certain partial differential equation free boundary problems. The solution to the latter problem leads to a single optimum boundary although the original decision problem involved two sets of stopping regions. It is suggested that this discrepancy is due to the fact that the two sets of optimal boundaries in the present decision problem are extremely asymmetric with respect to the neutral line. Although it is possible to reduce one set of boundaries to non-degenerate limits by this normalization, it is not possible to obtain finite co-ordinates for the other set by the same transformations.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1. Sampling Inspection by Attributes and Prevailing Concepts in this Area</td>
<td>2</td>
</tr>
<tr>
<td>1.1. Producer's and Consumer's Risks</td>
<td>2</td>
</tr>
<tr>
<td>1.1. Indifference Quality Level</td>
<td>5</td>
</tr>
<tr>
<td>1.1. Dodge-Romig Point of View</td>
<td>5</td>
</tr>
<tr>
<td>1.1. Minimax Principle</td>
<td>7</td>
</tr>
<tr>
<td>1.2. The Sampling Inspection Model due to Hald</td>
<td>8</td>
</tr>
<tr>
<td>1.2.1. Costs of Sampling Inspection</td>
<td>10</td>
</tr>
<tr>
<td>1.2.2. Some Criticisms of Hald's Model</td>
<td>13</td>
</tr>
<tr>
<td>1.2.3. Reproducible Distributions</td>
<td>16</td>
</tr>
<tr>
<td>1.3. A Decision Theoretic Formulation of Pfanzagl's Model of Sampling Inspection, and Statement of our Problem</td>
<td>26</td>
</tr>
<tr>
<td>1.4. Review of Some Sequential Plans by Attributes and Earlier Relevant Work</td>
<td>29</td>
</tr>
<tr>
<td>1.5. Outline of the Thesis</td>
<td>31</td>
</tr>
</tbody>
</table>
## II. SEQUENTIAL BAYES PROCEDURES WITH PFANZAGL'S MODEL OF SAMPLING INSPECTION

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1. Characterization of the Bayes Sequential Procedures with Pfanzagl's Model</td>
<td>34</td>
</tr>
<tr>
<td>2.1.1. Terminal Decision Rule for the Bayes Procedure</td>
<td>35</td>
</tr>
<tr>
<td>2.1.2. Stopping Rule for the Bayes Procedure</td>
<td>39</td>
</tr>
<tr>
<td>2.1.3. An Useful Result for the Computation of the Optimum Boundary</td>
<td>48</td>
</tr>
<tr>
<td>2.1.4. Determination of an Upper Bound on the Point of Truncation for the Bayes Sequential Procedure</td>
<td>52</td>
</tr>
<tr>
<td>2.1.5. Improved Upper Bounds on the Point of Truncation: An Example</td>
<td>60</td>
</tr>
<tr>
<td>2.2. Procedure for the Determination of the Optimal Boundary for the Bayes Sequential Procedure</td>
<td>65</td>
</tr>
<tr>
<td>2.2.1. A Modification of the Computing Procedure: the Use of Certain Transformed Functions</td>
<td>67</td>
</tr>
<tr>
<td>CHAPTER</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>2.3. Determination of Operating Characteristics for the Bayes Sequential Procedure</td>
<td>69</td>
</tr>
<tr>
<td>Determination of the OC-function</td>
<td>69</td>
</tr>
<tr>
<td>Derivation of the ASN-function</td>
<td>71</td>
</tr>
<tr>
<td>2.4. Some Numerical Examples of Bayes Sampling Inspection Plans, and their Properties</td>
<td>72</td>
</tr>
<tr>
<td>2.4.1. Determination of the Acceptance and Rejection Numbers</td>
<td>73</td>
</tr>
<tr>
<td>2.4.2. Some Representative Plans</td>
<td>73</td>
</tr>
<tr>
<td>2.4.3. Graphs of the Stopping Boundaries for Three Particular Plans</td>
<td>88</td>
</tr>
<tr>
<td>2.5. Comparison of the Optimal Sequential Sampling Plans with the Optimum Single and Double Sampling Plans</td>
<td>92</td>
</tr>
<tr>
<td>2.5.1. Comparison of the Bayes Risks</td>
<td>93</td>
</tr>
<tr>
<td>2.5.2. Graphs of the OC- and ASN-Curves for an Optimal Sequential Procedure and the Corresponding Fixed Sample Procedure</td>
<td>95</td>
</tr>
</tbody>
</table>
CHAPTER III. A FREE BOUNDARY PROBLEM RELATED TO BAYES PROCEDURES FOR PFANZAGL'S MODEL.

3.1. Introduction

3.2. Free Boundary Problem Associated with Bayes Sequential Procedures for Pfanzagl's Model when the Lot Size Approaches Infinity

BIBLIOGRAPHY

Page

98

98

101

134
CHAPTER I
INTRODUCTION

In this thesis, we consider a particular model for sampling inspection by attributes which was first proposed by Hald [21]. In his formulation, he postulated certain definite forms for the losses which are associated with the acceptance or rejection of a lot, and the costs due to taking a single observation. He further assumed that there exists an apriori distribution of the number of defectives, the distribution being known to the experimenter, and derived the associated optimal single sampling procedure which minimizes the expected costs averaged with respect to the known apriori distribution. Pfanzagl [33], using Hald's model, but specializing further to a particular type of apriori distribution, obtained the double sampling inspection procedures which are optimal in the same sense. He also made some comparisons between the optimal double and single sampling procedures. Here we generalize Pfanzagl's work by extending his analysis to sequential sampling plans and obtain the optimal sequential procedures.

Before making a formal statement of our problem, we briefly review the concepts which presently prevail in the area of sampling inspection by attributes and reproduce some of the arguments which led Hald to adopt his model. The statistical literature on sampling inspection plans is very vast and no attempt will be made here to review individually the work of scores of contributors in this area. However, since our results are concerned with sequential procedures, we include in a later section of this chapter a brief survey of some of such plans proposed in the literature.
Sampling Inspection by Attributes and Prevailing Concepts in this Area

It is a common practice in industry to subject a small portion of a lot of manufactured product to inspection before it is used. The reasons for carrying out a sampling inspection as opposed to 100% inspection, according to Duncan [16], are the following:

(1) Inspection is often expensive.
(2) 100% inspection is fatiguing and frequently results in poor standards of work.
(3) When testing is destructive, inspection has to be limited to a small portion of the lot.

The type of acceptance sampling plan that merely grades the product as defective or non-defective (also called effective) is called a sampling plan by attributes. Several different concepts have been developed by various statisticians in this area, and attribute sampling plans have been designed in keeping with these ideas. We shall discuss these concepts very briefly below and point out some of the arbitrariness of the existing plans.

Producer's and Consumer's Risks

For a lot size $N$ containing $\theta$ defectives, the probability of accepting the lot, expressed as a function of the fraction defective $p = \frac{\theta}{N}$, is called the operating characteristic or the OC of the particular sampling
plan which is being used. Plans are designed so that a lot which is considered to be of good quality (i.e., having a low fraction defective) has a low probability of being rejected, a low "producer's risk" as it is so called. The maximum fraction defective which is considered satisfactory as a process average is called the Acceptable Quality Level or AQL. Likewise it is specified that material of bad quality will have a low probability of being accepted, i.e., a low "consumer's risk". The fraction defective, chosen to denote the border line between satisfactory and unsatisfactory lots from the consumer's point of view, is called the lot tolerance proportion defective or LTPD.

Let the AQL be denoted by $\theta_0$ and the associated producer's risk be $\alpha$. Similarly let the LTPD be $\theta_1$ and the corresponding consumer's risk be $\beta$ ($0 < \theta_1 < \theta_0 < 1$ and $0 < \alpha, \beta < 1$). Also to avoid trivialities, we assume that $\alpha + \beta < 1$. According to the probability requirements stated above, we wish our plans to satisfy

\[
\begin{align*}
P_{\theta_0} \{ \text{lot is accepted} \} &= 1 - \alpha \\
P_{\theta_1} \{ \text{lot is accepted} \} &= \beta.
\end{align*}
\]

Thus we are specifying two points $(\theta_0, 1 - \alpha)$ and $(\theta_1, \beta)$ on the OC-curve. When the sampling inspection problem is defined in the above manner, it is usual to make the assumption that the lot size is indefinitely large so that the binomial approximation to the hypergeometric distribution can be applied.
There exist many sampling plans which guarantee the probability requirements at the two given points. From among these, a plan is chosen which is optimum in some desired sense. These inspection plans therefore insure that under no condition more than a given proportion of unsatisfactory lots submitted for inspection are accepted or more than a given proportion of satisfactory lots are rejected. This formulation of the inspection problem is typical of those of the Neyman-Pearson type (although it should be pointed out that these ideas preceded the Neyman-Pearson formulation by several years) and is, therefore, subject to the same criticism which is raised against their approach. Tippett [38] describes his reservations concerning the above formulation in the following words: "one procedure postulates rather narrow-minded producers and customers whose imagined interests are protected independently. It is in the interest of the producer to ensure only that an undue proportion of satisfactory batches are not rejected; he does not mind how many satisfactory batches are accepted. Likewise, it is in the interest of the consumer to ensure only that an undue proportion of unsatisfactory batches are not accepted; he does not mind how many satisfactory batches are rejected. The producer and consumer so described are notional; no producer or consumer in real life can define his interests so narrowly." In practice, it is also hard to determine the two points described above and the associated risks in a rational manner. To a large extent, therefore, any sampling plan derived from these considerations has to be regarded as somewhat arbitrary.
Indifference Quality Level

Instead of determining the sampling plan from two points on the O.C. curve as described above, Hammaker [22] has proposed to use one point, the "point of control" or "indifference quality level" corresponding to an acceptance probability of $\frac{1}{2}$ and the relative slope of the curve at this point. From a computational point of view, this method of specifying the O.C. curve is often just as satisfactory as the previous one but in practice it is just as difficult to specify logically the indifference quality level and the slope as the producer's and consumer's risks.

Dodge-Romig Point of View

Dodge and Romig ([14], [15]) in their work have confined attention to rectifying sampling inspection plans, i.e., those schemes where rejected lots are to be completely inspected and all defective items found replaced by good ones. Thus their scheme is limited to non-destructive sampling only. The basic idea in the Dodge-Romig system is to choose a sampling plan which minimizes expected total inspection costs for normal production and at the same time guarantees required consumer protection of which the two following different kinds are envisaged:

(1) Lot Quality Protection in which some chosen value of the lot tolerance proportion defective is given and a value for the probability of accepting a submitted lot that has a percent defective equal to the LTPD is specified.
(2) **Average Quality Protection** in which some chosen value of average percent defective in the production after inspection is specified that will not be exceeded in the long run no matter what may be the level of fraction defective in the product submitted for inspection.

Dodge and Romig have developed plans guaranteeing each of the above two kinds of consumer protection. If \( N \) is the size of the lot, \( n \) is the number of items inspected, \( P_a(p) \) is the probability of accepting the lot when the fraction defective in the lot is \( p \) and the cost of inspection per unit = 1, and the total cost of inspection is \( I \), then \( I \) is a random variable and

\[
E\{I\} = n + (N - n)(1 - P_a).
\]

It is evident that \( I \) besides being a function of the parameter of the sampling plan is a function of \( p \), the "true" fraction defective of the lot. According to their system, we try to minimize \( E\{I\} \) where \( E\{I\} \) is computed with respect to the process average (or the normal expected quality) subject either to the constraint that

\[
P_a(p_t) = \beta \quad \text{where} \quad p_t \quad \text{denotes the lot tolerance percent defective and} \quad \beta \quad \text{(usually taken to be 0.10) is the required consumer's risk or} \quad P_L < k \quad \text{where} \quad P_L \quad \text{is the maximum of the average quality level after inspection (the so-called Average Outgoing Quality) and} \quad k \quad \text{is a specified constant lying between 0 and 1.}
\]

It is to be emphasized that the average amount of inspection under the above scheme is minimized corresponding to the expected quality. A satisfactory estimate of this level of quality can probably be obtained by reviewing data for the past period during
which normal conditions existed, and by utilizing other relevant information bearing on the manufacturing performance under the present conditions. If this level were constant, the system would be in statistical control and inspection would have been rendered unnecessary. But practically, while such a level may be adhered to most of the time, spasmodic changes in the product quality are bound to occur. When such changes occur, the sampling scheme automatically increases or decreases the inspection by an amount depending on the degree of deterioration or improvement of quality. If quality degenerates, inspection costs increase and provide an incentive for the elimination of the causes of trouble; on the other hand, if the level of quality improves, inspection schedules may be relaxed.

The system proposed by Dodge and Romig is subject to the following criticism. Instead of taking all cost elements into consideration, they have limited themselves to the cost of inspection only. The loss due to accepting defective items has not been taken care of explicitly. Further, the choice of the LTPD and the associated consumer's risk or the Average Outgoing Quality Limit remains more or less arbitrary.

**Minimax Principle**

If the lots submitted for inspection come from a system which is in a state of statistical control, it is likely that the number of defectives in a lot will have an apriori distribution. Often, however, the apriori distribution is unknown and decisions have to be made by statisticians in ignorance of it. The minimax principle
has been used by certain authors, e.g., Moriguti [31], Breakwell [8], etc. to cope with the lack of knowledge of the apriori distribution of the number of defectives. Assuming certain values for the cost parameters, one can compute the risk associated with using a certain procedure $\delta$ for different values of $p$, the fraction defective. Let this quantity be denoted by $r(p; \delta)$. One then chooses that plan which minimizes $\max_p r(p; \delta)$. Various criticisms have been raised against this approach also. First the value of $p$ corresponding to the minimax solution may well lie outside the range of values of $p$ which usually occur and therefore may not be particularly meaningful for a given problem. The second objection is the usual one -- namely that it is overly conservative.

Numerous authors have suggested various modifications of these concepts and formulated single, multiple or sequential sampling schemes that take them into account. It is not our purpose to go through them here in detail. For a discussion of the basic principles of sampling by attributes, the reader is referred to Hammaker [22] and for a list of the literature on this subject, he is referred to Horsnell [34].

1.2. The Sampling Inspection Model due to Hald

In this section, we discuss the sampling inspection model due to Hald [21] which we adopt for our study in the later chapters. Hald acknowledges that the above-mentioned sampling plans are in widespread use and have proved to be of great practical value. Nevertheless, he considers all of them to be somewhat arbitrary in character and takes the view that none of them fully takes into
account the economic implications of the size of the lot and the consequences of rejection or acceptance of the lot. Very often in deriving the appropriate sampling plans, he also remarks, one uses the Poisson or the binomial approximation to the hypergeometric distribution so that the parameters obtained for the sampling plans become independent of the lot size. This practice, which is in common usage, tends to disregard some of the costs associated with the size of the lot and, according to Hald, is therefore not very satisfactory. He also mentions that all sampling plans and especially those which we have described above (excepting the minimax procedures) implicitly assume the existence of a prior distribution. For example, he feels that in the Dodge and Romig scheme, the authors are considering a situation in which the apriori distribution is the weighted average of two distributions, one of which is binomial and the other is a discrete distribution which assigns higher probability to the larger numbers of defectives than does the binomial; the form of the discrete distribution and the relative weights assigned to it and to the binomial distribution are not known. The one parameter in their system is the process average \( \bar{p} \) in the binomial part of the prior distribution; the other parameter is the consumer's risk point (the LTPD) which is usually associated with a probability of 0.10. It appears that it is Hald's basic thesis that a more appropriate way of proceeding is to make some reasonable assumptions concerning the apriori distribution and then perform a formal Bayesian analysis. The approach advocated by Hald not only employs an apriori distribution which is completely known but also takes into account various economic factors associated with sampling inspection.
In what follows we shall describe in some detail the elements which enter into his model.

1.2.1. Costs of Sampling Inspection.

According to Hald, the basic question while deciding on the subject of costs is: What happens to rejected lots and what happens to defective items in either accepted or rejected lots. In the case of non-destructive sampling, Hald enumerates certain possibilities which we reproduce here.

```
Rejected lots
  └── Sorted └── Not Sorted
      └── Effective items └── Defective items
          └── Scrapped └── Sold at reduced price └── Returned to supplier
                          └── Repaired or replaced by effective
                                            └── Not repaired
                                                └── Scrapped └── Sold at reduced price
```

a. Costs due to accepting a defective item

Hald uses the average loss caused by an accepted defective item as the economic unit. It is assumed that defective items in accepted lots cause some damage which is measurable in economic terms. If the items under consideration are to be used in further production, the
loss by accepting a defective item may consist of the price paid per item, costs of handling and identifying the defective item plus possible costs of rework. If, however, the items represent finished goods, the loss incurred by the acceptance of the defective item may involve service and replacement costs plus loss of goodwill which, in general, will be difficult to measure.

b. Costs due to rejected lots

It is seen from the above diagram that there exist many possibilities for dealing with rejected lots. Hald assumes in his model that the cost associated with rejected lots (after sampling) are proportional to the number of items remaining in the lot. Thus if \( N \) is the lot size, \( n \) the sample size, the costs due to rejection of a lot will be equal to \((N - n) k_r\), where \( k_r \) is a constant to be determined. This constant of proportionality \( k_r \) is determined as follows in different situations:

1. In case of the sorting situation (non-destructive sampling), \( k_r \) is found as the sorting costs per item divided by the costs of accepting a defective item.

2. In case of destructive sampling which precludes sorting, \( k_r \) is found as the manufacturing costs (or market price) per item divided by the costs of accepting a defective item.

c. Costs due to sampling inspection

The last element of cost considered by Hald is the cost of inspection and is denoted by \( k_s \). In terms of the unit of cost which he assumes, it is evaluated as the sampling and testing costs per item divided by the costs of accepting a defective item.
Several remarks will be in order before we point out some of the criticisms which have been levelled against this model.

**Remark 1**

If \( p \) is the fraction defective and \( N \) is the lot size, then the expected loss due to acceptance of the lot without any inspection is \( Np \); on the other hand, loss due to rejection of the entire lot without inspection is \( Nk_r \). Thus \( k_r \) defines a break-even quantity in the sense that for \( p < k_r \), it will be cheaper to accept without inspection whereas for \( p > k_r \), the opposite will be true.

**Remark 2**

It is expected that in most practical situations the cost element \( k_r \) will be less than or equal to \( k_s \). In the sorting case this is obvious since \( k_r \) includes the sampling costs in addition to the costs due to testing. For destructive sampling also this relation will be true since \( k_s \) besides manufacturing costs also contains the costs of sampling and testing.

Before leaving the discussion of the subject of costs, we should point out that relatively simple and realistic extensions of this model are possible for which the same analytical techniques can be used. One such linear model has been suggested by Hald [22] himself and another has been given by Guthrie and Johns [19]. The latter authors have suggested a scheme in which six elements of cost are considered instead of the three considered by Hald. In their notation, these costs are as follows:

(1) Costs due to acceptance of the \( i \)-th item without inspection

\[ = a_1 x_i + a_2 \]
(2) Costs due to rejection of the $i$-th item without inspection

\[ = r_1 x_i + r_2 \]

(3) Cost of inspection of the $i$-th item

\[ = s_1 x_i + s_2 \]

In case of sampling by attributes, $X_i = 1$ or 0 according as the item is defective or non-defective. $s_1$ represents the cost of replacing a bad item by a good one and $s_2$ the cost of testing the item. In this model if it is assumed that cost of accepting an item = $a_1$ if it is defective and = 0 otherwise, then $a_2 = 0$. Similarly if it is postulated that the cost of rejecting the uninspected remainder of the lot is proportional to the number of defective items remaining in the lot, then $r_1 = 0$. Similar formulations of the cost parameter have been given by Moriguti [31], Breakwell [8], and Vaghoklar [38]. Hald has shown that the expression for the expected cost in the model considered by Guthrie and Johns [20], presumably also for the other linear models, can be reduced very simply to a linear function of that of Hald's model. From the point of view of determination of the optimal plan, the analytical expressions are very similar.

1.2.2. Some Criticisms of Hald's Model.

Many different objections have been set forth against the model proposed by Hald. Barnard [4] has pointed out that in real life the cost functions which are taken as constant in Hald's model do not in fact remain so, and may be extremely variable depending on various factors, e.g., the demand for goods being inspected, storage capacity,
raw material position, quality of other component items, etc. In addition to this, these costs are very often hard to evaluate precisely. It also appears that the traditional concepts associated with sampling plans, e.g., producer's and consumer's risks may be more readily understood and specified by management than the rather sophisticated cost concepts introduced in Hald's model. Anscombe [2] states yet another objection. In his opinion, one should regard the cost of sampling inspection as a function of the quality of the lot as well as of the size of the sample. His remark is probably appropriate when the testing is destructive because in that situation the value of the net product will be reduced more by inspecting an item if the quality were good than if it were bad. It is not very clear, Anscombe's remark notwithstanding, how to formulate a cost function which takes into account this aspect of dependence on the unknown parameter. We shall consider this question further in Chapter II in the light of its implications on the construction of optimum sequential procedures.

Hald's model also assumes that the apriori distribution of the number of defectives is known to the statistician and that a formal Bayesian analysis can therefore be carried out. This point of view may also not be found acceptable to many statisticians. The controversy which currently exists between the Bayesian and Non-Bayesian groups of statisticians on this point is well known, and we do not wish to dwell on it in this dissertation.
We thus see that there exist many arguments for and against Hald's model of sampling inspection. Whether or not it can be recommended for applications in any given practical situation will probably depend on circumstances. We do not wish to pass judgment on the applicability of the model in real life situations.

In this thesis, we shall be mainly concerned with obtaining the optimal sequential plans for Hald's model with a special type of apriori distribution. Having obtained the plans we compare the average risks associated with the procedures with those of the optimal single and double sampling plans. In doing this, we chose certain representative values of the parameters involved and found that for these values, very little is gained by taking observations sequentially when the average risk for the Bayes sequential procedure is compared with the corresponding quantity for the optimal one-stage and two-stage plans. This may appear to contradict the impression one obtains from classical sequential analysis -- namely that sequential procedures are much superior to single-stage sampling plans. It should however be borne in mind that in classical statistics, the Bayesian point of view is not adopted, and the usual criterion for a comparison of the "goodness" of the single-stage and sequential plans is the ASN when both sets of plans guarantee certain preassigned values of probabilities of correct decisions. In Chapter II of this thesis, we also make a comparison of the ASN of the Bayes sequential plans with the optimum sample size of single-stage procedures for the same values of the cost and distribution parameters. Similar to the results given in classical sequential analysis, we also find here that the
ASN-function for the sequential plans is much lower than the optimum sample size of the one-stage procedures over a wide range of values of the total number of defectives in the lot.

1.2.3. Reproducible Distributions.

Although it is theoretically possible to obtain the optimum single, multiple or sequential plans and the related Bayes risks with respect to any given apriori distribution, Hald chooses mainly to deal with those distributions which are reproducible to hypergeometric sampling. It was shown by him that such distributions lead to particularly convenient solutions to the problem of determination of the optimal single sampling plans for his model.

There exist many such distributions which are reproducible to hypergeometric sampling, and Pfanzagl [33] chose from among them a two-parameter family which it appears is most convenient to deal with mathematically. In our thesis, we shall also be concerned with this particular family.

The family of distributions reproducible to hypergeometric sampling and its counterpart in the case of independently and identically distributed random variables which has been alternatively referred to as natural conjugate priors by Raiffa and Schlaifer [34], possesses some convenient closure properties which are very useful in sequential analysis. The choice of a distribution belonging to this particular class as the apriori distribution has the effect of reducing considerably the dimensions of the random walk associated with sequential plans. We give below a formal definition of the distributions reproducible to hypergeometric sampling and the natural conjugate priors.
Let \( \xi_N(\theta) \) denote the apriori probability that a lot of size \( N \) contains \( \theta \) defectives and let \( p_n(x|\theta) \) denote the conditional probability that a sample of size \( n \) from the lot contains \( x \) defectives when the lot contains \( \theta \) defectives. Let \( \theta^{(n)} = \theta - x \) denote the number of defectives remaining in the non-inspected part of the lot. Then the marginal distribution of \( x \), with respect to the apriori distribution \( \xi_N(\theta) \), i.e., the unconditional probability that a lot of size \( N \) will contain \( x \) defectives is given by

\[
(1.1) \quad g_n^c(x) = \sum_{\theta=x}^{N-n-x} \binom{\theta}{x} \binom{N-\theta}{n-x} \xi_N(\theta)
\]

\[
(1.2) \quad = \binom{n}{x} \sum_{\theta^{(n)} \geq 0} \xi_n(\theta) \frac{\binom{N-n-\theta^{(n)}}{\theta^{(n)}}}{\binom{N}{\theta}}
\]

\( g_n^c(x) \) is called the compound hypergeometric distribution since it is derived by averaging the hypergeometric distribution for given \( x \) over all possible values of \( \theta \) according to the apriori distribution.

**Definition**

If the compound hypergeometric distribution \( g_n^c(x) \) as defined by (1.1) or (1.2) is equal to \( \xi_n(x) \), then the apriori distribution \( \xi_N(\theta) \) is said to be reproducible to hypergeometric sampling.

The above definition is particularly relevant for single samples of size \( n \) which Hald has considered. In the context of sequential sampling also, this property holds as we naturally expect.

For a proof, see Lemma 1.1.
Hald has given a detailed characterization of the distribution reproducible to hypergeometric sampling for any given lot size $N$ and proves that a sufficient condition for a distribution to satisfy the above property is that it be a hypergeometric or a binomial or a rectangular or a Polya or a mixed binomial distribution or any weighted combination of these with weights independent of $N$ and $\theta$. Also, if the apriori distribution belongs to any of the above families, then for any $n$, the distribution of $\theta^{(n)}$ (the number of defectives remaining in the lot) for given $x$ is also of the same type but with parameters depending on $x$. Thus we observe that in the case of reproducible apriori distributions, the a posteriori distribution of $\theta^{(n)}$ given $x$ belongs to the same family as that of the apriori distribution of $\theta$.

Recently the use of family of prior distributions which possesses such nice closure properties has become very widespread in statistical analysis. When an apriori distribution is such that all posterior distributions after observing i.i.d. random variables $X_1, X_2, \ldots$ belong to the same family as that of the apriori distribution, it is called a natural conjugate prior (n.c.p.) to the distribution of $X_1$. This nomenclature is due to Raiffa and Schlaifer. Such a family of prior distributions has also been called closed under sampling (see Wetherill [44]). Raiffa and Schlaifer [34] have developed a theory for the natural conjugate priors, details of which can be found in their book. They do not, however, mention the distributions repro-
ducible to hypergeometric sampling presumably because in this case observations are not independent. Some examples of n.c.d. distributions are given as follows:

1. \( X - \text{Normal } N(\theta; 1) \)
   \[ \xi(\theta) \propto (1/b) e^{-(\theta - a)^2/2b^2} \]

2. \( X - \text{Binomial } (\theta) \)
   \[ \xi(\theta) \propto \theta^{a-1} (1-\theta)^{b-1} \]

3. \( X - \text{Poisson } (\theta) \)
   \[ \xi(\theta) \propto e^{-a\theta} \theta^{b-1} \]

J. Pfanzagl in his extension of Hald's work to double sampling plans has assumed in particular the prior distribution to be a Polya distribution which is given as follows:

\[
(1.3) \quad \xi_{N}(\theta; \alpha, \beta) = \binom{N}{\theta} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \Gamma(\beta)} \frac{\Gamma(\alpha + \theta) \Gamma(\beta + N - \theta)}{\Gamma(\alpha + \beta + N)}
\]

\[
= \binom{N}{\theta} \frac{B(\alpha + \theta, \beta + N - \theta)}{B(\alpha, \beta)}, \quad (\theta = 0, 1, 2, \ldots, N; \alpha, \beta \geq 1)
\]

where

\[
B(p, q) = \int_0^1 x^{p-1} (1-x)^{q-1} \, dx, \quad (p, q \geq 1).
\]

Pfanzagl [33] remarks that the use of the Polya distribution is widespread in the context of sampling plans because of the fact that it has two parameters (\(\alpha, \beta\)) and can therefore be conveniently used to approximate many types of empirical distributions. In evidence of this assertion, he gives examples where distributions belonging to the Polya family with different values of parameters serve as very close approximations to a bell-shaped curve and a J-shaped curve (see Figures 1 and 2, Pfanzagl [33]). Nevertheless, there do not exist any theoretical or empirical reasons to believe that a Polya distribution is better suited to fit any empirical
prior distribution than any other suitable family. As a matter of fact, published information on apriori data appears to be scarce, and it is very difficult to justify on solid grounds the choice of Polya distribution as the prior distribution. We shall, however, take the stand that the apriori information will probably be very rough and can be quantified in terms of a given member of this large family. This will in many cases be an approximation, perhaps gross on occasion, but we shall adhere to this assumption throughout this thesis mainly because of the resulting mathematical simplicity. In passing, we mention that the Polya distribution has also been used in conjunction with the hypergeometric distribution quite a few times in theoretical statistics (see for example, Hodges and Lehmann [24], where they obtain minimax point estimates for the number of defectives in a lot).

Below we list and prove some important facts about this family of distributions, which we shall use in Chapter II.

Lemma 1.1

Let the apriori distribution of the number of defectives in a lot of size \( N \) be given by

\[
\xi_N(\theta; \alpha, \beta) = \binom{N}{\theta} \frac{B(\alpha + \theta, \beta + N - \theta)}{B(\alpha, \beta)} \quad (\theta = 0, 1, 2, \ldots, N)
\]

Let \( X_i \ (i = 1, 2, \ldots, N) = 1 \) or 0 according as the \( i \)-th item sampled is defective or non-defective. Let \( \theta^{(j)} \) denote the number of defectives remaining in the non-inspected part of the lot after \( X_1, X_2, \ldots, X_j \) have been
observed, and let $\theta^{(0)} = \theta$. Then the a posteriori distribution of $\theta^{(j)}$ after observing $x_1 = x_1 (i = 1, 2, \ldots, j)$ belongs to the same Polya family and is given by

$$\xi_{N-j} \left( \theta^{(j)}; \alpha + \sum_{i=1}^{j} x_i, \beta + j - \sum_{i=1}^{j} x_i \right)$$

**Proof**

We give the proof for the a posteriori distribution of $\theta^{(1)}$ only. The proof for the a posteriori distribution of $\theta^{(j)}$ follows from a straightforward induction.

The conditional probability of observing $x_1 = x_1$ given that $\theta$ is the true number of defectives is given by

$$p(x_1 | \theta) = \frac{\binom{\theta}{x_1} \binom{N-\theta}{N-x_1}}{N}$$

$$= \begin{cases} 1 & \text{if } x_1 = x_1, \ldots, N-1+x_1, \\ 0 & \text{otherwise} \end{cases}$$

where $x_1 = 0$ or $1$.

The joint distribution of $\theta$ and $x_1$ is therefore given by

$$p(X_1 = x_1, \theta) = \frac{\binom{\theta}{x_1} \binom{N-\theta}{N-x_1}}{N} \xi_N (\theta; \alpha, \beta)$$

$$= \begin{cases} 1 & \text{if } x_1 = x_1, \ldots, N-1+x_1, \\ 0 & \text{otherwise} \end{cases}$$

The marginal distribution of $X_1$ when summed over all possible values of $\theta$ is
\[ p(x_1 = x_1) = \sum_{\theta = x_1}^{N-1+x_1} \frac{\binom{\theta}{x_1} \binom{N-\theta}{1-x_1}}{N} \xi_N(\theta; \alpha, \beta) \]

\[ = \sum_{x_1}^{N-1+x_1} \frac{\theta!}{x_1!(\theta-x_1)!} \frac{(N-\theta)!}{(1-x_1)! (N-\theta-1-x_1)!} \frac{N!}{\theta! (N-\theta)!} \frac{B(\alpha+\theta, \beta+N-\theta)}{B(\alpha, \beta)} \]

\[ = \frac{1}{B(\alpha, \beta)} \sum_{x_1}^{N-1+x_1} \frac{\theta!}{x_1!(\theta-x_1)!} \frac{(N-\theta)!}{(1-x_1)! (N-\theta-1-x_1)!} \frac{N!}{\theta! (N-\theta)!} \int_0^\theta (\alpha+\theta-1)(1-p)^{\beta+N-\theta-1} dp \]

\[ = \frac{1}{B(\alpha, \beta)} \int_0^\theta \frac{1}{x_1!(1-x_1)!} \frac{\alpha+x_1-1}{p} \frac{\beta-x_1}{(1-p)} \left[ \sum_{x_1}^{N-1+x_1} \frac{(N-1)!}{(\theta-x)! [(N-1-\theta-x)!]} \right] \]

\[ = \left( \frac{1}{x_1!} \right) \frac{1}{B(\alpha, \beta)} \int_0^\theta \frac{\alpha+x_1-1}{p} \frac{\beta-x_1}{(1-p)} dp \]

(1.6) \[ = \left( \frac{1}{x_1!} \right) \frac{B(\alpha+x_1, \beta+1-x_1)}{B(\alpha, \beta)} = \xi_1(x_1; \alpha, \beta) . \]

This shows that the Pólya distribution is reproducible to hypergeometric sampling. The a posteriori distribution of \( \theta \) is now given by
\[ p(\theta | x_1 = x_1) = \frac{\binom{\theta}{x_1} \binom{N-\theta}{\theta-x_1})^{N} \theta(\alpha, \beta)}{\text{B}(\alpha; \beta), \text{B}(\beta+1-x_1), \ldots, \text{B}(N-1+x_1)}. \] (1.7)

Therefore, the a posteriori distribution of \( \theta(1) = \theta-x_1 \), is given by

\[ p(\theta(1) | x_1 = x_1) = \frac{\binom{N-1}{\theta(1)} \text{B}(\alpha+\theta(1)+x_1, \beta+N-1-\theta(1)+1-x_1)}{\text{B}(\alpha+x_1, \beta+1-x_1)} \] (1.8)

The rest of the lemma is now proved by a straightforward induction.

**Lemma 1.2.**

The first two moments of \( \xi_{\theta} \) are given by

\[ E(\theta) = \frac{N \alpha}{\alpha + \beta}. \] (1.9)
\begin{align*}
(1.10) \quad \text{Var}(\theta) &= \frac{N \alpha \beta}{(\alpha+\beta)(\alpha+\beta+1)} + \frac{N^2 \alpha \beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.

\text{Proof} \quad \text{E}(\theta) &= \frac{1}{B(\alpha,\beta)} \sum_{\theta=0}^{N} \binom{N}{\theta} B(\alpha+\theta, \beta+N-\theta) \\
&= \frac{1}{B(\alpha,\beta)} \sum_{\theta=0}^{N} \binom{N}{\theta} \int_0^1 p^{\alpha+\theta-1}(1-p)^{\beta+N-\theta-1} \, dp \\
&= \frac{1}{B(\alpha,\beta)} \int_0^1 p^{\alpha-1}(1-p)^{\beta-1} \left[ \sum_{\theta=0}^{N} \binom{N}{\theta} p^\theta (1-p)^{N-\theta} \right] \, dp \\
&= \frac{N}{B(\alpha,\beta)} \int_0^1 p^\alpha (1-p)^{\beta-1} \, dp \\
&= \frac{NB(\alpha+1, \beta)}{B(\alpha, \beta)} = \frac{N \alpha}{\alpha + \beta}.
\end{align*}
Similarly, we have

\[
E(\theta^2) = \frac{N^2 \alpha (\alpha + 1)}{(\alpha + \beta)(\alpha + \beta + 1)} + \frac{N \alpha \beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}
\]

so that

\[
\text{Var}(\theta) = \frac{N \alpha \beta}{(\alpha + \beta)(\alpha + \beta + 1)} + \frac{N^2 \alpha \beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}.
\]

If we have any apriori information about the mean and variance of \(\theta\), it can be used to find the particular Polya distribution to be fitted, using (1.9) and (1.10), by the Method of Moments.

Another result which we shall need later is given below.

**Lemma 1.3**

Suppose that a lot of size \(N\) has \(\theta\) defectives where \(\theta\) has an apriori distribution given by \(\xi_N(\theta; \alpha, \beta)\). Then the marginal probability of an item sampled being defective is equal to \(\frac{\alpha}{\alpha + \beta}\).

**Proof**

We have

\[
p(X_1 = 1 \mid \theta) = \frac{\theta}{N}.
\]

Let the marginal probability that \(X_1 = 1\) when the apriori distribution of \(\theta\) is \(\xi_N(\theta; \alpha, \beta)\) be denoted by \(P_\xi(X = 1)\). Then we have from (1.9) that

\[(1.11) \quad P_\xi(X = 1) = \frac{1}{N} E(\theta) = \frac{\alpha}{\alpha + \beta}\]

and therefore

\[(1.12) \quad P_\xi(X = 0) = 1 - P_\xi(X = 1) = \frac{\beta}{\alpha + \beta}.
\]
Other important properties of the Polya distribution can be found in Hald [22] and in Feller [17].

In the sequel, we shall refer to the above model of sampling inspection when the apriori distribution belongs to the Polya family as Pfanzagl's model.

We end this section with the following remark. We do not know whether this model considered by Hald will eventually be applied in practical situations. If it is used, care should be exercised to incorporate the changes of the parameters of cost or those of the apriori distributions in the various sampling schemes. Without such updating of parameters, faulty conclusions may result frequently.

1.3. A Decision Theoretic Formulation for Pfanzagl's Model of Sampling Inspection, and Statement of our Problem

In this section we write down the decision theoretic formulation for the sampling inspection model due to Hald as adopted by Pfanzagl.

(1) Model

Let \( X_i \) (\( i=1,2,...,N \)) equal 0 or 1 according as the \( i \)-th observation taken without replacement from a lot of size \( N \) containing \( \theta \) defectives, is non-defective or defective. Then \( X_i \) has the following probability law:

\[
P(X_i = x_i) = \frac{\theta^{(i-1)} x_i^{x_i} (N-i-\theta^{(i-1)} + 1)^{1-x_i}}{N - i + 1}, \quad (\theta^{(i-1)} = x_i, \ldots, N-i+1-x_i)
\]

= 0 otherwise
(2) Parameter Space \( (\Omega) \)

\[ \Omega = \{ \theta | \theta = 0, 1, 2, \ldots, N \} \]

(3) Action Space

There are two decisions \( a_1 \) and \( a_2 \). \( a_1 \) corresponds to
acceptance of lot; \( a_2 \) corresponds to rejection of lot.

(4) Costs per observation

In general as noted earlier, the cost of an observation may
depend on the unknown parameter \( \theta \) and the observations. In
recognition of this possibility, we denote the cost function per
observation as \( c(\theta; x) \). In this particular case

\[ (1.13) \quad c(\theta; x) = k_s \]

(5) Terminal Losses

From the foregoing discussion of the model proposed by Hald,
it is apparent that the loss incurred at any stage depends on the
observed outcome. Let the loss incurred due to taking action \( a_i \)
at the \( j \)-th stage after observing \( x_1, x_2, \ldots, x_j \) when \( \theta \) is the
true parameter be denoted by

\[ L_j(\theta; a_i) = L(\theta; a_i; x_1, x_2, \ldots, x_j), (i=1,2; j=1,2,\ldots,N; \theta \in \Omega) \]

For consistency of notation, we denote the loss incurred
by taking action \( a_1 \) without making any observation as \( L_0(\theta; a_1) \).
Then according to the situation envisaged by Hald,

\[ (1.14) \quad \begin{align*}
L_j(\theta; a_1) &= \varepsilon(j) \\
L_j(\theta; a_2) &= (N-j) k_r
\end{align*} \quad \left( \theta \in \Omega, j = 0,1,2,\ldots,N \right) \]
We have already made the convention that $\theta^{(0)} = \theta$ (see Lemma 1.1) and our notations are therefore consistent.

The overall loss function, if a decision is made to stop at the \textit{j}-th stage, is therefore $\theta^{(j)} + jk_{\alpha}$ if $a_1$ is chosen and $(N-j)k_{\alpha} + jk_{\beta}$ if $a_2$ is chosen.

(6) \textbf{Apriori distribution on } $\Omega$

The apriori distribution on the parameter space $\Omega$ belongs to the Polya family given by

\begin{equation}
\xi_N(\theta, \alpha_0, \beta_0) = \binom{N}{\theta} \frac{B(\alpha_0 + \theta, \beta_0 + N - \theta)}{B(\alpha_0, \beta_0)} \quad [\theta = 0, 1, 2, \ldots, N]
\end{equation}

For simplicity, $\alpha_0$ and $\beta_0$ will be taken to be integers.

Hald considered the above problem and gave optimum procedures which, in the class of all single sample plans, gives the minimum average risk. Pfanzagl has given double sampling inspection plans which are optimum in a similar sense. Our object in this thesis will be to construct the Bayes sequential procedures for the above model. The desirability of such an extension to sequential sampling plans has been pointed out by Hald himself (see [22]) and Anscombe [2].

It turns out that a rather simple characterization is possible for the Bayes sequential procedures for the above model. This is mainly because of the fact that the apriori distribution has been chosen to belong to a family which is reproducible to hypergeometric sampling. In this connection, we mention that the mathematical convenience resulting from the use of the family of the n.c.p. distributions in sequential analysis has been demonstrated recently by Wetherill [44], Mikhailevich [30], Moriguti and Robbins [32], and
Lechner [29]. Their analysis has been confined mainly to the case of independently and identically distributed random variables. The mathematical simplicity which we obtain here is comparable to that obtained by them in the problems which they consider.

1.4. Review of Some Sequential Sampling Plans by Attributes and Earlier Relevant Work.

The sequential sampling plan which is perhaps most familiar in industry is the one based on Wald's sequential probability ratio test (WSPRT) [40]. For this test, two points on the OC-curve are specified. It is also assumed that the lot size is indefinitely large so that the binomial approximation to the hypergeometric distribution is valid.

Let the producer's and consumer's risk points be denoted by \( \theta_{-1} \) and \( \theta_1 \), respectively, and let \( 1 - \alpha \) and \( \beta \) be the associated probabilities of accepting the lot. The particular WSPRT is chosen which guarantees that

\[
P_{\theta_{-1}} \{\text{lot is accepted}\} \geq 1 - \alpha
\]

\[(1.16)\]

\[
P_{\theta_1} \{\text{lot is accepted}\} \leq \beta.
\]

It turns out that due to the monotonicity properties of the OC-curve associated with the WSPRT as applied to the parameter \( \theta \) of a binomial distribution, the WSPRT which guarantees (1.16) also guarantees the following:

\[
P_{\theta} \{\text{lot is accepted}\} \geq 1 - \alpha \quad \text{if } \theta \leq \theta_{-1}
\]

\[(1.17)\]

\[
P_{\theta} \{\text{lot is accepted}\} \leq \beta \quad \text{if } \theta \geq \theta_1.
\]
Further, Wald's plan has the important property that among all plans satisfying (1.17), whether sequential or not, the expected sample size is a minimum at $\theta_l$ and $\theta_u$. Recently, Weiss [42], and Freeman and Weiss [18] have given some sequential sampling plans which minimize the maximum expected sample size among all plans satisfying (1.17). Barnard [5] and Anscombe [1] have also described simple sequential plans based on some scoring systems, which do not possess any such optimum property but are perhaps administratively easier to handle. These sampling plans confine themselves to a Neyman-Pearson theoretic formulation and as such differ from the decision-theoretic formulation of minimizing the risk associated with a plan where the different economic consequences of acceptance or rejection are taken into account. Chernoff and S offering [13] state that such an approach was first taken by Wurtele [45] in her unpublished thesis. Champernowne [10] also analyzed the economics of sequential sampling plans by assuming large lots and postulating the existence of a break-even quantity $p_0$ for the proportion defective, and assuming a linear loss structure. If the acceptance of a lot is denoted by $a_1$ and the rejection by $a_2$, his loss function is

$$L(p; a_1) = \max \left[ 0, k(p-p_0) \right]$$

$$L(p; a_2) = \max \left[ 0, k(p_0-p) \right]$$

where $k$ is a positive constant and the cost per observing an item = 1. He assumed further that the apriori distribution belongs to the Beta family. His results appear to be approximate and it is not clear how he computes the boundaries for his sequential decision rules. Recently,
Moriguti and Robbins [32] gave a detailed analysis for the above problem when \( p_0 = \frac{1}{2} \), and Ray [35] extended their results to any given value of \( p_0 \), and showed how to compute the optimum boundaries for the sequential decision problems in this case. Similar formulations of this problem have also been considered by Moriguti [31] and Breakwell [8], but they do not postulate the existence of an apriori distribution; they have instead found the sampling procedures which minimize the maximum risk associated with a plan. Chernoff and Ray [13] have considered Bayes sequential plans with a formulation similar to that of ours but they have restricted their attention to one terminal action only, namely acceptance of the lot. Our work here can therefore also be regarded as a generalization of their results to the two-action case.

The works reviewed above are those which to some extent are relevant to our problem; it is not possible to include all the major papers connected with either sampling inspection or sequential analysis. For a review of the present state of affairs in the subject of sequential analysis, the reader is referred to Johnson [26] and Jackson [27].

1.5 Outline of the Thesis

Chapter 1

The basic problem of sampling inspection is stated. Existing sampling inspection plans within the framework of Neyman-Pearson theory are described briefly and their main defects and arbitrariness pointed out. Dodge and Romig's point of view is explained and the sampling inspection plans proposed by them are described. It is
pointed out that these plans do not take into account fully the economic consequences of rejection or acceptance of a lot and the economic implication of the size of a lot. Further all these plans implicitly assume the existence of an apriori distribution. Hald's model which purports to remove all the deficiencies is next described in detail. The concept of reproducible apriori distributions is introduced and the problem is then formulated as a decision-theoretic one within the framework of Wald's theory. The basic problem of the thesis is stated and existing sequential plans reviewed.

Chapter 2

The Bayes sequential procedures for Pfanzagl's model are characterized and it is shown that the usual characterization due to Wald and Wolfowitz [41] in the set-up of independent and identically distributed random variables and "stage-independent" loss function also applies in our decision problem although the random variables under question are dependent and not identically distributed, and the loss at each stage depends on the observed outcome. We describe how to obtain the Bayes sampling rule and hence the optimum (Bayes) boundary. It is shown that the region of sampling may be narrowed considerably, and an upper bound to the point of truncation for the Bayes procedure is given. The relations for obtaining the operating characteristics of the Bayes sequential sampling plan are derived. These and the Bayes risks are compared with those of the optimal single and double sampling plans.
Chapter 3

The limiting behavior in large samples of the optimum boundaries for the decision problem studied in Chapter 2 is considered. Extending Chernoff and Ray's [13] approach the problem of finding the normalized optimal boundaries with appropriate normalization is reduced to that of solving certain partial differential equation free boundary problems. It turns out that the solution for this problem in our particular case give rise to somewhat unexpected results. The implication of these answers is considered in some detail.
CHAPTER II

SEQUENTIAL BAYES PROCEDURES WITH
PFANZAGL'S MODEL OF SAMPLING INSPECTION

In this chapter we describe how to obtain the Bayes sequential procedures for the decision problem described in Section 1.3. We characterize these procedures in some detail and give a description of the optimal sampling plans for some representative values of the parameters involved. The operating characteristics for these plans are also obtained for these values and compared with those of the optimal single sampling plans.

2.1. Characterization of the Bayes Sequential Procedures with Pfanzagl's Model

Since according to our assumptions the lot is finite containing $N$ items, and the sampling is without replacement, the number of potential observations is bounded and any sequential procedure for this problem is necessarily truncated. It is well known how to construct the Bayes procedures, within a class of procedures truncated at a fixed stage, by an application of the so-called "working backwards" technique. Blackwell and Girshick [7] describe how to obtain such procedures for quite general cost and loss structures. But their method of construction is rather complicated; here we shall obtain a simpler characterization, similar to the one which was derived by Wald and Wolfowitz [41]. In our case, unlike most of those treated in the literature, successive observations are dependent and not identically distributed. Furthermore, our loss depends at each stage on the previous observations. In the sequel we shall derive,
in some detail, the Bayes stopping and terminal decision rules for our problem. In so doing, we closely follow a similar development by Ray [35] for a special type of "modified loss" structure. We retain his notation insofar as it is consistent with our problem.

2.1.1. Terminal Decision Rule for the Bayes Procedure

Let $\mathcal{S}$ be any decision rule. Let $\psi_0$ denote the probability that $\mathcal{S}$ does not take any observations, and let $\psi_j = \psi_j(x_j) \quad (j = 1, 2, \ldots, N)$ denote the probability that $\mathcal{S}$ terminates at the $j$-th stage after observing $X_i = x_i \quad (i = 1, 2, \ldots, j)$. Let $\phi_0$ denote the probability that $\mathcal{S}$ chooses $a_2$ (i.e., rejects the lot) without taking any observations, and similarly let $\phi_j = \phi_j(x_j) \quad (j = 1, 2, \ldots, N)$ be the probability that $\mathcal{S}$ takes action $a_2$ after observing $X_i = x_i \quad (i = 1, 2, \ldots, j)$ conditional upon $\eta = j$, where $\eta$ is the sample size (a random variable). Let $\mathcal{S}_{\xi} = (\psi_{\xi,0}, \psi_{\xi,1}, \psi_{\xi,2}, \ldots, \psi_{\xi,N}; \phi_{\xi,0}, \phi_{\xi,1}, \ldots, \phi_{\xi,N})$ denote the Bayes decision rule for our problem. Here, $\psi_{\xi,0}$ denotes the probability that $\mathcal{S}_{\xi}$ will not take any observations, and $\phi_{\xi,j}$ the conditional probability that it will choose $a_2$ given that it takes no observations; also, $\psi_{\xi,j} = \psi_{\xi,j}(x_j) \quad (j = 1, 2, \ldots, N)$ denotes the probability, when $\mathcal{S}_{\xi}$ is used, that $\eta = j$ when $X_i = x_i \quad (i = 1, 2, \ldots, j)$ have been observed, and similarly, $\phi_{\xi,j} = \phi_{\xi,j}(x_j)$ is the probability of choosing $a_2$ after observing $X_i = x_i \quad (i = 1, 2, \ldots, j)$ conditional upon $\eta = j \quad (j = 1, 2, \ldots, N)$.

The terminal decision rule for the Bayes procedure is now given in the following lemma.
Lemma 2.1

If the Bayes sequential procedure for the decision problem of Section 1.3 stops after observing $X_i = x_i$ ($i = 1, 2, \ldots, j$), then the terminal decision function $\varphi_{x, j}$ ($j = 1, 2, \ldots, N$) is given by equation (2.1); if it does not take any observations, the terminal decision function $\varphi_{x, 0}$ is given by equation (2.2).

\begin{align*}
(2.1) \quad \varphi_{x, j} &= 0 \quad \text{if} \quad (\alpha_0 + \frac{\sum_{i=1}^j x_i}{\alpha_0 + \beta_0 + j}) < k_r \\
&\qquad \text{if} \quad (\alpha_0 + \frac{\sum_{i=1}^j x_i}{\alpha_0 + \beta_0 + j}) = k_r \\
&\qquad 1 \quad \text{if} \quad (\alpha_0 + \frac{\sum_{i=1}^j x_i}{\alpha_0 + \beta_0 + j}) > k_r
\end{align*}

\begin{align*}
(2.2) \quad \varphi_{x, 0} &= 0 \quad \text{if} \quad \frac{\alpha_0}{\alpha_0 + \beta_0} < k_r \\
&\qquad \text{if} \quad \frac{\alpha_0}{\alpha_0 + \beta_0} = k_r \\
&\qquad 1 \quad \text{if} \quad \frac{\alpha_0}{\alpha_0 + \beta_0} > k_r
\end{align*}

In the above, $q$ is any arbitrary number between 0 and 1.

Let $r_N (\theta; \delta)$ be the risk associated with $\delta$ when $\theta$ is the true parameter and $N$ is the lot size. We now introduce the following additional notation:

$x_j$ : the $j$-vector $(x_1, x_2, \ldots, x_j)$, $\quad (j = 1, 2, \ldots, N)$. 
\( x_j \) : the space consisting of all possible \( x_j \), \( j = 1, 2, \ldots, N \).

\( p_{\theta,j} = p_{\theta,j}(x_j) = p(X_1 = x_1, X_2 = x_2, \ldots, X_j = x_j | \theta), \)
\( j = 1, 2, \ldots, N \).

: the conditional probability of \( X_i = x_i \) \( (i = 1, 2, \ldots, j) \)
when \( \theta \) is the true parameter.

\( p_{\theta,0} = 1 \), by definition.

\( \pi(\theta) \) : an apriori distribution on the unknown states of Nature; in
our problem, \( \pi(\theta) = \pi_N(\theta; \alpha_0, \beta_0) \).

\( p_{\pi,j} = p_{\pi,j}(x_j) = \sum_{\theta} p_{\theta,j}(\theta) = \sum_{\theta} p_{\theta}(x_j) \pi_N(\theta; \alpha_0, \beta_0), \)
\( j = 1, 2, \ldots, N \).

: the marginal probability of \( X_i = x_i \) \( (i = 1, 2, \ldots, j) \)
computed with respect to the prior distribution \( \pi(\theta) \).

\( \tau_{x_j}(\pi(\theta)) \) : the a posteriori distribution of \( \theta \) after \( X_i = x_i \)
\( (i = 1, 2, \ldots, j) \) have been observed.

Therefore, \( \tau_{x_j}(\pi(\theta)) = \frac{p_{\theta,j}(\pi(\theta))}{p_{\pi,j}} \), \( j = 1, 2, \ldots, N \).

\( \tau_{x_j}(\pi(\theta)(j)) \) : the a posteriori distribution of \( \theta(j) \) after \( X_i = x_i \)
\( (i = 1, 2, \ldots, j) \) have been observed, \( (j=1, 2, \ldots, N) \).

**Proof**

It is readily seen from the above definitions that

\( (2.3) \quad x_N(\theta; \delta) = \text{Expected loss} + \text{Expected cost of observations} \)
\[ r_N(\xi; \delta) = \psi_0 \left[ (1 - \varphi_0) \theta + \varphi_0 (Nk_r) \right] + \sum_{j=1}^{N} \sum_{\theta=0}^{\alpha_0} \psi_j \left[ (1 - \varphi_j) \theta(j) + \varphi_j (N - j) k_r \right] + jk_s \right] p_{\theta,j} \xi_N(\theta; \alpha_0, \beta_0) \]

The expected risk associated with \( \delta \), when \( \xi_N(\theta; \alpha_0, \beta_0) \) is the apriori distribution of \( \theta \), is given by

\[ (2.4) \]

Now denote \( \sum_{i=1}^{j} x_i \) by \( z_j \). From Lemma 1.1, it follows that

\[ r_N(\xi; \delta) = \psi_0 \left[ (1 - \varphi_0) \frac{N\alpha_0}{\alpha_0 + \beta_0} + \varphi_0 (Nk_r) \right] + \sum_{j=1}^{N} \sum_{\theta=0}^{\alpha_0} \psi_j \xi_N(\theta; \alpha_0, \beta_0) \]

Recalling that \( \tau_{x_j} \xi(\theta(j)) = \xi_{N-j}(\theta(j); \alpha_0 + z_j, \beta_0 + j - z_j) \), we obtain
\[(2.5) \quad r_N (\xi; \delta) = \psi_0 \left[ (1 - \varphi_0) \frac{\frac{N \alpha_0}{\alpha_0 + \beta_0} + \varphi_0 (N k_r)}{\varphi_j (N - j) \frac{\alpha_0 + \beta_0}{\alpha_0 + \beta_0 + j}} + \varphi_j (N - j) k_r + j k_s \right]^\psi_j, j, \]

and hence

\[(2.6) \quad r_N (\xi; \delta) \geq \psi_0 \left[ \min \left( \frac{\frac{N \alpha_0}{\alpha_0 + \beta_0}, \frac{N k_r}{\alpha_0 + \beta_0 + j} \right) + \sum_{j=1}^{N} \psi_j \sum_{j=1}^{N} \chi_j \right] \min \left\{ \frac{(N - j) (\alpha_0 + \beta_0 + j)}{\alpha_0 + \beta_0 + j}, (N - j) k_r + j k_s \right\}^\psi_j, j. \]

Equality holds in (2.6) if \( \varphi_j, j \) (\( j = 0, 1, 2, \ldots, N \)) is defined in accordance with equations (2.1) and (2.2). This lemma is thus proved.

2.1.2. Stopping Rule for the Bayes Procedure

In this section, we characterize the stopping rule for the Bayes procedure.

Let \( \Delta_n \ (n = 0, 1, 2, \ldots, N) \) denote the class of sequential decision functions for this problem truncated with probability one at \( \eta = n \), i.e., with \( \psi_0 + \psi_1 + \ldots + \psi_n = 1 \) identically in \( x_n \); and let

\[(2.7) \quad \rho_n (\xi) = \min_{\delta \in \Delta_n} r (\xi; \delta) \quad \text{be the Bayes risk in} \ \Delta_n. \]

For the class of prior distributions which we consider, we write \( \rho_n (\xi) = \rho_n (\alpha; \beta) \) when \( \xi (\theta) = \xi_n (\theta; \alpha, \beta) \). Let

\[(2.8) \quad \rho_n^0 (\xi) = \rho_n^0 (\alpha; \beta) = \min \left( \frac{k \alpha}{\alpha + \beta}, nk_r \right), \quad (n=0, 1, 2, \ldots, N). \]

We first prove the following lemma.
Lemma 2.2

For the Bayes sequential procedure corresponding to the decision problem of Section 1.3,

\[(2.9) \, \rho_N^*(\alpha_0; \beta_0) = \min \left[ \rho_N^0(\alpha_0; \beta_0), \frac{\alpha_0}{\alpha_0 + \beta_0} \rho_{N-1}^*(\alpha_0 + 1; \beta_0) + \frac{\beta_0}{\alpha_0 + \beta_0} \rho_{N-1}^*(\alpha_0; \beta_0 + 1) + k_s \right].\]

\[\text{Proof}\]

From (2.5) we have that for a decision rule \( \delta \) belonging to \( \Delta_N' \),

\[(2.10) \, r_N^*(\xi; \beta) = \psi_0 \left[ (1 - \varphi_0) \frac{Nz_0}{\alpha_0 + \beta_0} + \varphi_0 (Nk_x) \right] + (1 - \psi_0) \sum_{x_1} \frac{\psi_1}{1 - \psi_0} \]

\[+ \left[ (1 - \varphi_1) \frac{(N - 1) (\alpha_0 + x_1)}{\alpha_0 + \beta_0 + 1} + \varphi_1 (N - 1) k_x + k_s \right] \sum \frac{\psi_j}{1 - \psi_0} \left[ (1 - \varphi_j) \frac{(N - j) (\alpha_0 + z_j)}{\alpha_0 + \beta_0 + j} + \varphi_j (N - j) k_x + jk_s \right] \sum \frac{\psi_j}{1 - \psi_0} \]

Now let

\[(2.11) \, \psi_j^{*} = \frac{\psi_j}{1 - \psi_0}; \, \varphi_j = \psi_j^{*}, \quad (j = 1, 2, \ldots, N).\]

It is obvious that \( \sum_{j=1}^{N} \psi_j^{*} (x_j) = 1 \) identically in \( x_j \); hence it follows that
(2.12)  \[ r_N (\xi; \sigma) = \psi_0 \left[ (1 - \varphi_0) \frac{N_0'}{\alpha_0 + \beta_0} + \varphi_0 (Nk_r) \right] + \psi_1 \left[ k_s + \sum_{x_1} \psi_0 \left( 1 - \varphi_0^* \right) \frac{(N - 1) (\alpha_0 + x_1)}{\alpha_0 + \beta_0 + 1} + \varphi_0^* (N - l) k_r \right] p_{l,1} + (1 - \psi_0) \sum_{j=2}^N \sum_{x_j} \psi_j \left[ (1 - \varphi_j^*) \frac{(N - j) (\alpha_0 + z_j)}{\alpha_0 + \beta_0 + j} + \varphi_j^* (N - j) k_r + (j - 1) k_s \right] p_{l,1}^*.

Let

(2.13)  \[ x_{j-1}^* = x_{j-1}^* (x_1) \]

denote the set of all possible \((j - 1)\)-vectors \((x_2, x_3, \ldots, x_j)\) obtained from \(j\)-vectors \(x_j\) for which \(X_1 = x_1\); let \(z_{j-1} = x_2 + x_3 + \ldots + x_j\) denote the sum of the components of the particular vector in \(x_{j-1}^*\) for which \(X_2 = x_2, \ldots, X_j = x_j\), \((j = 2, 3, \ldots, N)\).

Let \(p_{l,1}^*\) denote the conditional probability of observing \(X_1 = x_1^\ast\) \((i = 2, 3, \ldots, j)\) given that \(X_1 = x_1\); therefore, \(p_{l,1}^* = p_{l,1} / p_{l,1}^\ast\).

Finally, let \(j' = j - 1\).

From (2.12) we can therefore write

(2.14)  \[ r_N (\xi; \sigma) = \psi_0 \left[ (1 - \varphi_0) \frac{N_0'}{\alpha_0 + \beta_0} + \varphi_0 (Nk_r) \right] + (1 - \psi_0) \left[ k_s + \sum_{x_1} \psi_0 \left( 1 - \varphi_0^* \right) \frac{(N - 1) (\alpha_0 + x_1)}{\alpha_0 + \beta_0 + 1} + \varphi_0^* (N - l) k_r \right] p_{l,1} + \psi_1 \left[ \sum_{j=2}^{N-1} \sum_{x_j} \psi_j \left[ (1 - \varphi_j^*) \frac{(N - j - 1) (\alpha_0 + x_1 + z_{j-1}^*)}{\alpha_0 + \beta_0 + j + 1} + \varphi_j^* (N - j - 1) k_r + (j - 1) k_s \right] p_{l,1}^* \right].\]
\[
\begin{align*}
&= \psi_0 \left[ (1 - \varphi_0) \frac{N \lambda_0}{\alpha_0 + \beta_0} + \varphi_0 \left( N k_r \right) \right] + (1 - \psi_0) \left[ k_s + \sum_{\chi_1} \right] \\
&\quad \{ \psi_0^* \left[ (1 - \varphi_0^* \left( N - 1 \right) \left( \alpha_0 + x_1 \right) \frac{1}{\alpha_0 + \beta_0 + 1} + \varphi_0^* \left( N - 1 \right) k_r \right] \\
&\quad + \sum_{j' = 1}^{N-1} \sum_{\chi_j^*} \psi_{j'}^* \left( N - 1 - j' \right) \left( \alpha_0 + x_1 + z_{j'}^* \right) \frac{1}{\alpha_0 + \beta_0 + 1 + j'} \\
&\quad \quad \quad \times \varphi_{j'}^* \left( N - 1 - j' \right) k_r + j' k_s \right] p_{r', j'}, \right\} p_{r', 1}.
\end{align*}
\]

Let \( \delta^* \) be a sequential decision rule such that

\( \delta^* = (\psi_0^*, \psi_1^*, \ldots, \psi_{N-1}; \varphi_0^*, \varphi_1^*, \varphi_2^*, \ldots, \varphi_{N-1}) \). Identifying the right hand side of (2.14) with (2.5), we obtain

\begin{equation}
(2.15) \quad r_N (\xi; \delta) = \psi_0 \left[ (1 - \varphi_0) \frac{N \lambda_0}{\alpha_0 + \beta_0} + \varphi_0 \left( N k_r \right) \right] + (1 - \psi_0) \\
\quad \left[ k_s + \sum_{\chi_1} \frac{N \lambda_0}{\alpha_0 + \beta_0} r_{N-1} (\tau_{x_1} \xi (\theta^{(1)}); \delta^*) \right] p_{r', 1}.
\end{equation}

From (2.7), we also have

\begin{equation}
(2.16) \quad r_{N-1} (\tau_{x_1} \xi (\theta^{(1)}); \delta^*) \geq \rho_{N-1} (\tau_{x_1} \xi (\theta^{(1)})).
\end{equation}

Therefore it follows from (2.16) that

\[
\begin{align*}
r_N (\xi; \delta) &\geq \psi_0 \left[ (1 - \varphi_0) \frac{N \lambda_0}{\alpha_0 + \beta_0} + \varphi_0 \left( N k_r \right) \right] \\
&\quad + (1 - \psi_0) \left[ k_s + \sum_{\chi_1} \rho_{N-1} (\tau_{x_1} \xi (\theta^{(1)})) \right] p_{r', 1},
\end{align*}
\]

and hence from (2.2) that
Thus from (2.17) we see that

\[
(2.18) \quad \rho_N(x) = \inf_{\delta \in \Delta_N} r_N(x; \delta) = \min \{ \rho_N^0(x), k_s + \sum_{x_1} \rho_N(x, \delta(1)) \}.
\]

For our problem we have

\[
(2.19) \quad \rho_N(x) = \rho_N(\alpha_0; \beta_0); \quad \rho_N^0(x) = \rho_N(\alpha_0; \beta_0).
\]

From Lemma 1.3, we find that (2.18) reduces to

\[
\rho_N(\alpha_0; \beta_0) = \min \left[ \rho_N^0(\alpha_0; \beta_0), \frac{\alpha_0}{\alpha_0 + \beta_0} \rho_N(x, \delta(1)) \right. \left. + \frac{\beta_0}{\alpha_0 + \beta_0} \rho_N(x, \beta_0 + 1) + k_s \right],
\]

where \( \rho_N^0(\alpha_0; \beta_0) = \min \left( \frac{N\alpha_0}{\alpha_0 + \beta_0}, Nk \right) \).

The lemma is thus proved.

**Corollary 2.1**

The probability \( \psi_{x,0} \) that the Bayes procedure will not take any observations is given by equation (2.20).
\[ = 0 \text{ if } \rho_N (\alpha_0; \beta_0) < \rho_N^0 (\alpha_0; \beta_0) \]
\[ \psi_{s,0} = 1 \text{ if } \rho_N (\alpha_0; \beta_0) = \rho_N^0 (\alpha_0; \beta_0). \]

The proof is obvious and follows readily from (2.17).

In the above lemma, we confine our consideration to non-randomized Bayes procedures having tacitly assumed in (2.20) that if two members of the right hand side of (2.9) are equal, we take one more observation. More generally, we might have allowed for randomization when ties occur. However, here and in the sequel, we shall without any loss of generality consider non-randomized procedures only.

It can also be shown that for a lot of size \(N - j\), the Bayes risk with respect to \(\xi_{N-j} (\theta; \alpha, \beta)\) will be given by

\[ \rho_{N-j} (\alpha; \beta) = \min \left[ \rho_{N-j}^0 (\alpha; \beta), \frac{\alpha}{\alpha + \beta} \rho_{N-j-1} (\alpha + 1; \beta) + \frac{\beta}{\alpha + \beta} \rho_{N-j-1} (\alpha; \beta + 1) + k_s \right], \quad (j = 1, 2, \ldots, N) \]

and

\[ \rho_0 (\alpha; \beta) = 0. \]

Let \(\pi_{s, j}\) denote the conditional probability that the Bayes procedure stops after taking \(j\) observations given that it takes at least \(j\) observations.
Thus,

\[(2.22)\]

\[\pi_{\xi, j} \equiv \pi_{\xi, j} (x_j) = \frac{\psi_{\xi, j}}{\psi_{\xi, j} + \psi_{\xi, j+1} + \ldots + \psi_{\xi, N}}, \quad (j=0, 1, 2, \ldots, N).\]

We recall that \(\tau_{x_j} \xi (\theta(j)) = \xi_{N-j} (\theta(j); \alpha_0 + z_j, \beta_0 + j - z_j)\).

It now follows easily from (2.20) and (2.21) that

\[(2.23)\]

\[\pi_{\xi, j} = 1 \quad \text{if} \quad \rho_{N-j} (\alpha_0 + z_j; \beta_0 + j - z_j) = \rho_{N-j}^0 (\alpha_0 + z_j; \beta_0 + j - z_j),\]

\[\pi_{\xi, j} = 0 \quad \text{if} \quad \rho_{N-j} (\alpha_0 + z_j; \beta_0 + j - z_j) < \rho_{N-j}^0 (\alpha_0 + z_j; \beta_0 + j - z_j).\]

The above results can be conveniently summarized in the following theorem.

**Theorem 2.1**

The Bayes risk \(\rho_N (\alpha_0; \beta_0)\) for the decision problem of Section 1.3 with respect to the Polya apriori distribution \(\xi_N (\theta; \alpha_0, \beta_0)\) is computed recursively from the equations

\[(2.24)\]

\[\rho_j (\alpha; \beta) = \min \{\rho_j^0 (\alpha; \beta), \frac{\alpha}{\alpha + \beta} \rho_{j-1} (\alpha + 1; \beta)\]

\[+ \frac{\beta}{\alpha + \beta} \rho_{j-1} (\alpha; \beta + 1) + k_s\}, \quad (j = N, N-1, \ldots, 1),\]

and

\[(2.25)\]

\[\rho_0 (\alpha; \beta) = 0.\]

Further, the Bayes stopping rule with respect to the above prior distribution calls for stopping after \(j\) observations if and only if
(2.26)\[ \rho_{N-j} (a_0 + \sum_{i=1}^{j} x_i; \beta_0 + j - \sum_{i=1}^{j} x_i) = \rho_{N-j} (a_0 + \sum_{i=1}^{j} x_i; \beta_0 + j - \sum_{i=1}^{j} x_i), \]
and then the corresponding terminal decision rule is given by (2.1) and (2.2).

**Remark**

It may be appropriate to point out here that this simple characterization of the Bayes procedure has been possible, in spite of the many complications indicated in the beginning of this chapter, because of the particular form of the loss function (1.14) and the a priori distribution (1.3). If the loss functions were slightly different, e.g., \( L_j (\theta; a_2) = (N - j/2)k \), the above characterization would not have been possible.

We can therefore always compute \( \rho_N (a_0; \beta_0) \), at least in principle, starting with the known values given in (2.25) and then applying successively the recursion equations (2.24). Using (2.26) we can in this process also determine the optimum boundary. It is possible to simplify the computations further. We observe from (2.26) that if we start with the a priori distribution \( \bar{g}_N (\theta; a_0', \beta_0) \), the relevant points for which \( \rho_{N-j} (a; \beta) \) has to be defined after taking \( j \) observations are those for which

(2.27)\[ N - j + \alpha + \beta = N + \alpha_0 + \beta_0 = M \text{ (say),} \quad (j = 1, 2, \ldots, N). \]

In other words, the set of all a priori distributions which can arise following the Bayes rule from the a priori distribution \( \bar{g}_N (\theta; a_0; \beta_0) \) satisfy (2.27) and can be represented by the set of all lattice points in
a three-dimensional Euclidean space such that each of the co-ordinates is non-negative and they sum to \( M \).

Let \( \rho^* (\alpha; \beta) \) denote the Bayes risk for our decision problem when the apriori distribution is \( \xi_{M-\alpha-\beta}^\theta (\theta; \alpha, \beta) \) (i.e., the lct is of size \( M - \alpha - \beta \)). In terms of this notation we can now express the recursion relations (2.24) and (2.25) in the following alternative way:

(2.28)

\[
\rho^* (\alpha; \beta) = 0 \quad \text{for} \quad \alpha + \beta = M, \quad \text{and} \\
\rho^* (\alpha; \beta) = \min \left[ \rho_0 (\alpha; \beta), \frac{\alpha}{\alpha + \beta} \rho^* (\alpha+1; \beta) + \frac{\beta}{\alpha + \beta} \rho^* (\alpha; \beta+1) + k_s \right]
\]

for \( \alpha + \beta < M \),

where

(2.30)

\[
\rho_0 (\alpha; \beta) = \min \left[ (M - \alpha - \beta) \frac{\alpha}{\alpha + \beta}, (M - \alpha - \beta) k_r \right].
\]

Further, \((\alpha, \beta)\) will be a stopping point if and only if

(2.31)

\[
\rho^* (\alpha; \beta) = \rho_0 (\alpha; \beta).
\]

It is evident therefore that the set of stopping points thus characterized in the \((\alpha, \beta)\)-plane is sufficient to determine the stopping rules for the Bayes procedures with respect to all Polya distributions \( \xi_{N'}^\theta (\theta; \alpha', \beta') \) such that \( N' + \alpha' + \beta' = M \). For each different \( M \), however, one has to compute the optimal boundaries separately.
Using computers, it is quite simple to evaluate the Bayes risks and the optimal boundary from equations (2.28) to (2.31). For most of the plans which we describe later in this chapter, it took less than 1 minute using a 1604 CDC-computer to obtain the Bayes solution.

In this connection, we recall the comments of Anscombe [1] in which he suggested that the cost parameter \( k_s \) should be defined in such a way that it depends on \( \theta(i) \). As we noted earlier, it is not clear how we should formulate a model which takes this dependence into account. However, given that we succeed in obtaining some realistic model which takes care of this dependence, it should not be difficult to find out theoretically the Bayes stopping and terminal decision rules or to compute the optimum boundary associated with them. In such a situation, the recursion relation (2.29) will be altered to the extent that \( k_s \) will no longer be a constant but at each stage will be a function of \( \alpha \) and \( \beta \).

Since the problem of determining the Bayes sequential procedure is not expected to alter substantially with the introduction of this additional feature, we do not propose to consider it in this thesis. For a description of some Bayes sequential procedures for the binomial and the trinomial models where the cost depends on the unknown parameter, the reader is referred to Ray [35].

It is possible to prove the convexity of the Bayes stopping regions for our model using the same arguments as used by Blackwell and Girshick [7]. We omit the proof.

2.1.3. An Useful Result for the Computation of the Optimum Boundary

In Section 1.2 where we discussed the model proposed by Hald, we considered at length the significance of the cost parameters, and indicated why, in most practical situations, \( k_r \) should be less than \( k_s \).
If \( k_r \leq k_s \), it is possible to narrow the region of sampling to a considerable extent. We show this in the following lemma.

**Lemma 2.3**

Let

\[
M - \beta^* = \alpha^* = \begin{cases} 
M_{k_r} & \text{if } M_{k_r} \text{ is an integer} \\
\lceil M_{k_r} \rceil + 1 & \text{otherwise}.
\end{cases}
\]

Then under the assumption that

\[
k_r \leq k_s \leq 1,
\]

all lattice points \((\alpha, \beta)\) for which \(\alpha \geq \alpha^*\) or \(\beta \geq \beta^* + 1\) are stopping points, and

\[
\rho^*(\alpha; \beta) = (M - \alpha - \beta) \frac{\alpha}{\alpha + \beta} \quad \text{for } \beta \geq \beta^* + 1
\]

and

\[
\rho^*(\alpha; \beta) = (M - \alpha - \beta) k_r \quad \text{for } \alpha \geq \alpha^*.
\]

Thus the recursion relation (2.29) is needed for computing \(\rho^*(\alpha; \beta)\) only for lattice points in the rectangle \(\alpha \leq \alpha^* - 1, \beta \leq \beta^*; \alpha, \beta \geq 0\).

A result similar to the above has been stated without proof by Chernoff and Ray [13] for their one-action sampling inspection problem.

**Proof**

The proof is carried out by induction. First we prove (2.34).

Consider the set of lattice points \((\alpha^0, \beta^0)\) on the line \(\alpha^0 + \beta^0 = M - 1\) for which \(\beta^0 \geq \beta^* + 1\). Then by an application of (2.29) we obtain

\[
\rho^*(\alpha^0; \beta^0) = \min \{\frac{\alpha^0}{\alpha^0 + \beta^0}, k_r, k_s\}.
\]
From (2.32) we see that
\[ \beta^* + 1 \geq M - [ Mk_r ] \geq M (1 - k_r). \]

Also since \( \beta^0 \geq \beta^* + 1 \), we have
\[ \frac{\alpha^0}{\alpha^0 + \beta^0} = \frac{M - 1 - \beta^0}{M - 1} \leq \frac{M - 1 - M (1 - k_r)}{M - 1} = \frac{Mk_r - 1}{M - 1} \leq k_r. \]

Therefore from (2.33), (2.36), and (2.37) it follows that
\[ \rho^* (\alpha^0; \beta^0) = \frac{\alpha^0}{\alpha^0 + \beta^0}. \]

Now we make the induction assumption that for all lattice points \((\alpha, \beta)\) for which \( \alpha + \beta = M' + 1 < M \) and \( \beta \geq \beta^* + 1 \), we have
\[ \rho^* (\alpha; \beta) = (M - \alpha - \beta) \frac{\alpha}{\alpha + \beta}. \]

Let \((\alpha, \beta)\) be a lattice point satisfying \( \alpha + \beta = M \) and \( \beta \geq \beta^* + 1 \).

From (2.29) and (2.39) we obtain
\[ \rho^* (\alpha; \beta) = \min [(M - \alpha - \beta) \frac{\alpha}{\alpha + \beta}, (M - \alpha - \beta) k_r, (M - \alpha - \beta) \frac{\alpha + 1}{\alpha + \beta + 1} \]
\[ + \frac{\beta}{\alpha + \beta} (M - \alpha - \beta - 1) \frac{\alpha}{\alpha + \beta + 1} + k_s] \]
\[ = \min [(M - \alpha - \beta) \frac{\alpha}{\alpha + \beta}, (M - \alpha - \beta) k_r, (M - \alpha - \beta) \frac{\alpha}{\alpha + \beta} + (k_s - \frac{\alpha}{\alpha + \beta})]. \]

Since \( \beta \geq \beta^* + 1 \), it follows from (2.37) that \( \frac{\alpha}{\alpha + \beta} \leq k_r \). Now using (2.33) we obtain
\[ \rho^* (\alpha; \beta) = (M - \alpha - \beta) \frac{\alpha}{\alpha + \beta}. \]
We have therefore shown that if \( \rho^* (\alpha; \beta) = (M - \alpha - \beta) \frac{\alpha}{\alpha + \beta} \) for any lattice point \((\alpha, \beta)\) which satisfies \( \alpha + \beta = M' + 1 < M \) and \( \beta \geq \beta^* + 1 \), then \( \rho^* (\alpha; \beta) = (M - \alpha - \beta) \frac{\alpha}{\alpha + \beta} \) for all lattice points \((\alpha, \beta)\) for which \( \alpha + \beta = M' \) and \( \beta \geq \beta^* + 1 \). But since this result holds when \( \alpha + \beta = M - l \) and \( \beta \geq \beta^* + 1 \) [see (2.38)], it follows that (2.34) must be true for all lattice points in the positive quadrant for which \( \alpha + \beta \leq M \) and \( \beta \geq \beta^* + 1 \).

The proof of (2.35) is also carried out in a similar way.

Let \((\alpha, \beta)\) be a lattice point satisfying \( \alpha + \beta = M - l \) and \( \alpha \geq \alpha^* \). From (2.29) we find similarly that

\[
(2.42) \quad \rho^* (\alpha; \beta) = \min \left[ \frac{\alpha}{\alpha + \beta}, k_r, k_s \right].
\]

But

\[
(2.43) \quad \frac{\alpha}{\alpha + \beta} = \frac{\alpha}{M - 1} \geq \frac{\alpha}{M} \geq \frac{\alpha^*}{M} \geq \frac{Mk_r}{M} = k_r.
\]

Hence from (2.43) we see that

\[
(2.44) \quad \rho^* (\alpha; \beta) = k_r \quad \text{for} \quad \alpha + \beta = M - 1 \quad \text{and} \quad \alpha \geq \alpha^*.
\]

Now assume that

\[
(2.45) \quad \rho^* (\alpha; \beta) = (M - \alpha - \beta) k_r \quad \text{for} \quad \alpha \geq \alpha^* \quad \text{and} \quad \alpha + \beta = M' + l \leq M.
\]

Therefore for any lattice point \((\alpha, \beta)\) satisfying \( \alpha + \beta = M' \) and \( \alpha \geq \alpha^* \), we have from (2.29), (2.33), and (2.43) that
(2.46)
\[
\rho^* (\alpha; \beta) = \min \left[ \frac{(M-\alpha-\beta)\alpha}{\alpha+\beta}, (M-\alpha-\beta) k_r, \frac{\alpha}{\alpha+\beta} (M-\alpha-\beta-1) k_r + \frac{\beta}{\alpha+\beta} (M-\alpha-\beta-1) k_r + k_s \right]
\]
\[
= \min \left[ \frac{(M-\alpha-\beta)\alpha}{\alpha+\beta}, (M-\alpha-\beta) k_r, (M-\alpha-\beta) k_r + k_s - k_r \right]
\]
\[
= (M-\alpha-\beta) k_r.
\]

Using an induction argument similar to the above we see that (2.35) follows from (2.43) and (2.46).

Lemma 2.3 was found to be of particular use for the construction of the sampling plans which are derived in a later section.

2.1.4. **Determination of an Upper Bound on the Point of Truncation for the Bayes Sequential Procedure.**

A sequential procedure is said to be truncated if there exists an integer \( n' \) such that \( P [\eta \leq n'] = 1 \). If \( n_0 \) is the minimum of all such integers \( n' \), then \( n_0 \) is called the exact stage of truncation.

For our particular problem, any sequential procedure is necessarily truncated since no more than \( N \) observations can be taken by any plan. Nevertheless, for the purpose of constructing the Bayes procedure by backward induction, it is natural to inquire whether it is truncated at an earlier stage and if so, whether it is possible to determine the exact stage of truncation or to specify a smaller upper bound than \( N \). In recent years, the determination of the upper bound on the maximum sample size of a Bayes sequential procedure has received considerable attention.
in the literature (see, for example, Wetherill [44], Weiss [43], and Ray [36]). In this section we determine an upper bound on the point of truncation by applying techniques developed by Ray [36].

We noted earlier that the set of all posterior distributions, which can arise when the Bayes procedure with respect to $\xi_N (\theta; \alpha_0, \beta_0)$ is used, can be represented by the set of lattice points in the $(\alpha, \beta)$-plane; also, the Bayes risk associated with any lattice point $(\alpha', \beta')$ can be denoted by the symbol $\rho^* (\alpha'; \beta')$ where $\rho^* (\alpha'; \beta') = \rho_{M-\alpha'-\beta'} (\alpha'; \beta')$. Following Ray [36] let us denote by $\lambda (\alpha; \beta)$ the gain due to stopping after one free observation rather than stopping immediately when the apriori distribution is $\xi_{M-\alpha-\beta} (\theta; \alpha, \beta) = \xi_{\alpha, \beta}$ (say). Let $E_\xi [g (X)]$ denote the expectation of $g(X)$ computed with respect to the marginal distribution of $X$ when $\xi_{\alpha, \beta}$ is the apriori distribution of $\theta$. In terms of these definitions and those given earlier, we have

\begin{equation}
\lambda (\alpha; \beta) = \rho_0 (\xi_{\alpha, \beta}) - E_\xi [\rho_0 (\tau_X, \xi_{\alpha, \beta})].
\end{equation}

Also

\begin{equation}
\rho_0 (\xi_{\alpha, \beta}) = \min \left\{ (M-\alpha-\beta) \frac{\alpha}{\alpha+\beta} , (M-\alpha-\beta) k_x \right\}
= \frac{1}{2} (M-\alpha-\beta) \left( \frac{\alpha}{\alpha+\beta} + k_x \right) - \frac{1}{2} (M-\alpha-\beta) \left| \frac{\alpha}{\alpha+\beta} - k_x \right|.
\end{equation}

Similarly, when $X = 1$,

\begin{equation}
\rho_0 (\tau_X, \xi_{\alpha, \beta}) = \rho_0 (\xi_{\alpha+1, \beta})
= \min \left\{ (M-\alpha-\beta-1) \frac{\alpha+1}{\alpha+\beta+1} , (M-\alpha-\beta-1) k_x \right\}
= \frac{1}{2} (M-\alpha-\beta-1) \left( \frac{\alpha+1}{\alpha+\beta+1} + k_x \right) - \frac{1}{2} (M-\alpha-\beta-1) \left| \frac{\alpha+1}{\alpha+\beta+1} - k_x \right|,
\end{equation}
and when \( X = 0 \),

\[
\rho_0 (\tau_x \xi_{\alpha, \beta}) = \rho_0 (\xi_{\alpha+1, \beta}) = \frac{1}{2} (M-\alpha-\beta-1) \left( \frac{\alpha}{\alpha+\beta+1} + k_r \right) - \frac{1}{2} (M-\alpha-\beta-1) \left\lvert \frac{\alpha}{\alpha+\beta+1} - k_r \right\rvert .
\]

Therefore

\[
\lambda (\alpha; \beta) = \rho_0 (\xi_{\alpha, \beta}) - \frac{\alpha}{\alpha+\beta} \rho_0 (\xi_{\alpha+1, \beta}) - \frac{\beta}{\alpha+\beta} \rho_0 (\xi_{\alpha, \beta+1}).
\]

We shall now evaluate (2.51) using (2.48) to (2.50) and taking into account all possible cases for which the quantities \( \frac{\alpha}{\alpha+\beta} - k_r \), \( \frac{\alpha+1}{\alpha+\beta+1} - k_r \) and \( \frac{\alpha}{\alpha+\beta+1} - k_r \) have different signs.

**Case 1:**

When \((\alpha, \beta)\) is such that

\[
\frac{\alpha}{\alpha+\beta} - k_r \geq 0, \quad \frac{\alpha+1}{\alpha+\beta+1} - k_r \geq 0 \quad \text{and} \quad \frac{\alpha}{\alpha+\beta+1} - k_r \geq 0.
\]

In this case,

\[
\lambda (\alpha; \beta) = \frac{1}{2} \left( k_r + \frac{\alpha}{\alpha+\beta} \right) + \frac{1}{2} \left( k_r - \frac{\alpha}{\alpha+\beta} \right) = k_r .
\]

**Case 2:**

When \((\alpha, \beta)\) is such that

\[
\frac{\alpha}{\alpha+\beta} - k_r \leq 0, \quad \frac{\alpha+1}{\alpha+\beta+1} - k_r \leq 0 \quad \text{and} \quad \frac{\alpha}{\alpha+\beta+1} - k_r \leq 0.
\]

In this case

\[
\lambda (\alpha; \beta) = \frac{1}{2} \left( k_r + \frac{\alpha}{\alpha+\beta} \right) + \frac{1}{2} \left( \frac{\alpha}{\alpha+\beta} - k_r \right) = \frac{\alpha}{\alpha+\beta} .
\]
Case 3:

When \((\alpha, \beta)\) is such that \(\frac{\alpha}{\alpha+\beta} - k_r = 0\) and consequently \(\frac{\alpha+1}{\alpha+\beta+1} - k_r > 0\), \(\frac{\alpha+1}{\alpha+\beta+1} - k_r < 0\), i.e., when \((\alpha, \beta)\) is on the so-called "neutral line."

In this case

\[
\lambda (\alpha; \beta) = \frac{1}{\beta} (k_r + \frac{\alpha}{\alpha+\beta}) + \frac{1}{\beta} (M-\alpha-\beta-1) \frac{\alpha}{\alpha+\beta} \left( \frac{\alpha+1}{\alpha+\beta+1} - k_r \right)
\]

\[
+ \frac{1}{\beta} (M-\alpha-\beta-1) \frac{\beta}{\alpha+\beta} (k_r - \frac{\alpha}{\alpha+\beta+1})
\]

\[
= k_r + \frac{(M - \alpha - \beta - 1) \alpha \beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)}
\]

\[
= k_r + \frac{(M - \alpha - \beta - 1) k_r (1 - k_r)}{\alpha + \beta + 1},
\]

and hence

\[
(2.52) \quad \lambda (\alpha; \beta) < k_r + \frac{(M - \alpha - \beta - 1) k_r (1 - k_r)}{\alpha + \beta + 1}.
\]

Case 4:

When \((\alpha, \beta)\) satisfies \(\frac{\alpha}{\alpha + \beta + 1} < k_r < \frac{\alpha}{\alpha + \beta}\).

In this case

\[
\lambda (\alpha; \beta) = \frac{1}{\beta} (k_r + \frac{\alpha}{\alpha+\beta}) - \frac{1}{\beta} (M-\alpha-\beta) \left( \frac{\alpha}{\alpha+\beta} - k_r \right) +
\]

\[
\frac{1}{\beta} (M-\alpha-\beta-1) \frac{\beta}{\alpha+\beta} (k_r - \frac{\alpha}{\alpha+\beta+1}) + \frac{1}{\beta} (M-\alpha-\beta-1) \frac{\alpha}{\alpha+\beta} \left( \frac{\alpha+1}{\alpha+\beta+1} - k_r \right)
\]

\[
= k_r + (M-\alpha-\beta-1) \left[ \frac{\beta}{\alpha+\beta} k_r - \frac{\alpha \beta}{(\alpha+\beta) (\alpha+\beta+1)} \right]
\]

\[
< k_r + (M-\alpha-\beta-1) \left[ \frac{\alpha \beta}{(\alpha+\beta)^2} - \frac{\alpha \beta}{(\alpha+\beta) (\alpha+\beta+1)} \right]
\]

\[
= k_r + \frac{(M-\alpha-\beta-1) \alpha \beta}{(\alpha+\beta)^2 (\alpha+\beta+1)}.
\]
Thus

\[(2.53) \quad \lambda (\alpha; \beta) < k_r + (M - \alpha - \beta - 1) \frac{k_r (1 - k_r)}{\alpha + \beta} . \]

**Case 5:**

When \((\alpha, \beta)\) satisfies \(\frac{\alpha}{\alpha + \beta} < k_r < \frac{\alpha + 1}{\alpha + \beta + 1} .\)

In this case

\[\lambda (\alpha; \beta) = \frac{1}{2} (k_r + \frac{\alpha}{\alpha+\beta}) - \frac{1}{2} (M-\alpha-\beta) (k_r - \frac{\alpha}{\alpha+\beta}) + \]

\[\frac{1}{2} (M-\alpha-\beta-1) \frac{\alpha}{\alpha+\beta} \left(\frac{\alpha+1}{\alpha+\beta+1} - k_r\right) + \frac{1}{2} (M-\alpha-\beta-1) \frac{\beta}{\alpha+\beta} \left(k_r - \frac{\alpha}{\alpha+\beta+1}\right) \]

\[= \frac{\alpha}{\alpha+\beta} + (M-\alpha-\beta-1) \left[\frac{\alpha (\alpha+1)}{(\alpha+\beta) (\alpha+\beta+1)} - \frac{\alpha^2}{(\alpha+\beta)^2}\right] \]

\[< k_r + (M-\alpha-\beta-1) \left[\frac{\alpha (\alpha+1)}{(\alpha+\beta) (\alpha+\beta+1)} - \frac{\alpha^2}{(\alpha+\beta)^2}\right] \]

\[= k_r + \frac{(M-\alpha-\beta-1) \alpha \beta}{(\alpha+\beta)^2} (\alpha+\beta+1) .\]

Thus

\[(2.54) \quad \lambda (\alpha; \beta) < k_r + (M - \alpha - \beta - 1) \frac{k_r (1 - k_r)}{\alpha + \beta} . \]

These five cases exhaust all of the possibilities for which the quantities \(\frac{\alpha}{\alpha + \beta} - k_r, \frac{\alpha + 1}{\alpha + \beta + 1} - k_r\) and \(\frac{\alpha}{\alpha + \beta + 1} - k_r\) have different signs. To summarize, we now have
(2.55) \[ \lambda (\alpha; \beta) = k_r \]

\[ = \frac{\alpha}{\alpha + \beta} \]

in Case 1

\[ = k_r + \frac{(M - \alpha - \beta - 1) k_r (1 - k_r)}{\alpha + \beta + 1} \]

in Case 2

\[ = k_r + (M - \alpha - \beta - 1) \left[ \frac{\beta}{\alpha + \beta} k_r - \frac{\alpha}{\alpha + \beta + 1} \right] \]

in Case 3

\[ = k_r + (M - \alpha - \beta - 1) \left[ \frac{\alpha (\alpha + 1)}{\alpha + \beta + 1} - \frac{\alpha}{\alpha + \beta} k_r \right] \]

in Case 4

\[ = \frac{\alpha}{\alpha + \beta} + (M - \alpha - \beta - 1) \left[ \frac{\alpha (\alpha + 1)}{\alpha + \beta + 1} - \frac{\alpha}{\alpha + \beta} k_r \right] \]

in Case 5

Also from (2.52), (2.53), (2.54) and the fact that \( \frac{\alpha}{\alpha + \beta} > k_r \) in Case 2, it follows that for all points \((\alpha, \beta)\) we have

\[ (2.56) \lambda (\alpha; \beta) < k_r + \frac{(M - \alpha - \beta - 1) k_r (1 - k_r)}{\alpha + \beta} \]

If the apriori distribution of \( \theta \) is \( \xi_N (\theta; \alpha_0, \beta_0) \), then after \( n \) observations have been taken, the a posteriori distribution of \( \theta (n) \) is given by \( \xi_{N-n} (\theta (n); \alpha_0 + z_n; \beta_0 + n - z_n) \). Identifying \( \alpha_0 + z_n \) with \( \alpha \) and \( \beta_0 + n - z_n \) with \( \beta \) we find that for our problem, \( \alpha + \beta = \alpha_0 + \beta_0 + n \). Rewriting (2.56) and recalling that \( M = N + \alpha_0 + \beta_0 \), we finally have

\[ (2.57) \sup_{n \text{ fixed}} \lambda (\alpha; \beta) < k_r + \frac{(N - n - 1)}{\alpha_0 + \beta_0 + n} k_r (1 - k_r), \]

the right hand side of the inequality being a decreasing function of \( n \).

Now we are in a position to apply Ray's Theorem 3.1 (see [36]) according to which an upper bound on the point of truncation is given by \( n^* \) where \( n^* \) is the smallest integer for which
\[
\sup_{n \text{ fixed}} \lambda (\alpha; \beta) \leq k_s \text{ uniformly in } n
\]

Therefore, using (2.57) we have that

\[(2.58)\]

\[
0 \quad \text{if } n' \leq 0
\]

\[
n^* = n' \quad \text{if } n' \text{ is a positive integer}
\]

\[
[n'] + 1 \quad \text{if } n' \text{ is not a positive integer}
\]

where

\[(2.59)\]

\[
n' = \frac{(N - 1) k_r (1 - k_r) - (\alpha_0 + \beta_0) (k_s - k_r)}{k_s - k_r^2}.
\]

and \([a]\) is the largest integer contained in \(a\).

The following lemma now summarizes the above results.

**Lemma 2.4**

An upper bound on the stage of truncation for the Bayes sequential procedure with respect to the apriori distribution \(\tau_N (\theta; \alpha_0, \beta_0)\) is given by \(n^*\) where \(n^*\) is determined from (2.58).

**Corollary 2.2**

In the special case when \(k_r = k_s\), an upper bound on the point of truncation for the Bayes sequential procedure is \(N - 1\).

The proof is immediate and obtained by direct substitution in (2.58).

**Remark**

We note here that in order to obtain the above corollary, such an involved argument as that of Lemma 2.4 is not necessary. This result follows directly from (2.28) and (2.29). When \(n = N - 1\), then we have \(\alpha + \beta = M - 1\), and
(2.60) \[ \rho_0(\alpha; \beta) = \min \left( \frac{\alpha}{\alpha + \beta}; k \right) \leq k \]
\[ = k + \frac{\alpha}{\alpha + \beta} \rho^*(\alpha + 1; \beta) + \frac{\beta}{\alpha + \beta} \rho^*(\alpha; \beta + 1). \]

Therefore, \( \rho^*(\alpha; \beta) = \rho_0(\alpha; \beta) \), and all points \((\alpha, \beta)\) for which \(\alpha + \beta = M - 1\) are stopping points.

We can now combine the results of Lemma 2.3 and Lemma 2.4, and state the following theorem which implies that the region of sampling can be narrowed down considerably in many situations.

**Theorem 2.2**

The continuation region for the Bayes sequential procedure corresponding to the decision problem of Section 1.3 is contained within the pentagon \(P\) in the \((\alpha, \beta)\)-plane where \(P\) is bounded by the lines \(\alpha = 0, \beta = 0, \alpha = \alpha^* - 1, \beta = \beta^*\) and \(\alpha + \beta = n^*\), the stopping risks outside this region being given by (2.34) and (2.35).

The saving in computation as a result of starting the backward induction process from \(n^*\) instead of \(N\) will depend on the values of the parameters \(N, k_r, k_s, \alpha_0,\) and \(\beta_0\). If \(k_s\) is considerably larger than \(k_r\), \(n^*\) will be much less than \(N\) and therefore there will be a sizeable decrease in the computing time; on the other hand if \(k_s = k_r\), \(n^* = N - 1\) and very little gain is achieved.

Thus (2.58) provides us with an upper bound on the point of truncation which is less than \(N\), but whether or not this bound is sufficiently close to the exact stage of truncation will, it appears, depend
on the values of $M$ and the cost parameters. In many situations, however, knowledge of the exact stage of truncation may be important and lead to considerable savings.

It is often possible to locate the exact truncation point or failing that, to obtain improved bounds by inspection and by performing somewhat more precise computations than those employed in Lemma 2.4. We proceed to illustrate this point by means of an example. For simplicity we refer to the truncation point as the one obtained with respect to the apriori distribution $\xi_M (\theta; 0, 0)$ and denote it by $N^*$. In order to obtain the stage of truncation corresponding to $\xi_N (\theta; \alpha_0, \beta_0)$, where $N + \alpha_0 + \beta_0 = M$, the frame of reference is shifted and the maximum sample size in this case is given by $N^* + \alpha_0 + \beta_0$.

2.1.5. Improved Upper Bounds on the Point of Truncation: An Example.

We consider an example for which $M = 450$, and $k_s = k_r = .02$. In terms of our previous notation, we obtain $\alpha^* = 9$ and $\beta^* = 441$. From Theorem 2.2, we know that $M - 1$ is an upper bound on the point of truncation, and any point $(\alpha, \beta)$ for which $\alpha \geq \alpha^*$ or $\beta \geq \beta^* + 1$ is a stopping point. For this particular example, the points on the neutral line are given by $(\alpha, .02\alpha)$ and we note that there can be only 9 lattice points on the neutral line in the positive quadrant for $\alpha + \beta \leq M$. The points are as follows: $(1, 49); (2.98); (3, 147); (4, 196); (5, 245); (6, 294); (7, 343); (8, 392); (9, 441).

We also know that $(9, 441)$ is a stopping point since it lies on the line $\alpha + \beta = 450$. For the purpose of obtaining an upper bound on the point of truncation in Lemma 2.4, we replaced $\lambda (\alpha; \beta)$ by the maximum value which it can assume for a given $n$, namely that on the neutral line.
Since there are no lattice points on the neutral line between $\alpha + \beta = 401$ and $\alpha + \beta = 449$, it is obvious that the exact truncation point will be considerably less than 449, which is the upper bound given in Corollary 2.2.

On the other hand, from (2.54) and the fact that for $\alpha + \beta = 400$ there is a lattice point on the neutral line, one may expect that the exact stage of truncation will be greater than 400. In order to determine the exact stage of truncation we therefore investigate the lattice points $(\alpha, \beta)$ which are between the lines $\alpha + \beta = 401$ and $\alpha + \beta = 449$ and for which either $\alpha/(\alpha + \beta + 1) < k_r < \alpha/(\alpha + \beta)$ or $\alpha/(\alpha + \beta) < k_r < (\alpha + 1)/(\alpha + \beta + 1)$ [Cases 4 and 5 of Lemma 2.4]. This set of points, together with the lattice points on the neutral line, constitute the so-called "extended neutral boundary" (see Ray [35]). One can afford to ignore all other points for this purpose since according to this lemma, $\lambda(\alpha; \beta) \leq k_r$ for all these points.

We noted earlier that all lattice points $(\alpha, \beta)$ for which $\alpha \geq \alpha^* = 9$ are stopping points, and therefore we need only consider those points on the extended neutral boundary for which $\alpha \leq 8$. We note further that between the lines $\alpha + \beta = 401$ and $\alpha + \beta = 448$, there will be no lattice points with $\alpha = 8$ for which the inequality $\alpha/(\alpha + \beta + 1) < k_r < \alpha/(\alpha + \beta)$ holds, and there will be a sequence of lattice points $(8, N', 8), (N' = 401, 402, \ldots, 448)$ for which the inequality $\alpha/(\alpha + \beta) < k_r < (\alpha + 1)/(\alpha + \beta + 1)$ holds. For $N' = 449$ there does not exist any point on the extended neutral boundary. In order to investigate the value of $\lambda(\alpha; \beta)$ for points on the extended neutral boundary, we need only consider the expression obtained for Case 5 of Lemma 2.4. From (2.55) we have
(2.61) \[ \lambda(\alpha; \beta) = \frac{\alpha}{\alpha + \beta} + \frac{(M - \alpha - \beta - 1) \alpha}{\alpha + \beta} \left( \frac{\alpha + 1}{\alpha + \beta + 1} - k_r \right). \]

This is a decreasing function of \(\alpha + \beta\) so that if for a given value of \(\alpha + \beta = N'\), \(\lambda(\alpha; \beta) \leq k_s\) for all values of \(\alpha\) and \(\beta\), this inequality will also be satisfied for all \(N > N'\). Now for any point \((\alpha, \beta)\) for which \(\alpha = \delta, \alpha + \beta = N'\) and \(\alpha/(\alpha + \beta) < k_r < (\alpha + 1)/(\alpha + \beta + 1)\), we obtain from (2.61)

(2.62) \[ \lambda(\alpha; \beta) = \frac{\delta}{N} + (M - N' - 1) \left[ \frac{8}{N} \left( \frac{9}{N' + 1} - k_r \right) \right]. \]

When \(M = 450\) and \(k_r = .02\), it is easily found using (2.62) that the smallest \(N'\) value satisfying the inequality \(\lambda(\alpha; \beta) \leq k_s = .02\) is \(N' = 418\) and this inequality holds uniformly for all \(N' > 418\). Therefore \(N' = 418\) is an improved upper bound on the point of truncation.

It is however not possible to assert with the help of the sufficient condition given by Ray [36] that this is also the exact stage of truncation. This condition tells us that in order to show that \(N'\) is the exact stage of truncation, one must find a set of \((x_1, x_2, \ldots, x_{N'-1})\) such that

(2.63) \[ P \left[ \lambda \left( \frac{\tau}{x_i} \right) > k_s \right] > 0 \quad (i = 0, 1, 2, \ldots, N' - 1). \]

For our problem when \(k_r = k_s = k\), it is not possible to find a path such that the condition (2.63) is satisfied. This is clear from the following consideration. If a line \(\alpha + \beta = c\) intersects the line \(\alpha/(\alpha + \beta) = k\) at a lattice point, then the intersection of the extended neutral boundary with the line \(\alpha + \beta = c - 1\) is empty. We have shown before that \(\lambda(\alpha; \beta)\) can be greater than \(k\) only at points belonging
to the extended neutral boundary. From these considerations and the preceding enumeration of neutral points, we observe that for all points \((\alpha, \beta)\) on the line \(\alpha + \beta = 49, 99, 149, 199, 249, 299, 349\) and \(399\), the inequality (2.63) will never be satisfied.

It is thus clear that the sufficient condition (2.53) is not satisfied in this example; we can therefore only assert that \(N' = 418\) is an upper bound on the point of truncation. This value of \(N'\), however, happens to be the exact stage of truncation as we can easily verify by a direct application of (2.28) and (2.29).

We conclude from the preceding example that in many situations it will be possible to obtain improved upper bound on the point of truncation than that given in Lemma 2.4. We also note that the bound obtained in the above example turns out to be the same as the exact stage of truncation although we are not able to conclude this from strictly theoretical grounds alone. In one special situation, however, namely when \(k_r = k_s = 0.5\), the exact stage of truncation can be determined theoretically by another set of sufficient conditions also due to Ray [36]. We now discuss this case briefly.

**Special case** \[k_s = k_r = k = 0.5.\]

In this special situation, the extended neutral boundary consists of points on the neutral line only. Therefore we do not have to consider Cases 4 and 5 of Lemma 2.4. We have from (2.55) that in this particular situation,
\[ \lambda (\alpha; \beta) = \frac{1}{2} \]  

in Case 1 \[(2.64) \quad \lambda (\alpha; \beta) = \frac{\alpha}{\alpha + \beta} \]  

in Case 2

\[ \lambda (\alpha; \beta) = \frac{1}{2} + \frac{(M - \alpha - \beta - 1)}{4(\alpha + \beta + 1)} \]  

in Case 3.

Following Ray [36], let us now define

\[(2.65) \quad \beta (\alpha; \beta) = [\lambda (\alpha; \beta) - k + E_x \{ \lambda (\tau_x; \xi_{\alpha, \beta}) - k \} ]_+ - [\lambda (\alpha; \beta) - k]_+ \]

where

\[ [a]_+ = \max (a, 0). \]

It can be easily verified that \( \beta (\alpha; \beta) = 0 \) for all \((\alpha, \beta)\) except those for which \( |\alpha - \beta| = 1 \), and that

\[ \beta (\alpha; \alpha - 1) = [\frac{(\alpha - 1)(M - 2 \alpha)}{4(2 \alpha - 1)^2}]_+ \]

\[(2.66) \quad \beta (\alpha - 1; \alpha) = [\frac{\alpha(M - 2 \alpha)}{4(2 \alpha - 1)^2} - \frac{1}{2(2 \alpha - 1)}]_+ \]

We now state the following lemma.

**Lemma 2.6**

For the special case \( k_1 = k_2 = \frac{1}{2} \), the Bayes sequential procedure with respect to \( t_N(\theta; \alpha_0, \beta_0) \) is truncated at \( N - 1 \).

**Proof**

Using an argument similar to that employed in the proof of Lemma 2.4, we can show from (2.66) that
(2.67)
\[
\sup_{n \text{ fixed}} \beta(\alpha; \beta) = \left[ \frac{(n - 1)(M - n - 1)}{8n^2} \right]_+ = g(n) \quad \text{(say),}
\]
and therefore \( g(n) = 0 \) uniformly for \( n \geq M - 1 \).

Applying Ray's [36] Theorem 3.3, we find that \( M - 1 \) is an upper bound on the point of truncation. It is easily verified from (2.64) and (2.66) that there exists a set of \( \{x_1, x_2, \ldots\} \) with positive probability such that

(2.68)
\[
\lambda(\tau_{x_1, \xi}) > \frac{1}{2} \quad \text{for every even } i \text{ less than } M - 1
\]
and \( \beta(\tau_{x_i, \xi}) > 0 \) for every odd \( i \) less than \( M - 1 \).

Thus the sufficient conditions (3.10) and (3.11) of Ray [36] are satisfied, and we conclude that \( N - 1 \) is the exact stage of truncation.


In this section we sketch the procedure which we followed in determining the optimal boundary for some representative plans which are described in Section 2.4.

Since we are dealing with discrete variables in our problem, the term "optimal boundary" requires some clarification. Following Ray [35], we now define the optimal boundary as the set of points \((\alpha, \beta)\) in the stopping region such that \((\alpha - 1, \beta)\) or \((\alpha, \beta - 1)\) is in the optimal continuation set.
From Lemma 2.3 we know that all points for which \( \alpha \geq \alpha^* \) or \( \beta \geq \beta^* + 1 \) hold are stopping points, and therefore the continuation region cannot lie outside the rectangle formed by \( \alpha = \alpha^* - 1, \beta = \beta^* \) in the positive quadrant. Also, on the line \( \beta = \beta^* + 1 \) all lattice points are stopping points, and the values of \( \rho^*(\alpha; \beta) \) on this line are given by (2.34). To determine the optimum boundary, we use (2.34) to compute \( \rho^*(\alpha; \beta) \) on the line \( \beta = \beta^* + 1 \); we next compute \( \rho^*(\alpha; \beta) \) on the line \( \beta = \beta^* \), starting with \( \alpha = \alpha^* \) and decreasing the \( \alpha \)-values by one at each step. This is done using the recursion relations (2.29). For each point on the line we test using (2.31) to determine whether or not it is a stopping point. We continue in this manner until we reach zero or a continuation point. If the former occurs, we follow the same procedure on the line \( \beta = \beta^* - 1 \), etc., until we reach a continuation point. At this stage we conclude that the stopping point immediately preceding this continuation point is on the optimal boundary. The process is then repeated for the successive points on this line until a stopping point is reached. As soon as we find a stopping point, we conclude that it is on the optimal boundary. We do not need to interrogate any further points on the line since the stopping region is convex. After having located the points on the boundary for a fixed value of \( \beta \), we move down the row one step, i.e., decrease the value of \( \beta \) by 1, and carry on this procedure again. We continue this procedure until \( \beta = 0 \) when we shall have located all the points on the optimal boundary. Our program on the computer was written in CORC, the Cornell computing language, which was found to be adequate for solving our problem. We omit the details.

It is easily seen that the optimal boundary does not depend on the absolute magnitude of \( \rho^* \) but rather on the relative magnitudes of the quantities on the right hand side of (2.29). This therefore suggests that we may well obtain the optimal boundary more conveniently if instead of considering the absolute magnitude of \( \rho^* \), we consider certain derived quantities the computation of which will involve fewer arithmetic operations. With this object in mind we define following Ray [35] the translated quantity

\[(2.69) \quad V(\alpha; \beta) = \rho_0(\alpha; \beta) - \rho^*(\alpha; \beta).\]

We can now express (2.29) in terms of \( V(\alpha; \beta) \) in the following way:

\[(2.70) \quad V(\alpha; \beta) = \max \left[ 0, \frac{\alpha}{\alpha + \beta} V(\alpha + 1; \beta) + \frac{\beta}{\alpha + \beta} V(\alpha; \beta + 1) - k_s \right. \]

\[+ \left. \{ \rho_0(\alpha; \beta) - \frac{\alpha}{\alpha + \beta} \rho_0(\alpha + 1; \beta) - \frac{\beta}{\alpha + \beta} \rho_0(\alpha; \beta + 1) \} \right].\]

The expression \( \rho_0(\alpha; \beta) - \frac{\alpha}{\alpha + \beta} \rho_0(\alpha + 1; \beta) - \frac{\beta}{\alpha + \beta} \rho_0(\alpha; \beta + 1) \) is the same as our \( \lambda(\alpha; \beta) \) defined by (2.51). Using (2.55) we can write

\[(2.71) \quad V(\alpha; \beta) = \max \left[ 0, \frac{\alpha}{\alpha + \beta} V(\alpha + 1; \beta) + \frac{\beta}{\alpha + \beta} V(\alpha; \beta + 1) + k_r - k_s \right] \text{ in Case 1} \]

\[= \max \left[ 0, \frac{\alpha}{\alpha + \beta} V(\alpha + 1; \beta) + \frac{\beta}{\alpha + \beta} V(\alpha; \beta + 1) + \frac{\alpha}{\alpha + \beta} - k_s \right] \text{ in Case 2} \]
\[
\begin{align*}
= \max & \left[ 0; \frac{\alpha}{\alpha+\beta} V (\alpha+1, \beta) + \frac{\beta}{\alpha+\beta} V (\alpha, \beta+1) + k_r + (M-\alpha-\beta-1) \frac{M-\alpha-\beta-1}{\alpha+\beta+1} (1-k_r) - k_s \right] \quad \text{in Case 3} \\
= \max & \left[ 0; \frac{\alpha}{\alpha+\beta} V (\alpha+1, \beta) + \frac{\beta}{\alpha+\beta} V (\alpha, \beta+1) + k_r + (M-\alpha-\beta-1) \frac{\beta}{\alpha+\beta} k_r - \frac{\alpha}{\alpha+\beta} (\alpha+1) \frac{\alpha}{\alpha+\beta+1} - k_s \right] \quad \text{in Case 4} \\
= \max & \left[ 0; \frac{\alpha}{\alpha+\beta} V (\alpha+1, \beta) + \frac{\beta}{\alpha+\beta} V (\alpha, \beta+1) + k_r + (M-\alpha-\beta-1) \frac{\alpha}{\alpha+\beta} (\alpha+1) \frac{\alpha}{\alpha+\beta+1} - k_s \right] \quad \text{in Case 5.}
\end{align*}
\]

Also,

\[(2.72) \quad V (\alpha; \beta) = 0 \quad \text{for } \alpha + \beta = M, \]

and furthermore, \((\alpha, \beta)\) is a stopping point if and only if

\[(2.73) \quad V (\alpha; \beta) = 0.\]

Since there can be at most \(M\) lattice points which belong to Cases 3, 4 and 5, it appears from a comparison of (2.29) and (2.71) that fewer arithmetic operations will be required if we use the latter set of recursion formulae.
2.3. Determination of Operating Characteristics for the Bayes Sequential Procedure.

It is of some interest to obtain the operating characteristics of the Bayes sequential procedures. Such curves have obvious interpretations to non-Bayesian statisticians.

**Determination of the OC-function.**

Now we indicate how to obtain the OC-function for our Bayes sequential procedures. Let \( P(N; \theta_0; \alpha_0; \beta_0) \) denote the probability of accepting the lot using the Bayes procedure with respect to \( \xi_N(\theta; \alpha_0; \beta_0) \) when \( \theta = \theta_0 \). This function plotted against different values of \( \theta \) will yield the Operating Characteristic curve.

In the previous section we indicated how to determine the optimal boundary and the stopping region in the \((\alpha, \beta)\)-plane. Let \( \Omega_1(\xi) \) denote the stopping region in the \((\alpha, \beta)\)-plane where the lot is accepted, and \( \Omega_2(\xi) \) denote the stopping region in the \((\alpha, \beta)\)-plane where the lot is rejected. We now denote \( P(\theta; N; \alpha_0, \beta_0) \) by \( P_M(\theta; \alpha_0, \beta_0) \) where \( M = N + \alpha_0 + \beta_0 \). Also we let \( P_M(\theta; \alpha, \beta) = P(\theta; M - \alpha - \beta; \alpha, \beta) \).

Using the techniques described in Section 2.2, \( \Omega_1(\xi) \) and \( \Omega_2(\xi) \) can be determined. The complement of \( \Omega_1(\xi) \cup \Omega_2(\xi) \) gives the continuation region which is known as soon as the stopping sets are obtained.

It is evident from the definition of \( P_M(\theta; \alpha, \beta) \) that for \( \theta = 0, 1, 2, \ldots, N, \)

\[
(2.7^k) \quad P_M(\theta; \alpha, \beta) = \begin{cases} 1 & \text{if } (\alpha, \beta) \in \Omega_1(\xi) \\ 0 & \text{if } (\alpha, \beta) \in \Omega_2(\xi). \end{cases}
\]
Denote the continuation region by $\Omega^c(\xi)$. We now seek a suitable formula for the determination of $P_M(\theta; \alpha, \beta)$ in the continuation region. We are aided in this process if we draw an analogy between the optimal sequential plan and the following random walk in the $(\alpha, \beta)$-plane.

A random walk in the $(\alpha, \beta)$-plane starts at the point $(\alpha_0, \beta_0)$. If $V(\alpha_0; \beta_0) = 0$, the walk ceases; if $V(\alpha_0; \beta_0) > 0$, $X_1$ is observed and the walk moves to $(\alpha_0 + 1, \beta_0)$ or to $(\alpha_0, \beta_0 + 1)$ according as $X_1 = 1$ or 0. After the first transition, the value of the function $V$ is computed at the new point. If $V = 0$, the walk ceases; if $V > 0$, $X_2$ is observed and the walk moves one step horizontally or vertically, as above. This procedure continues until a point is reached for which $V = 0$ at which point the random walk ceases.

From the correspondence between the optimal sequential procedure and the above random walk, we see that $P_M(\theta; \alpha, \beta)$ will be given by the following difference equation:

\[
(2.75) \\
P_M(\theta; \alpha, \beta) = p_\theta(\alpha; \beta) P_M(\theta; \alpha + 1, \beta) + [1 - p_\theta(\alpha; \beta)] P_M(\theta; \alpha, \beta + 1)
\]

for all $(\alpha, \beta)$ in $\Omega^c(\xi)$ and $\theta = 0, 1, 2, \ldots, N$.

In the above, $p_\theta(\alpha; \beta)$ denotes the probability of moving one step to the right given that the random walk is currently at the point $(\alpha, \beta)$, and the initial number of defectives at the point $(\alpha_0, \beta_0)$ was equal to $\theta$. 

When the process is at \((\alpha, \beta)\), the lot size is equal to \(M - \alpha - \beta\) and the corresponding number of defectives is \(\theta - (\alpha - \alpha_0)\). Also we note that \(\theta - (\alpha - \alpha_0)\) cannot be less than 0 or greater than \(M - \alpha - \beta\). Therefore we have

\[
(2.76)
\]

\[
p_\theta(\alpha; \beta) = \text{Prob} \left[ X = 1 \text{ given that the "random walk" is at the point } (\alpha, \beta) \right]
\]

\[
= \max \left\{ 0, \min \left[ \frac{\theta - (\alpha - \alpha_0)}{M - \alpha - \beta}, 1 \right] \right\}.
\]

To summarize, we see that \(P_M(\theta; \alpha, \beta)\) is given by (2.74) if \((\alpha, \beta)\) is in the stopping set, and is given by (2.75) if \((\alpha, \beta)\) is in the continuation region. Once the optimal boundary is known, we can determine the values of the OC-function in the continuation region for different values of \(\theta\) by a "backward induction" procedure similar to the one used for the construction of the optimum sequential rule.

**Derivation of the ASN-function**

The determination of the ASN-function is similar to that of the OC-function. Let \(n(\theta_0; N, \alpha_0, \beta_0)\) denote the expected sample size associated with the Bayes procedure with respect to \(\xi_M(\theta; \alpha_0, \beta_0)\) when \(\theta = \theta_0\). As before we denote \(n(\theta; N, \alpha_0, \beta_0)\) by \(n_M(\theta; \alpha_0, \beta_0)\). Similar to the expressions for the OC-function, we obtain the following equations for the ASN-function

\[
(2.77)
n_M(\theta; \alpha, \beta) = 0 \text{ if } (\alpha, \beta) \in \Omega_1(\xi) \cup \Omega_2(\xi)
\]

\[
= p_\theta(\alpha; \beta) n_M(\theta; \alpha+1, \beta) + [1-p_\theta(\alpha; \beta)]n_M(\theta; \alpha, \beta+1) + 1
\]

if \((\alpha, \beta) \in \Omega^c(\xi)\)

\((\theta = 0, 1, 2, \ldots, N)\)
where \( p_\theta (\alpha; \beta) \) is given by (2.75).

\[ n_M (\theta; \alpha_0, \beta_0) \] is obtained in a similar way as \( p_M (\theta; \alpha_0; \beta_0) \) was obtained by using a "backward induction" method starting from the optimal boundary and using (2.77) recursively.

Numerical values of the OC- and ASN-functions for different values of \( \theta \) for some representative sampling plans are given in the next section.

2.4. Some Numerical Examples of Bayes Sampling Inspection Plans, and their Properties.

In this section we list several different Bayes sequential sampling inspection plans. To facilitate comparison of these plans with the corresponding optimum single and double sampling plans, for the most part we confine ourselves to a description of the sequential plans for those parameter-values which have been considered by Pfanzagl. For each such Bayes sequential plan the following data are given:

1. Values of \( N, \alpha_0, \beta_0, M, k_r \) and \( k_s \).
2. The largest number of observations the plan can actually take.
4. A list of acceptance numbers \( \{A_n\} \) and rejection numbers \( \{R_n\} \) for the plan. \( A_n \) (\( R_n \)) are defined such that if
   \[
   \sum_{i=1}^{n} x_i = z_n \leq A_n \ (\geq R_n),
   \]
   then the plan stops sampling and accepts (rejects) the lot.
5. Values of the OC- and ASN-functions for selected values of \( \theta \).
6. The number of observations required to be taken by the corresponding optimum single sampling plan.
2.4.1. Determination of the Acceptance and Rejection Numbers.

In Section 2.2 we described how to obtain the optimum boundary in the \((\alpha, \beta)\)-plane for \(\alpha + \beta \leq M\). Following Chernoff and Ray [13], we define \(m = \alpha + \beta\). A direct correspondence can be set up between the points in the \((\alpha, \beta)\)-plane and those in the \((\alpha, m)\)-plane. If the apriori distribution of \(\theta\) is \(\xi_{\alpha_0, \beta_0}\), and if after taking \(n\) observations the a posteriori distribution of \(\theta^{(n)}\) is \(\xi_{\alpha', \beta'}\), then \(m = \alpha_0 + \beta_0 + n\), and \(\alpha = \alpha_0 + z_n\). Thus a set of stopping points on the optimal boundary in the \((\alpha, m)\)-plane will have co-ordinates \((\alpha_0 + z_n; \alpha_0 + \beta_0 + n)\) where for a fixed \(n\), \(z_n\) gives the acceptance or rejection number according as the point is in \(\Omega_1(\xi)\) or \(\Omega_2(\xi)\). The origin in the \((\alpha, m)\)-plane can now be shifted to the point \((\alpha_0; \alpha_0 + \beta_0)\) to yield the acceptance and rejection numbers for each \(n\).

2.4.2. Some Representative Plans.

We have computed ten optimal sequential sampling plans each one of which is associated with various possible values of the parameters \(N\), \(k_r\), \(k_s\), \(\alpha_0\) and \(\beta_0\). The particular parameter-values which were employed are listed below.

<table>
<thead>
<tr>
<th>Plan No.</th>
<th>(N)</th>
<th>(k_r)</th>
<th>(k_s)</th>
<th>(\alpha_0)</th>
<th>(\beta_0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>400</td>
<td>.01</td>
<td>.01</td>
<td>1</td>
<td>49</td>
</tr>
<tr>
<td>2</td>
<td>400</td>
<td>.02</td>
<td>.02</td>
<td>1</td>
<td>49</td>
</tr>
<tr>
<td>3</td>
<td>400</td>
<td>.01</td>
<td>.01</td>
<td>2</td>
<td>98</td>
</tr>
<tr>
<td>4</td>
<td>400</td>
<td>.02</td>
<td>.02</td>
<td>2</td>
<td>98</td>
</tr>
<tr>
<td>5</td>
<td>800</td>
<td>.01</td>
<td>.01</td>
<td>1</td>
<td>49</td>
</tr>
<tr>
<td>6</td>
<td>800</td>
<td>.02</td>
<td>.02</td>
<td>1</td>
<td>49</td>
</tr>
<tr>
<td>Plan No.</td>
<td>N</td>
<td>$k_r$</td>
<td>$k_b$</td>
<td>$\alpha_0$</td>
<td>$\beta_0$</td>
</tr>
<tr>
<td>---------</td>
<td>-----</td>
<td>-------</td>
<td>-------</td>
<td>-------------</td>
<td>-----------</td>
</tr>
<tr>
<td>7</td>
<td>800</td>
<td>0.01</td>
<td>0.01</td>
<td>2</td>
<td>98</td>
</tr>
<tr>
<td>8</td>
<td>800</td>
<td>0.02</td>
<td>0.02</td>
<td>2</td>
<td>98</td>
</tr>
<tr>
<td>9</td>
<td>1600</td>
<td>0.01</td>
<td>0.01</td>
<td>1</td>
<td>49</td>
</tr>
<tr>
<td>10</td>
<td>400</td>
<td>0.02</td>
<td>0.03</td>
<td>1</td>
<td>49</td>
</tr>
</tbody>
</table>

Plan 10 was not one of those which was considered by Pfanzagl, and therefore we do not give the number of observations taken by the optimal fixed sample size plan corresponding to this set of parameter-values.
PLAN 1

1. \(N = 400; \alpha_0 = 1; \beta_0 = 49; k_s = k_r = 0.01; M = 450\).

2. Largest number of observations the plan can take: 371.


4. Acceptance and Rejection numbers:
   The plan cannot accept before taking 120 observations.

\[
A_n = \begin{cases} 
0 & \text{for } 120 \leq n \leq 213 \\ 
1 & \text{for } 214 \leq n \leq 217 \\ 
2 & \text{for } 218 \leq n \leq 370 \\ 
3 & \text{for } n = 371 
\end{cases}
\]

The plan cannot reject before taking 4 observations.

\[
R_n = 4 \text{ for } n \geq 4
\]

5. Values of the OC- and ASN-functions for selected values of \(\theta\).

\[
\begin{array}{cccccccccc}
\theta & 0 & 4 & 8 & 12 & 20 & 32 & 40 & 200 & 400 \\
OC & 1.0000 & .7671 & .5342 & .5068 & .5003 & .5000 & .4835 & 0 & \\
ASN & 120.00 & 242.82 & 170.43 & 122.24 & 76.35 & 48.61 & 19.80 & 6.66 & 4.00 \\
\end{array}
\]

6. Number of observations taken by the Bayes single sampling plan: 113.
1. \( N = 400; \alpha_0 = 1; \beta_0 = 49; k_s = k_r = 0.02; M = 450. \)

2. The exact stage of truncation: 369.


4. Acceptance and Rejection numbers:

   The plan cannot accept before taking 51 observations.

   \[
   A_n = \begin{cases} 
   0 & \text{for } 51 \leq n \leq 103 \\ 
   1 & \text{for } 104 \leq n \leq 152 \\ 
   2 & \text{for } 153 \leq n \leq 199 \\ 
   3 & \text{for } 200 \leq n \leq 244 \\ 
   4 & \text{for } 245 \leq n \leq 267 \\ 
   5 & \text{for } 268 \leq n \leq 329 \\ 
   6 & \text{for } 330 \leq n \leq 368 \\ 
   7 & \text{for } n = 369. 
   \end{cases}
   \]

   The plan cannot reject before taking 7 observations

   \[
   R_n = \begin{cases} 
   7 & \text{for } 7 \leq n \leq 16 \\ 
   8 & \text{for } n \geq 17 
   \end{cases}
   \]

5. Values of the OC- and ASN-functions for selected values of \( \theta \):

<table>
<thead>
<tr>
<th>( \theta )</th>
<th>0</th>
<th>4</th>
<th>8</th>
<th>12</th>
<th>16</th>
<th>20</th>
<th>24</th>
<th>32</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>OC</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.8060</td>
<td>0.2977</td>
<td>0.1381</td>
<td>0.0700</td>
<td>0.0369</td>
<td>0.0108</td>
<td>0.0032</td>
</tr>
<tr>
<td>ASN</td>
<td>51.00</td>
<td>89.93</td>
<td>180.53</td>
<td>190.76</td>
<td>167.39</td>
<td>143.75</td>
<td>124.26</td>
<td>93.30</td>
<td>78.02</td>
</tr>
</tbody>
</table>

6. Number of observations taken by the Bayes single sampling plan: 73.
1. \( N = 400; \alpha_0 = 2; \beta_0 = 98; k_s = k_r = 0.01; M = 500. \)

2. Largest number of observations the plan can take: 337.


4. Acceptance and Rejection numbers:
   The plan cannot accept before taking 171 observations.
   
   \[
   A_n = \begin{cases} 
   0 & \text{for } 171 \leq n \leq 257 \\
   1 & \text{for } 258 \leq n \leq 336 \\
   2 & \text{for } n = 337 
   \end{cases}
   \]

   The plan cannot reject before taking 3 observations
   
   \[
   R_n = 3 \text{ for } n \geq 3
   \]

5. Values of the OC- and ASN functions for selected values of \( \theta \):

   \[
   \begin{array}{cccccccccccc}
   \theta & 0 & 4 & 8 & 12 & 16 & 20 & 24 & 32 & 40 \\
   OC & 1.0000 & 0.2369 & 0.0133 & 0.0011 & 0.0001 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
   ASN & 171.00 & 219.34 & 132.75 & 92.48 & 70.76 & 57.29 & 48.12 & 36.45 & 29.34 \\
   \end{array}
   \]

6. Number of observations taken by the corresponding Bayes single sampling plan: 164.
Plan 4

1. \( N = 400; \alpha_0 = 2; \beta_0 = 98; k_s = k_r = 0.02; M = 500. \)

2. Largest number of observations the plan can take: 369.


4. Acceptance and Rejection numbers:
   The plan cannot accept before taking 58 observations.

\[
A_n = \begin{cases} 
0 & \text{if } 58 \leq n \leq 106 \\
1 & \text{if } 107 \leq n \leq 153 \\
2 & \text{if } 154 \leq n \leq 200 \\
3 & \text{if } 201 \leq n \leq 244 \\
4 & \text{if } 245 \leq n \leq 288 \\
5 & \text{if } 289 \leq n \leq 329 \\
6 & \text{if } 330 \leq n \leq 368 \\
7 & \text{if } n = 369 
\end{cases}
\]

The plan cannot reject before taking 8 observations.

\[
R_n = 8 \text{ for } n \geq 8
\]

5. Values of the OC- and ASN-functions for selected values of \( \theta \).

<table>
<thead>
<tr>
<th>( \theta )</th>
<th>0</th>
<th>4</th>
<th>8</th>
<th>12</th>
<th>16</th>
<th>20</th>
<th>24</th>
<th>32</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>OC</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.7918</td>
<td>0.2560</td>
<td>0.1073</td>
<td>0.0487</td>
<td>0.0231</td>
<td>0.0055</td>
<td>0.0013</td>
</tr>
<tr>
<td>ASN</td>
<td>58.00</td>
<td>98.91</td>
<td>193.25</td>
<td>200.05</td>
<td>172.75</td>
<td>146.69</td>
<td>125.85</td>
<td>96.76</td>
<td>78.15</td>
</tr>
</tbody>
</table>

6. Number of observations taken by the corresponding Bayes single sampling plan: 78.
PLAN 5

1. $N = 800; \alpha_0 = 1; \beta_0 = 49; k_s = k_r = 0.01; M = 850.$
2. Largest number of observations the plan can take: 772.
4. Acceptance and Rejection numbers:

   The plan cannot accept before taking 150 observations.

   $A_n = 0$ if $150 \leq n \leq 254$
   $= 1$ if $255 \leq n \leq 351$
   $= 2$ if $352 \leq n \leq 444$
   $= 3$ if $445 \leq n \leq 534$
   $= 4$ if $535 \leq n \leq 617$
   $= 5$ if $618 \leq n \leq 699$
   $= 6$ if $700 \leq n \leq 771$
   $= 7$ if $n = 772$

   The plan cannot reject before taking 6 observations.

   $R_n = 6$ if $6 \leq n \leq 23$
   $= 7$ if $24 \leq n \leq 93$
   $= 8$ if $n \geq 94$

5. Values of the OC- and ASN-functions for selected values of $\theta$.

   \begin{tabular}{cccccccccc}
   $\theta$ & 0 & 8 & 16 & 24 & 32 & 40 & 48 & 64 & 80 \\
   OC & 1.0000 & 0.8136 & 0.5233 & 0.4955 & 0.4655 & 0.4107 & 0.3351 & 0.1768 & 0.0712 \\
   ASN & 150.00 & 48.86 & 363.39 & 254.27 & 192.04 & 152.44 & 125.03 & 90.31 & 70.28 \\
   \end{tabular}

6. Number of observations taken by the Bayes single sampling plan: 200.
PLAN 6

1. \( N = 800; \alpha_0 = 1; \beta_0 = 49; k_s = k_r = 0.02; M = 850. \)
2. Largest number of observations the plan can take: 769.
4. Acceptance and Rejection numbers:

   The plan cannot accept before taking 68 observations.

   \[
   A_n = 0 \text{ for } 68 \leq n \leq 125
   \]
   \[
   = 1 \text{ for } 126 \leq n \leq 178
   \]
   \[
   = 2 \text{ for } 179 \leq n \leq 228
   \]
   \[
   = 3 \text{ for } 229 \leq n \leq 278
   \]
   \[
   = 4 \text{ for } 279 \leq n \leq 326
   \]
   \[
   = 5 \text{ for } 327 \leq n \leq 374
   \]
   \[
   = 6 \text{ for } 375 \leq n \leq 421
   \]
   \[
   = 7 \text{ for } 422 \leq n \leq 467
   \]
   \[
   = 8 \text{ for } 468 \leq n \leq 513
   \]
   \[
   = 9 \text{ for } 514 \leq n \leq 558
   \]
   \[
   = 10 \text{ for } 559 \leq n \leq 603
   \]
   \[
   = 11 \text{ for } 604 \leq n \leq 646
   \]
   \[
   = 12 \text{ for } 647 \leq n \leq 689
   \]
   \[
   = 13 \text{ for } 690 \leq n \leq 730
   \]
   \[
   = 14 \text{ for } 731 \leq n \leq 768
   \]
   \[
   = 15 \text{ for } n = 769
   \]

   The plan cannot reject before taking 8 observations.

   \[
   R_n = 8 \text{ for } 8 \leq n \leq 21
   \]
   \[
   = 9 \text{ for } 22 \leq n \leq 46
   \]
   \[
   = 10 \text{ for } 47 \leq n \leq 76
   \]
   \[
   = 11 \text{ for } 77 \leq n \leq 112
   \]
   \[
   = 12 \text{ for } 113 \leq n \leq 156
   \]
   \[
   = 13 \text{ for } 157 \leq n \leq 211
   \]
   \[
   = 14 \text{ for } 212 \leq n \leq 283
   \]
   \[
   = 15 \text{ for } n \geq 284
   \]
5. Values of the OC- and ASN-functions for selected values of \( J \).

<table>
<thead>
<tr>
<th>( J )</th>
<th>0</th>
<th>8</th>
<th>16</th>
<th>24</th>
<th>32</th>
<th>40</th>
<th>48</th>
<th>54</th>
<th>80</th>
</tr>
</thead>
<tbody>
<tr>
<td>OC</td>
<td>1.000</td>
<td>1.000</td>
<td>.8395</td>
<td>.2053</td>
<td>.0749</td>
<td>.0309</td>
<td>.0135</td>
<td>.0027</td>
<td>.0005</td>
</tr>
<tr>
<td>ASN</td>
<td>68.00</td>
<td>132.97</td>
<td>368.24</td>
<td>425.67</td>
<td>358.38</td>
<td>289.96</td>
<td>235.31</td>
<td>154.56</td>
<td>123.91</td>
</tr>
</tbody>
</table>

6. Number of observations taken by the corresponding Bayes single sampling plan: 120.
1. \(N = 800; \alpha_0 = 2; \beta_0 = 98; k_g = k_r = 0.01; M = 900\).

2. Largest number of observations the plan can take: 737


4. Acceptance and Rejection numbers:

The plan cannot accept before taking 209 observations.

\[ A_n = 0 \text{ for } 209 \leq n \leq 306 \]
\[ = 1 \text{ for } 307 \leq n \leq 399 \]
\[ = 2 \text{ for } 400 \leq n \leq 489 \]
\[ = 3 \text{ for } 490 \leq n \leq 576 \]
\[ = 4 \text{ for } 577 \leq n \leq 659 \]
\[ = 5 \text{ for } 660 \leq n \leq 736 \]

The plan cannot reject before taking 6 observations.

\[ R_n = 6 \text{ for } 6 \leq n \leq 23 \]
\[ = 7 \text{ for } n \geq 24 \]

5. Values of the OC- and ASN-functions for selected values of \(\theta\).

<table>
<thead>
<tr>
<th>(\theta)</th>
<th>0</th>
<th>8</th>
<th>16</th>
<th>24</th>
<th>32</th>
<th>40</th>
<th>48</th>
<th>64</th>
<th>80</th>
</tr>
</thead>
<tbody>
<tr>
<td>OC</td>
<td>1.0000</td>
<td>0.3816</td>
<td>0.0114</td>
<td>0.0007</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>ASN</td>
<td>209.00</td>
<td>520.69</td>
<td>327.35</td>
<td>224.17</td>
<td>169.90</td>
<td>136.74</td>
<td>114.39</td>
<td>86.14</td>
<td>68.96</td>
</tr>
</tbody>
</table>

6. Number of observations taken by the Bayes single sampling plan: 261.
PLAN 8

1. \( N = 800; \alpha_0 = 2; \beta_0 = 96; k_r = k_s = 0.02; M = 900 \)

2. Largest number of observations the plan can take: 769

3. Value of the Bayes risk: 12.911162

4. Acceptance and Rejection numbers:

The plan cannot accept before taking 17 observations

\[
A_n = \begin{cases} 
0 & \text{for } 77 \leq n \leq 130 \\
1 & \text{for } 131 \leq n \leq 181 \\
2 & \text{for } 182 \leq n \leq 231 \\
3 & \text{for } 232 \leq n \leq 279 \\
4 & \text{for } 280 \leq n \leq 327 \\
5 & \text{for } 328 \leq n \leq 371 \\
6 & \text{for } 372 \leq n \leq 422 \\
7 & \text{for } 423 \leq n \leq 468 \\
8 & \text{for } 469 \leq n \leq 513 \\
9 & \text{for } 514 \leq n \leq 558 \\
10 & \text{for } 559 \leq n \leq 603 \\
11 & \text{for } 604 \leq n \leq 646 \\
12 & \text{for } 647 \leq n \leq 689 \\
13 & \text{for } 690 \leq n \leq 730 \\
14 & \text{for } 731 \leq n \leq 768 \\
15 & \text{for } n = 769 
\end{cases}
\]

The plan cannot reject before taking 10 observations

\[
R_n = \begin{cases} 
10 & \text{for } 10 \leq n \leq 19 \\
11 & \text{for } 20 \leq n \leq 53 \\
12 & \text{for } 54 \leq n \leq 93 \\
13 & \text{for } 94 \leq n \leq 140 \\
14 & \text{for } 141 \leq n \leq 198 \\
15 & \text{for } 199 \leq n \leq 273 \\
16 & \text{for } n \geq 274 
\end{cases}
\]
5. Values of the OC- and ASN-functions for selected values of $\theta$.

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>0</th>
<th>8</th>
<th>16</th>
<th>24</th>
<th>32</th>
<th>40</th>
<th>48</th>
<th>64</th>
<th>80</th>
</tr>
</thead>
<tbody>
<tr>
<td>OC</td>
<td>1.0000</td>
<td>1.0000</td>
<td>.8303</td>
<td>.1753</td>
<td>.0552</td>
<td>.0199</td>
<td>.0076</td>
<td>.0012</td>
<td>.0002</td>
</tr>
<tr>
<td>ASN</td>
<td>77.000</td>
<td>145.04</td>
<td>389.36</td>
<td>440.79</td>
<td>356.76</td>
<td>296.17</td>
<td>240.90</td>
<td>170.00</td>
<td>129.42</td>
</tr>
</tbody>
</table>

6. Number of observations taken by the Bayes single sampling plan: 128.
PL A N 9

1. $N = 1600; \alpha_0 = 1; \beta_0 = 49; k_r = k_s = 0.01; M = 1650.$

2. Largest number of observations the plan can take: 1572.


4. Acceptance and Rejection numbers:

The plan cannot accept before taking 184 observations.

\[
A_n = \begin{cases} 
0 & \text{for } 184 \leq n \leq 299 \\
1 & \text{for } 300 \leq n \leq 404 \\
2 & \text{for } 405 \leq n \leq 505 \\
3 & \text{for } 506 \leq n \leq 604 \\
4 & \text{for } 605 \leq n \leq 699 \\
5 & \text{for } 700 \leq n \leq 796 \\
6 & \text{for } 797 \leq n \leq 889 \\
7 & \text{for } 890 \leq n \leq 981 
\end{cases}
\]

\[
A_n = \begin{cases} 
8 & \text{for } 982 \leq n \leq 1072 \\
9 & \text{for } 1073 \leq n \leq 1162 \\
10 & \text{for } 1163 \leq n \leq 1250 \\
11 & \text{for } 1251 \leq n \leq 1336 \\
12 & \text{for } 1337 \leq n \leq 1420 \\
13 & \text{for } 1421 \leq n \leq 1499 \\
14 & \text{for } 1500 \leq n \leq 1571 \\
15 & \text{for } n = 1572 
\end{cases}
\]

The plan cannot reject before taking 7 observations.

\[
R_n = \begin{cases} 
7 & \text{for } 7 \leq n \leq 13 \\
8 & \text{for } 14 \leq n \leq 43 \\
9 & \text{for } 44 \leq n \leq 80 \\
10 & \text{for } 81 \leq n \leq 126 \\
11 & \text{for } 127 \leq n \leq 187 \\
12 & \text{for } 188 \leq n \leq 257 \\
13 & \text{for } 258 \leq n \leq 358 \\
14 & \text{for } 359 \leq n \leq 487 \\
15 & \text{for } 488 \leq n \leq 663 \\
16 & \text{for } n \geq 664 
\end{cases}
\]
5. Values of the OC- and ASN-functions for selected values of $\theta$:

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>0</th>
<th>16</th>
<th>32</th>
<th>48</th>
<th>64</th>
<th>80</th>
<th>96</th>
<th>128</th>
<th>160</th>
</tr>
</thead>
<tbody>
<tr>
<td>OC</td>
<td>1</td>
<td>0000</td>
<td>.8532</td>
<td>.3509</td>
<td>.0280</td>
<td>.0007</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
</tr>
<tr>
<td>ASN</td>
<td>184</td>
<td>00</td>
<td>937</td>
<td>17</td>
<td>733</td>
<td>20</td>
<td>454</td>
<td>96</td>
<td>314.76</td>
</tr>
</tbody>
</table>

6. Number of observations taken by the Bayes single sampling plan: **299**.
PLAN 10

1. \( N = 400; \alpha_0 = 1; \beta_0 = 49; k_r = .02; k_s = .03; M = 450. \)
2. Largest number of observations the plan can take: 210.
4. Acceptance and Rejection numbers:

   The plan cannot accept before taking 35 observations.

   \[ A_n = 0 \text{ for } 35 \leq n \leq 80 \]
   \[ = 1 \text{ for } 81 \leq n \leq 124 \]
   \[ = 2 \text{ for } 125 \leq n \leq 167 \]
   \[ = 3 \text{ for } 168 \leq n \leq 209 \]
   \[ = 4 \text{ for } n = 210 \]

   The plan cannot reject before taking 11 observations.

   \[ R_n = 1 \text{ for } 11 \leq n \leq 76 \]
   \[ = 2 \text{ for } 77 \leq n \leq 138 \]
   \[ = 3 \text{ for } 139 \leq n \leq 196 \]
   \[ = 4 \text{ for } n \geq 197 \]

5. Values of the OC- and ASN-functions for selected values of \( \theta \).

\[
\begin{array}{cccccccccc}
\theta & 0 & 4 & 8 & 12 & 16 & 20 & 24 & 32 & 40 \\
OC & 1.0000 & .8181 & .5799 & .3891 & .2563 & .1683 & .1108 & .0484 & .0212 \\
ASN & 35.00 & 38.97 & 36.64 & 32.37 & 27.99 & 24.08 & 20.77 & 15.77 & 12.34 \\
\end{array}
\]
2.4.3. Graphs of the Stopping Boundaries for Three Particular Plans.

In this section we exhibit graphs showing the optimal stopping boundaries in the \((\alpha, \beta)\)-plane for three sequential plans; using these graphs, one can obtain Plans 2 and 4 as well as others for which \(M = 450, 500\) or 100.

The parameters of the plans associated with the graphs are enumerated below:

<table>
<thead>
<tr>
<th>Graph No.</th>
<th>(M)</th>
<th>(k_r)</th>
<th>(k_s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>450</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>2</td>
<td>500</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>0.50</td>
<td>0.50</td>
</tr>
</tbody>
</table>

All three graphs show that the optimal boundaries for the plans under consideration are extremely asymmetric with respect to the neutral line. While the optimal acceptance boundary is approximately a straight line and is very close to the neutral boundary, the optimal rejection boundary is a convex curve and is much farther away from the neutral line than is the optimal acceptance boundary.
The Optimum Stopping Boundaries in the $(\alpha, \beta)$-plane for all Polya a priori distributions $\mathbb{P}_{\alpha, \beta}(\theta)$ with $\alpha + \beta \leq M = 450$ and $k_r = k_s = 0.02$. 

---

**GRAPH 1**

- **Optimal Acceptance Boundary**
- **Acceptance Region**
- **Continuation Region**
- **Neutral Line**
- **Rejection Region**
- **Optimal Rejection Boundary**
GRAPH 2

The Optimum Stopping Boundaries in the \((\alpha, \beta)\)-plane for all Polya apriori distributions \(\pi_n(\theta; \alpha, \beta)\) with \(\alpha + \beta \leq M = 500\) and \(k_r = k_s = 0.02\).
The Optimum Stopping Boundaries in the $(\alpha, \beta)$-plane for all Polya a priori distributions $\pi_\theta(\theta; \alpha, \beta)$ with $\alpha + \beta \leq M = 100$ and $k_r = k_s = 0.50$. 

**Acceptance Region**

**Continuation Region**

**Rejection Region**

**Optimal Acceptance Boundary**

**Neutral Line**

**Optimal Rejection Boundary**
2.5. Comparison of the Optimal Sequential Sampling Plans with the
Optimum Single and Double Sampling Plans.

Pfanzagl [33] made some comparisons between the optimum single
and double sampling plans for his model, and observed that the reduction
in the Bayes risk achieved by adopting a two-stage sampling plan rather
than a single-stage sampling plan is very small. Here we carry his analy-
ysis one step further and make similar comparisons among the corresponding
single, double and sequential sampling inspection plans. Table 2.5.1
gives the Bayes risks corresponding to the optimum single, double and se-
quential plans for certain selected values of the parameters considered
in the examples of the previous section. In the above table, we have,
following Pfanzagl [33], expressed the Bayes risks associated with the
different plans in terms of the ratios they bear to the total costs due to
complete sorting. The last column in this table gives the risks associ-
ated with the Ideal Procedure (see Hald [21]), which can only be applied
when the number of defectives is known in advance. The rule for accept-
ance or rejection of the lot according to the Ideal Procedure is

\[
\text{accept lot if } \theta \leq N_k r \\
\text{reject lot if } \theta > N_k r
\]

without taking any observations.

The values of the minimum average risks attainable with single and
double sampling plans have been taken from Pfanzagl [33]. The values of
the Bayes risks for six of the double sampling plans listed in this table
have not been provided by him, which explains the corresponding gaps in
the following table.
2.5.1. Comparison of the Bayes Risks.

**TABLE 2.5.1**

Bayes Risks Associated with the Optimal Single, Double and Sequential Plans for Pfanzagl's Model.

(Bayes Risks have been expressed as ratios to the cost due to complete sorting.)

<table>
<thead>
<tr>
<th>Lot Size N</th>
<th>( k_s )</th>
<th>( k_r )</th>
<th>Total cost due to complete sorting</th>
<th>( \alpha )</th>
<th>( \beta )</th>
<th>(Bayes risk)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Cost due to complete sorting)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Single</td>
</tr>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
</tr>
<tr>
<td>400</td>
<td>.01</td>
<td>.01</td>
<td>4</td>
<td>1</td>
<td>49</td>
<td>.9161</td>
</tr>
<tr>
<td>400</td>
<td>.01</td>
<td>.01</td>
<td>4</td>
<td>2</td>
<td>98</td>
<td>.9799</td>
</tr>
<tr>
<td>400</td>
<td>.02</td>
<td>.02</td>
<td>8</td>
<td>1</td>
<td>49</td>
<td>.7881</td>
</tr>
<tr>
<td>400</td>
<td>.02</td>
<td>.02</td>
<td>8</td>
<td>2</td>
<td>98</td>
<td>.8550</td>
</tr>
<tr>
<td>800</td>
<td>.01</td>
<td>.01</td>
<td>8</td>
<td>1</td>
<td>49</td>
<td>.8876</td>
</tr>
<tr>
<td>800</td>
<td>.01</td>
<td>.01</td>
<td>8</td>
<td>2</td>
<td>98</td>
<td>.9650</td>
</tr>
<tr>
<td>800</td>
<td>.02</td>
<td>.02</td>
<td>16</td>
<td>1</td>
<td>49</td>
<td>.7388</td>
</tr>
<tr>
<td>800</td>
<td>.02</td>
<td>.02</td>
<td>16</td>
<td>2</td>
<td>98</td>
<td>.8301</td>
</tr>
<tr>
<td>1600</td>
<td>.01</td>
<td>.01</td>
<td>16</td>
<td>1</td>
<td>49</td>
<td>.8643</td>
</tr>
<tr>
<td>1600</td>
<td>.01</td>
<td>.01</td>
<td>16</td>
<td>2</td>
<td>98</td>
<td>.9504</td>
</tr>
<tr>
<td>1600</td>
<td>.02</td>
<td>.02</td>
<td>32</td>
<td>1</td>
<td>49</td>
<td>.7131</td>
</tr>
<tr>
<td>1600</td>
<td>.02</td>
<td>.02</td>
<td>32</td>
<td>2</td>
<td>98</td>
<td>.8078</td>
</tr>
</tbody>
</table>
From Table 2.5.1, we see that for the model of Pfanzagl and the values of the cost and distribution parameters we have considered, the reduction in the Bayes risk obtained by adopting the optimal sequential plan rather than the optimal single or double sampling plan is very small. This is in agreement with the conclusions of Moriguti [31] and Breakwell [8], who seem to have found not much difference between the optimal one-stage and sequential plans for their model of sampling inspection by attributes. However, it should be pointed out that for many other formulations of the sampling inspection problem, e.g., those considered by Wald [40], Weiss [42], etc., the sequential plans have been shown superior to the corresponding single sampling plans. But as we have noted in Section 1.2, the criterion of "goodness" that has been adopted in the latter situations is the ASN, when the probabilities of incorrect decisions guaranteed by both sets of procedures do not exceed certain preassigned values at two given values of the parameter \( \theta \). If a comparison is now made between our optimal sequential and single sampling plans on the basis of ASN, it is seen from our computations (see Section 2.4) that the ASN for the sequential procedure is considerably lower over a large range of \( \theta \)-values as compared to the corresponding fixed sample size of the one-stage procedure.

Table 2.5.1 also gives some information on the robustness of the Bayes sequential procedures when the parameters of the apriori distributions are changed. For a fixed \( N \), we recall that \( \xi_N(\theta; 1, 49) \) and \( \xi_N(\theta; 2, 98) \) have the same expected value although they differ otherwise in shape and skewness. It may be argued that in many situations the expected value will be approximately known although the entire distribution function may not be known. Therefore, under the assumption that the number of defectives has
a Polya prior distribution, the difference in the magnitudes of the Bayes risks derived with respect to the above two sets of prior distributions under the same set of cost parameters may give an indication as to the robustness of these Bayes procedures, when the only prior information available is the value of the expected number of defectives in the lot. Considering the fact that \( \xi_N(\theta; 1, 49) \) and \( \xi_N(\theta; 2, 98) \) represent two widely different distributions (see Figures 1 and 2 of Pfanzagl [33]), the difference between the sets of Bayes risks for either the optimal sequential plans or the optimal single sampling plan does not appear to be very great. In any of the above examples, the difference in the two sets of Bayes risks does not exceed 10% of the total cost due to sorting.

2.5.2. Graphs of the OC- and ASN-Curves for an Optimal Sequential Procedure, and the Corresponding Fixed Sample Procedure.

In this section we exhibit two graphs which show the OC- and ASN-functions corresponding to the optimal sequential and single sampling plans for one particular case, namely \( N = 400, k_A = k_S = .02, \alpha_0 = 1 \) and \( \beta_0 = 49 \).

It is to be noted that the ASN for the optimal sequential plan is not uniformly (in \( \theta \)) lower than the corresponding optimal fixed sample size procedure although the \( \theta \)-region over which the sequential ASN exceeds the fixed sample size is small. On the other hand, the OC-function of the optimal sequential plan is higher than that of the corresponding fixed sample procedure for small values of the fraction defective but tapers off to zero at a much faster rate for higher values of the fraction defective. Thus the optimal sequential plan can be thought of having a sharper discriminatory ability than does the corresponding optimal one-stage plan.
The ASN-curves associated with the Optimum Sequential and Fixed-Sample Procedures with respect to \( \mu_{400} (\theta;1,49) \) when \( k_r = k_s = 0.02 \)
The OC-curves associated with the Optimum Sequential and Fixed-Sample Procedures with respect to $\xi_{400} (\theta;1,49)$ when $k_r = k_s = 0.02$. 

Optimal Sequential Procedure

Optimum Fixed Sample Procedure

Number of defectives
3.1. Introduction.

In recent years, investigations on the limiting behavior of optimal sequential decision procedures which tend to require large samples have been carried out quite extensively in the literature. In the case of random variables from an infinite population, the sample size is not one of the given parameters of a sequential decision problem and therefore the usual approach that has been taken for the large sample theory is to let the cost of taking an observation tend to zero. Following this approach, Schwarz [37] has developed a theory of the asymptotic shape of the Bayes stopping region relative to an apriori distribution $F$ for testing sequentially between the composite hypothesis $\theta \leq \theta_1$ and $\theta \geq \theta_2$ where $\theta$ is the real parameter of a distribution of the Koopman-Darmois type with indifference region in the open interval $(\theta_1, \theta_2)$. His results have been considerably generalized by Kiefer and Sacks [28] to distributions having any arbitrary form and cases where more than two decisions are permitted. On the other hand, Chernoff has treated the special case of testing sequentially the composite hypothesis $\mu < \mu_0$ vs. $\mu > \mu_0$ where $\mu$ is the mean of a Normal distribution with known variance $\sigma^2$, and the apriori distribution of $\mu$ is also normal. When the cost of taking individual observations is very small compared to the loss due to making a wrong terminal decision, the Bayes test takes a large number of observations with
high probability. For such a situation, it is argued by Chernoff that the asymptotic theory can be developed as if the sampling is continuous, and that the optimal plan is not greatly affected by making such an assumption. Thus when the cost tends to zero, observations from the Normal distribution with mean \( \mu \) and variance \( \sigma^2 \) can be considered to have arisen from a Wiener Process with drift \( \mu \) and variance \( \sigma^2 \) per unit time. Chernoff [11] has shown that the problem of determining the optimal boundary and the Bayes risks for testing sequentially between the composite hypothesis \( \mu \leq \mu_0 \) and \( \mu \geq \mu_0 \) for a Wiener Process with known variance can be reduced to that of solving a free boundary problem involving a partial differential equation. This reduction is exact whatever may be the cost of taking an observation or loss due to making wrong terminal decisions, and therefore the solution to the free boundary problem (which we will sometimes refer to as FBP in the sequel) will approximately yield the optimal boundary for sequentially testing the mean value of a Normal distribution when the cost tends to zero. Solutions for the FBP which arise when in particular the cost of taking an observation is constant and the loss function is linear in the neighborhood of \( \mu_0 \) have been given in the form of formal series expansions by Breakwell and Chernoff [9], Chernoff [12] and Bather [6]. Similar free boundary problems were also obtained by Moriguti and Robbins [32], and Ray [35] in the context of testing the binomial parameter with a Beta apriori distribution. Instead of letting the cost tend to zero, the latter authors took an equivalent approach of making the terminal loss tend to infinity and derived certain formal
series expansions for the optimal boundaries and the Bayes risks for their problems.

For the particular problem we have considered in Chapter II, however, the population is finite and the sample size can never exceed \( N \), the lot size. Thus a reasonable approach for obtaining asymptotic results in our case appears to be that of letting the lot size tend to infinity, which has also the effect of making the terminal losses very large for any finite stage of sampling. Chernoff and Ray [13] used this idea and were able from some heuristic considerations to reduce the problem of finding the asymptotic optimal boundary and the Bayes risk for their one-action sampling inspection problem to that of solving a FBP involving a diffusion equation. They also gave formal series expansions representing the asymptotic behavior of the optimum boundary and the Bayes risks, and indicated that these expansions are asymptotic in the sense that at any stage the error due to truncation is less than the last term. Here, we try to adapt their development for a generalization to the two-decision problem which we have considered so far. We succeed in similarly obtaining a free boundary problem related to the Bayes solution for our problem when the lot size is very large. Unfortunately, however, it turns out that the solution to this particular FBP gives rise to unexpected results. We shall comment in detail on the implications of these results later in this chapter after having derived the FBP and its solution.

From the examples we have considered in Section 2.4, and from physical considerations, it is apparent that as the lot size becomes larger, more and more sampling is required to reach the optimum boun-
boundary. We may therefore expect that as the lot size approaches infinity, the optimal boundary will also approach infinity. Therefore, in order to study the limiting tendency (if any) of the Bayes solution, a suitable normalizing transformation of the co-ordinates is required. Thus the first question to be resolved in the asymptotic behavior of the optimal boundary is what constitutes a suitable normalization which will lead to non-degenerate limits. Chernoff (see Section 3, [11]) has indicated how in general one can find suitable transformations for testing the drift of a Wiener Process. It turns out that for our problem the transformation used by Chernoff and Ray [13] is equally applicable.

In the sequel, we obtain the free boundary problem associated with the Bayes solution as $N$ approaches infinity. In our consideration, we shall confine ourselves to the case when $k_r = k_s = k$.

3.2. Free Boundary Problem Associated with Bayes Sequential Procedures for Pfanzagl's Model when the Lot Size Approaches Infinity

From (2.28) and (2.29) we have the following recursion relations for the Bayes risk when $k_s = k_r = k$:

\begin{equation}
(3.1) \quad \rho^* (\alpha; \beta) = 0 \quad \text{for} \quad \alpha + \beta = M
\end{equation}

\[ = \min \left[ \rho_0 (\alpha, \beta); \frac{\alpha}{\alpha + \beta} \rho^* (\alpha+1, \beta) + \frac{\beta}{\alpha + \beta} \rho^* (\alpha, \beta+1) + k \right] \]

for $\alpha + \beta < M$

where $\rho_0 (\alpha, \beta) = \min \left\{ (M - \alpha - \beta) \frac{\alpha}{\alpha + \beta}; (M - \alpha - \beta) k \right\}$

and $M = N + \alpha_0 + \beta_0$. 
Further, \((\alpha, \beta)\) is a stopping point if and only if \(\rho* (\alpha; \beta) = \rho_{0} (\alpha; \beta)\).

As before we can write

\[
\rho_{0} (\alpha; \beta) = \frac{1}{2} (M-\alpha-\beta) | \frac{\alpha}{\alpha + \beta} + k | - \frac{1}{2} (M-\alpha-\beta) \left| \frac{\alpha}{\alpha + \beta} - k \right|.
\]

We now define

\[
W_{0} (\alpha; \beta) = (M - \alpha - \beta) \left| \frac{\alpha}{\alpha + \beta} - k \right|
\]

and

\[
W (\alpha; \beta) = (M-\alpha-\beta) \left( \frac{\alpha}{\alpha + \beta} + k \right) - 2p^{m} (\alpha; \beta).
\]

In terms of these transformed functions (2.31) reduces to the following:

\((\alpha, \beta)\) is a stopping point if and only if

\[
W (\alpha; \beta) = W_{0} (\alpha; \beta).
\]

Substituting (3.2) and (3.3) in (2.29), we obtain

\[
\frac{1}{2} (M - \alpha - \beta) \left( \frac{\alpha}{\alpha + \beta} + k \right) - \frac{1}{2} W (\alpha; \beta)
\]

\[= \min \left( \frac{1}{2} (M-\alpha-\beta) \left( \frac{\alpha}{\alpha + \beta} + k \right) - \frac{1}{2} W_{0} (\alpha, \beta); \right.
\]

\[
\left. \frac{\alpha}{\alpha + \beta} \left[ \frac{1}{2} (M-\alpha-\beta-1) \left( \frac{\alpha}{\alpha + \beta + 1} + k \right) - \frac{1}{2} W (\alpha + 1; \beta) \right] \right. 
\]

\[+ \frac{\beta}{\alpha + \beta} \left[ \frac{1}{2} (M-\alpha-\beta-1) \left( \frac{\alpha}{\alpha + \beta + 1} + k \right) - \frac{1}{2} W (\alpha; \beta + 1) \right] + k \]

or

\[
- \frac{1}{2} W (\alpha, \beta) = \min \left[ - \frac{1}{2} W_{0} (\alpha, \beta); \frac{1}{2} \frac{\alpha}{\alpha + \beta} W (\alpha + 1, \beta) \right.
\]

\[\left. - \frac{1}{2} \frac{\beta}{\alpha + \beta} W (\alpha; \beta + 1) + k - \frac{1}{2} \left( k + \frac{\alpha}{\alpha + \beta} \right) \right] \]
or

\[(3.7) \quad W(\alpha; \beta) = \max \left[ W_0(\alpha; \beta), \frac{\alpha}{\alpha + \beta} W(\alpha + 1; \beta) + \frac{\beta}{\alpha + \beta} W(\alpha; \beta + 1) + \left( \frac{\alpha}{\alpha + \beta} - k \right) \right].\]

In the continuation region the following equality holds:

\[(3.8) \quad W(\alpha; \beta) = \frac{\alpha}{\alpha + \beta} W(\alpha + 1; \beta) + \frac{\beta}{\alpha + \beta} W(\alpha; \beta + 1) + \frac{\alpha}{\alpha + \beta} - k.\]

Following Chernoff and Ray [13], we now normalize with the following transformations:

\[(3.9a) \quad z = \frac{\alpha - (\alpha + \beta)k}{\sqrt{Mk} (1 - k)} \]
\[(3.9b) \quad t = \frac{\alpha + \beta}{M} \]
\[(3.9c) \quad B(z; t) = \frac{W(\alpha; \beta)}{\sqrt{Mk} (1 - k)} \]

It is evident for our particular problem, that \( t \) can take values between 0 and 1 only. We can write (3.8) as

\[(3.10) \quad W(\alpha; \beta) = \frac{1}{2} \left[ W(\alpha + 1; \beta) + W(\alpha; \beta + 1) \right] + \frac{\alpha - \beta}{2 (\alpha + \beta)} \left[ W(\alpha + 1; \beta) - W(\alpha; \beta + 1) \right] + \frac{\alpha - (\alpha + \beta)k}{(\alpha + \beta)}.\]

In the immediate sequel we use the symbol "\( \rightarrow \)" to denote "is replaced by." Thus, when \( \alpha \rightarrow \alpha + 1 \) or \( \beta \rightarrow \beta + 1 \), we obtain the following changes in \( z \) and \( t \).
\[ (3.11) \quad \alpha \rightarrow \alpha + 1; \quad z \rightarrow \frac{\alpha + 1 - k (\alpha + \beta + 1)}{\sqrt{M} (1 - k)} = z + \frac{1 - k}{\sqrt{Mk} (1 - k)}; \quad t \rightarrow t + \frac{1}{M} \]

\[ (3.12) \quad \beta \rightarrow \beta + 1; \quad z \rightarrow \frac{\alpha - k (\alpha + \beta + 1)}{\sqrt{Mk} (1 - k)} = z - \frac{k}{\sqrt{Mk} (1 - k)}; \quad t \rightarrow t + \frac{1}{M}. \]

Let
\[ (3.13) \quad \delta_1 = \frac{(1 - k)}{\sqrt{k} (1 - k)} \]
and
\[ (3.14) \quad \delta_2 = -\frac{k}{\sqrt{(1 - k) k}}. \]

Also from (3.9c), (3.11) and (3.12) we obtain

\[ (3.15) \quad W(\alpha + 1; \beta) = \sqrt{M} (1 - k) B \left( z + \frac{1 - k}{\sqrt{M} (1 - k)} ; t + \frac{1}{M} \right). \]

\[ = \sqrt{Mk} (1 - k) B \left( z + \delta_1 M ; t + \frac{1}{M} \right) \]
and
\[ (3.16) \quad W(\alpha; \beta + 1) = \sqrt{M} (1 - k) B \left( z - \frac{k}{\sqrt{Mk} (1 - k)} ; t + \frac{1}{M} \right). \]

\[ = \sqrt{Mk} (1 - k) B \left( z + \delta_2 M ; t + \frac{1}{M} \right). \]

Further
\[ (3.17) \quad \frac{z}{\sqrt{M} \sqrt{Mk (1 - k)}} = \frac{\alpha - (\alpha + \beta) k}{M} \]
\[ = \frac{\alpha - (\alpha + \beta) k}{(\alpha + \beta) \sqrt{k} (1 - k)} \]
and
\[ (3.18) \quad \frac{2 z}{\sqrt{M} \sqrt{k (1 - k)}} = \frac{\alpha - (\alpha + \beta) k}{\alpha + \beta}. \]
Therefore,

\[(3.19) \quad \frac{2z\sqrt{k}(1-k)}{\sqrt{Mt}} + (2k - 1) = 2 \left[ \frac{\alpha - (\alpha + \beta)k}{\alpha + \beta} \right] + (2k - 1) = \frac{\alpha - \beta}{\alpha + \beta}.\]

Now using (3.15), (3.16) and (3.19), we can rewrite (3.10) as follows:

\[(3.20) \quad B(z; t) = \frac{1}{2} \left[ B(z + \delta_1 M^{-\frac{1}{2}}; t + M^{-1}) + B(z + \delta_2 M^{-1}; t + M^{-1}) \right] + \frac{1}{3} \left[ (2k - 1) + 2\sqrt{k(1-k)} \frac{2M^{-\frac{1}{2}}}{t} \right] \left[ B(z + \delta_1 M^{-\frac{1}{2}}; t + M^{-1}) - B(z + \delta_2 M^{-1}; t + M^{-1}) \right] + \frac{2}{t} M^{-1} B_{tt}(z; t).\]

We now assume that the function \(B(z; t)\) is continuous everywhere and is sufficiently differentiable with respect to both of its arguments for every point in the continuation set. Expanding the right hand side of (3.20) in a Taylor series about the point \((z, t)\), we obtain

\[(3.21) \quad B(z; t) = \frac{1}{2} \left( B(z; t) + \delta_1 M^{-\frac{1}{2}} B_z(z; t) + M^{-1} B_t(z; t) \right) + \frac{1}{3} \left[ \delta_1^2 M^{-1} B_{zz}(z; t) + 2\delta_1 M^{-3/2} B_{zt}(z; t) + M^{-2} B_{tt}(z; t) \right] + \frac{2}{t} M^{-1} B_{tt}(z; t) + O(M^{-2})\]

We now write \(B(z; t)\) as a sum of two terms, with the main term containing the leading terms of the Taylor expansion and the remainder term containing all other terms.

\[B(z; t) = \text{main term} + \text{remainder}.\]
\[
+ \frac{1}{4} \left( (2k - 1) + 2 \sqrt{k(1-k)} \frac{z}{t} M^{-\frac{1}{3}} \right) \left( B(z; t) + \delta_1 M^{-\frac{1}{3}} B_z(z; t) \right) + M^{-1} B_t(z; t) + \frac{1}{2} \left[ \delta_1 M^{-1} B_{zz}(z; t) \right] + 2 \delta_1 M^{-\frac{3}{2}} B_{zt}(z; t) + M^{-2} B_{tt}(z; t) \right) - B(z; t) \\
- \delta_2 M^{-\frac{1}{3}} B_z(z; t) - M^{-1} B_t(z; t) - \frac{1}{2} \left[ \delta_2 M^{-1} B_{zz}(z; t) \right] + 2 \delta_2 M^{-\frac{3}{2}} B_{zt}(z; t) + O(M^{-2}) + \frac{z}{t} M^{-1}
\] or
\[
= B(z; t) + \frac{1}{2} \left( \delta_1 + \delta_2 \right) M^{-\frac{1}{3}} B_z(z; t) + M^{-1} B_t(z; t) \]
\[
+ \frac{1}{4} \left( \delta_1^2 + \delta_2^2 \right) M^{-1} B_{zz}(z; t) + \frac{1}{2} \left( \delta_1 - \delta_2 \right) M^{-\frac{3}{2}} B_{zt}(z; t) \\
+ \frac{1}{2} (2k - 1) \left[ (\delta_1 - \delta_2) M^{-\frac{1}{3}} B_z(z; t) + \frac{1}{2} (\delta_1^2 - \delta_2^2) M^{-1} B_{zz}(z; t) \right] + \frac{z}{t} M^{-1} + O(M^{-2}).
\]

Now collecting terms and simplifying we obtain
\[
(3.22)
0 = \frac{1}{2} M^{-1} B_z(z; t) \left[ (\delta_1 + \delta_2) + (2k - 1) (\delta_1 - \delta_2) \right] + M^{-1} B_t(z; t) \\
+ \frac{1}{4} M^{-1} B_{zz}(z; t) \left[ (\delta_1^2 + \delta_2^2) + (2k - 1) (\delta_1^2 - \delta_2^2) \right] \\
+ \frac{z}{t} M^{-1} + O(M^{-3}).
\]
Recalling the definition of $\delta_1$ and $\delta_2$ in (3.13) and (3.14), we obtain

$$\delta_1 + \delta_2 + (2k - 1)(\delta_1 - \delta_2)$$

$$= 1 - 2k + \frac{2k - 1}{\sqrt{k(1-k)}} = 0.$$  

and

$$(\delta_1^2 + \delta_2^2) + (2k - 1)(\delta_1^2 - \delta_2^2)$$

$$= \frac{(1-k)^2 + k^2 + (2k - 1)[(1-k)^2 - k^2]}{k(1-k)}$$

$$= \frac{1 - 2k + 2k^2 + (2k - 1)(1 - 2k)}{k(1-k)}$$

$$= \frac{2k(1-k)}{k(1-k)} = 2.$$  

Therefore (3.22) reduces to:

$$\mathbf{M}^{-1} B_t (z; t) + \frac{1}{2} \mathbf{M}^{-1} B_{zz} (z; t) + \frac{z}{t} \mathbf{M}^{-1} B_z (z; t) + \frac{z}{t} \mathbf{M}^{-1} = 0 (\mathbf{M}^{\frac{3}{2}})$$

or

$$(3.23) \quad \frac{1}{2} B_{zz} (z; t) + \frac{z}{t} B_z (z; t) + B_t (z; t) + \frac{z}{t} = 0 (\mathbf{M}^{-\frac{1}{2}}).$$

Thus in the limit as $M \to \infty$, $B(z; t)$ satisfies the following differential equation for every point in the continuation set in the $(z,t)$-plane:

$$(3.24) \quad \frac{1}{2} B_{zz} + \frac{z}{t} B_z + B_t + \frac{z}{t} = 0.$$  

Let the upper boundary of the continuation region be denoted by $\tilde{z}^+(t)$ and the lower boundary be denoted by $\tilde{z}^-(t)$. 
A boundary condition at $z = z^+(t)$ and $z = z^-(t)$ is given by the assumption of continuity of the function $B(z; t)$. Equation (2.31), rewritten in terms of the functions $W$ and $W_0$, reads as follows:

$$W(\alpha; \beta) = W_0(\alpha; \beta) \text{ if and only if } (\alpha; \beta) \text{ is a stopping point, so that}$$

$$W(\alpha; \beta) = \frac{(M - \alpha - \beta)}{\sqrt{Mk}} \frac{\alpha}{(1 - k)} \left| \frac{\alpha}{\alpha + \beta} - k \right|$$

$$= \frac{M - \alpha - \beta}{\alpha + \beta} \frac{\alpha - (\alpha + \beta) k}{\sqrt{Mk}} (1 - k) .$$

In particular, when $(\alpha, \beta)$ is on the boundary, we obtain from the continuity assumption and (3.9a), (3.9b), and (3.9c) that

$$B(z; t) = \frac{1 - \frac{t}{t}}{|z|} \text{ for } (z^+(t), t) \text{ or } (z^-(t), t).$$

Another set of boundary conditions is obtained from the following heuristic argument.

Let $(\alpha, \beta)$ be a stopping point on the lower optimal boundary in the $(\alpha, \beta)$-plane so that $(\alpha + 1, \beta)$ is a stopping point while $(\alpha, \beta + 1)$ is a point within the continuation region. Then we expect that for points on the boundary the two quantities on the right hand side of (3.7) will be approximately equal. Now from the fact that the points $(\alpha, \beta)$ and $(\alpha + 1, \beta)$ are in the stopping set, and in view of (3.4), the following is true:

$$W(\alpha; \beta) = W_0(\alpha; \beta) \sim \frac{\alpha}{\alpha + \beta} W_0(\alpha + 1; \beta) + \frac{\beta}{\alpha + \beta} W(\alpha; \beta + 1) + \left( \frac{\alpha}{\alpha + \beta} - k \right).$$
In the sequel we treat (3.27) as an equality. Since \((\alpha, \beta)\) is a point on the lower optimal boundary and \((\alpha + 1, \beta)\) is also a stopping point, it follows from the convexity of the stopping region that \(k < \frac{\alpha}{\alpha + \beta}\) and therefore

\[
(3.28) \quad \frac{\alpha + 1}{\alpha + \beta + 1} - k > \frac{\alpha}{\alpha + \beta} - k > 0.
\]

From (3.2), (3.27), and (3.28) it now follows that

\[
(3.29) \quad (M - \alpha - \beta) \left(\frac{\alpha}{\alpha + \beta} - k\right) = \frac{\alpha}{\alpha + \beta} (M - \alpha - \beta - 1) \left(\frac{\alpha + 1}{\alpha + \beta + 1} - k\right)
\]

\[+ \frac{\beta}{\alpha + \beta} W(\alpha; \beta + 1) + \left(\frac{\alpha}{\alpha + \beta} - k\right)\]

or

\[
(M - \alpha - \beta - 1) \left(\frac{\alpha - (\alpha + \beta) k}{\alpha + \beta}\right) = \frac{\alpha}{(\alpha + \beta)} \frac{(M - \alpha - \beta - 1)}{(\alpha + \beta + 1)}
\]

\[+ W(\alpha; \beta + 1).\]
Now using (3.9a), (3.9b) and (3.16), we can write (3.29) in terms of \( t, z \) and \( B \) as follows:

\[
(3.30) \quad \frac{M(1 - t) - 1}{Mt} z = \frac{M(1 - t) - 1}{Mt(Mt + 1)} \left[ z \sqrt{\frac{1 - k}{Mk}} + \frac{z}{\sqrt{Mk}} \left( 1 - k \right) \right]
\]

\[
+ B \left( z + \delta^2 M^{-\frac{1}{2}} ; t + M^{-1} \right)
\]

or

\[
\frac{1 - t}{t} z - \frac{z}{Mt} = \frac{1 - t}{t(Mt + 1)} z
\]

\[
+ \left( \frac{M(1 - t) k}{(Mt + 1) \sqrt{Mk(1 - k)}} \right) \left( - \frac{k}{t(Mt + 1) \sqrt{Mk(1 - k)}} \right) \left( - \frac{z}{Mt(Mt + 1)} \right)
\]

\[
+ B \left( z + \delta^2 M^{-\frac{1}{2}} ; t + M^{-1} \right).
\]

As before we expand \( B \left( z + \delta^2 M^{-\frac{1}{2}} ; t + M^{-1} \right) \) in a Taylor Series. Further from (3.28) and (3.9a) it is evident that for this particular case, \( z \) is positive and therefore from (3.26) the value of \( B(z, t) \) at the point \((z, t)\) on the optimal boundary (in the \((z, t)\)-plane) is given by

\[
(3.31) \quad B(z, t) = \frac{1 - t}{t} z.
\]

Therefore (3.30) can now be written as follows:

\[
\frac{1 - t}{t} z = \frac{1 - t}{t} \frac{k}{\sqrt{k(1 - k)}} M^{-\frac{1}{2}} + \frac{1 - t}{t} z + \delta^2 M^{-\frac{1}{2}} B_z(z, t) + O(M^{-1}).
\]
From (3.14) and (3.32), it follows that

\[(3.33)\quad B_z (z; t) = \frac{1 - t}{t} + O \left( M^{-\frac{1}{2}} \right).\]

Therefore as \( M \to \infty \), we expect that

\[(3.34)\quad B_z (z; t) = \frac{1 - t}{t} \text{ on the boundary } (Z^+ (t), t).\]

The other boundary condition is derived similarly. Let \((\alpha, \beta)\) be a point on the upper optimal boundary in the \((\alpha, \beta)\)-plane so that \((\alpha, \beta + 1)\) is also a stopping point and \((\alpha + 1, \beta)\) is a point lying within the continuation set. Then, as before, we treat the following as an equality.

\[(3.35)\quad \mathcal{W}_0 (\alpha; \beta) = \frac{\alpha}{\alpha + \beta} \mathcal{W} (\alpha + 1; \beta) + \frac{\beta}{\alpha + \beta} \mathcal{W}_0 (\alpha; \beta + 1) + \left( \frac{\alpha}{\alpha + \beta} - k \right).\]

Since \((\alpha, \beta)\) is now a point on the optimal boundary where the decision \(a_1\) is accepted, it follows from similar convexity considerations that

\[(3.36)\quad \mathcal{W}_0 (\alpha; \beta) = (M - \alpha - \beta) \left| \frac{\alpha}{\alpha + \beta} - k \right| = (M - \alpha - \beta) \left( k - \frac{\alpha}{\alpha + \beta} \right)\]

and

\[(3.37)\quad \mathcal{W}_0 (\alpha; \beta + 1) = (M - \alpha - \beta - 1) \left| \frac{\alpha}{\alpha + \beta + 1} - k \right| = (M - \alpha - \beta - 1) \left( k - \frac{\alpha}{\alpha + \beta + 1} \right).\]

Substituting these values in (3.36), we obtain

\[(3.38)\quad (M - \alpha - \beta) \left( k - \frac{\alpha}{\alpha + \beta} \right) = \frac{\alpha}{\alpha + \beta} \mathcal{W} (\alpha + 1; \beta) + \frac{\beta}{\alpha + \beta} (M - \alpha - \beta - 1) \left( k - \frac{\alpha}{\alpha + \beta + 1} \right) + \left( \frac{\alpha}{\alpha + \beta} - k \right).\]
or

\[
(M - \alpha - \beta - 1) \left( k - \frac{\alpha}{\alpha + \beta} \right) + 2 \left( k - \frac{\alpha}{\alpha + \beta} \right) \frac{\alpha + \beta}{\alpha} \]

\[(3.39)\]

\[= W(\alpha + 1; \beta) + \frac{\beta (M - \alpha - \beta - 1)}{(\alpha + \beta)(\alpha + \beta + 1)} .\]

Now using the transformations (3.9a), (3.9b) and (3.9c) we can rewrite (3.39) as follows:

\[(3.40)\]

\[\frac{M (1 - t) - 1}{M t} \left( - z \sqrt{M k (1 - k)} \right) \left( - \frac{2z \sqrt{M k (1 - k)}}{M k \sqrt{k} + z \sqrt{M k (1 - k)}} \right) \]

\[= \sqrt{M k (1 - k)} B \left( z + \delta \frac{M^{1/2}}{t + M^{-1}} \right) + \frac{M (1 - t) - 1}{M t} \left( \frac{M (1 - t) - 1}{M t} \right) \frac{M (1 - k)}{M t (M t + 1)} \]

Dividing by \(\sqrt{M k (1 - k)}\) throughout, we obtain

\[(3.41)\]

\[- \frac{1 - t}{t} z + O(M^{-1}) = B \left( z + \delta \frac{M^{1/2}}{t + M^{-1}} \right) + \frac{1 - t}{t} \frac{(1 - k)}{\sqrt{M k (1 - k)}} \]

Expanding \(B \left( z + \delta \frac{M^{1/2}}{t + M^{-1}} \right)\) in a Taylor series, and using the fact that \(B(z; t) = \frac{1 - t}{t} \cdot z\) on the optimal boundary in the \((z, t)\)-plane for \(z < 0\) (see 3.26), we obtain from (3.41)
\[
(3.42) \quad - \frac{1 - t}{t} z + O(M^{-1}) = - \frac{1 - t}{t} z + 8 \frac{1}{2} M^{-\frac{1}{2}} B_z + \frac{1 - t}{t} \frac{(1 - k)}{\sqrt{Mk}} (l-k) 
\]

From the definition of \( \delta_1 \) in (3.13), it now follows that

\[
(3.43) \quad B_z(z; t) = - \frac{1 - t}{t} + O(M^{-\frac{1}{2}}). 
\]

Therefore as \( M \to \infty \), we obtain

\[
(3.44) \quad B_z(z; t) = - \frac{1 - t}{t} \text{ on the boundary } (\tilde{z}(t), t). 
\]

To summarize now, we have that in the limit as \( M \to \infty \),

\( B(z; t) \)

satisfies the following partial differential equation

\[
(3.45) \quad \frac{1}{2} B_{zz} + \frac{z}{t} B_z + B_t + \frac{z}{t} = 0 
\]

in the continuation region in the \((z, t)\) plane, i.e., in the set consisting of all points \((z, t)\) such that \([\tilde{z}^-(t) \leq z \leq \tilde{z}^+(t); 0 \leq t \leq 1]\) subject to the following boundary conditions:

\[
B(z; t) = \frac{1 - t}{t} z \quad \text{for all points } (\tilde{z}^+(t), t), \\
B_z(z; t) = \frac{1 - t}{t} \quad \text{on the optimal upper boundary}, \\
\]

\[
B(z; t) = - \frac{1 - t}{t} z \quad \text{for all points } (\tilde{z}^-(t), t), \\
B_z(z; t) = - \frac{1 - t}{t} \quad \text{on the optimal lower boundary}. 
\]

The problem of finding the optimal boundary and the Bayes risk for large lots thus reduces to solving the free boundary problem given by (3.45) and (3.46).
For the sake of comparisons with similar results obtained by Chernoff [11], we introduce the following transformation. Let

\[(3.47) \quad \tilde{B}(z; t) = B(z; t) + z.\]

Then it is easy to see that as \( M \to \infty \), \( \tilde{B}(z; t) \) satisfies the homogeneous partial differential equation

\[(3.48) \quad \frac{1}{2} \tilde{B}_{zz} + \frac{z}{t} \tilde{B}_z + \tilde{B}_t = 0 \]

in the continuation region in the \((z,t)\)-plane given by \( \{(z,t); 0 \leq t \leq 1 \text{ and } \tilde{z}^- (t) \leq z \leq \tilde{z}^+(t)\} \), subject to the boundary conditions

\[(3.49) \quad \tilde{B}(z; t) = \frac{1}{t} z + z \quad \text{on the optimal boundary } (\tilde{z}^+(t); t).\]

\[\tilde{B}_z(z; t) = \frac{1}{t} \]

\[\tilde{B}(z; t) = -\frac{1}{t} z + z \quad \text{on the optimal boundary } (\tilde{z}^-(t); t).\]

\[(3.50) \quad -(2 - \frac{1}{t}) z \quad \text{on the optimal boundary } (\tilde{z}^-(t); t).\]

\[\tilde{B}_z(z; t) = 2 - \frac{1}{t}.\]

It is easily verified that \( \frac{z}{t} \) is a solution to the partial differential equation \((3.48)\) so that

\[(3.51) \quad \tilde{B}(z; t) = \frac{z}{t} - \tilde{B}(z; t)\]

will also satisfy \((3.48)\). For reasons which will become apparent later, we choose to express the free boundary problem which is obtained here in terms of the function \( \tilde{B}(z; t) \). We thus obtain the following transformed free boundary problem:
To solve for the unknown function $\tilde{B}(z; t)$ and the unknown boundary $\tilde{z}^+(t)$ and $\tilde{z}^-(t)$ [0 ≤ $t$ ≤ 1] given that $\tilde{B}(z; t)$ satisfies the partial differential equation

\[
(3.52) \quad \frac{1}{z} \tilde{B}_{zz} + \frac{z}{t} \tilde{B}_z + \tilde{B}_t = 0
\]

in the region

\[
(3.53) \quad \tilde{z}^-(t) \leq z \leq \tilde{z}^+(t); \quad 0 \leq t \leq 1,
\]

and such that $\tilde{B}(z; t)$ satisfies the boundary conditions

\[
(3.54) \quad \tilde{B}(z; t) \bigg|_{\tilde{z}^+(t)} = 0; \quad 0 \leq t \leq 1
\]

\[
(3.55) \quad \tilde{B}_z(z; t) \bigg|_{\tilde{z}^+(t)} = 0; \quad 0 \leq t \leq 1
\]

\[
(3.56) \quad \tilde{B}(z; t) \bigg|_{\tilde{z}^-(t)} = -2 \left(1 - \frac{1}{t}\right) z; \quad 0 \leq t \leq 1
\]

\[
(3.57) \quad \tilde{B}_z(z; t) \bigg|_{\tilde{z}^-(t)} = -2 \left(1 - \frac{1}{t}\right); \quad 0 \leq t \leq 1.
\]

Any free boundary problem of this nature is very difficult to solve and no standard techniques seem to be available with the help of which an exact or an approximate solution in a closed form can be given. As we have noted earlier, formal asymptotic series expansions have been given as solutions to certain free boundary problems of this nature. Nevertheless, it appears that a very good intuition is required to be able to guess what the expansions will look like before
it is possible to provide some solution to this class of FBPs. For
certain types of melting problems in the subject of heat transfer,
one also comes across free boundary problems which are, however,
somewhat different. Techniques for solving the latter type of
problems are discussed in Friedman [19].

It is, however, possible to give a solution to the free
boundary problem described in (3.52) to (3.57) with the help of
certain formal reasoning, and the results given in Chernoff and
Ray's [13] paper. The first step in deriving such a solution is
given in the following

**Theorem 3.1.**

The free boundary problem described in (3.52) to (3.57) is
exact for the determination of the optimal boundary and the Bayes
risk for the following sequential decision problem truncated at
t = 1, concerning the unknown drift µ of a Wiener process Z(t)
with unit standard deviation per unit time.

The elements of the decision problem are as follows:

1. **Action Space:** There are two decisions \( d_1 \) and \( d_2 \).

2. **Terminal Loss Function:**

\[
L(\mu; d_1) = 0
\]

\[
(3.58)
L(\mu; d_2) = -2(T\mu - \mu) = 2(\mu - T\mu),
\]

where \( T \) denotes the stopping time and lies between 0 and 1 (T
is a random variable which denotes the point of time when a given
sequential procedure ceases to observe \( Z(t) \).)
(3) Cost rate per unit time of observing the process:

(3.59) Cost rate \( c(z, t) = 0 \) for \( 0 \leq t \leq 1 \).

(4) Apriori distribution:

Apriori distribution of \( \mu \) is Normal with mean \( \mu_0 \) and variance \( \sigma_0^2 \).

The proof of this theorem follows rather trivially from equations (5.1), (5.2) and (6.7) of Chernoff [11], but for the sake of completeness we choose to give a detailed proof.

In what follows, we shall sometimes represent by \( N(a, b) \) the Normal distribution with mean \( a \) and variance \( b \).

For proving the above theorem, the following lemmas are needed.

**Lemma 3.1**

For a Wiener Process \( Z(s) \) with drift \( \mu \) and standard deviation \( \sigma \) per unit time, the apriori distribution of \( \mu \) is given to be \( N(\mu_0, \sigma_0^2) \). Then the a posteriori distribution of \( \mu \) after observing \( Z(s) \) on \([0, t]\) and given that the sufficient statistic \( Z(t) = z \), is

\[
N\left( \frac{z + \mu_0}{\sigma_0^2}, \frac{1}{t + \frac{1}{\sigma_0^2}} \right).
\]  

**Proof**

This result appears to be widely known and we omit the proof.

The computations for proving the lemma are straightforward.
Now let
\[ t_0 = \frac{1}{\sigma_0^2} \quad \text{and} \quad Z(t_0) = \frac{\mu_0}{\sigma_0^2}. \]

**Corollary 3.1.**

For any time \( t \geq t_0 \), the posteriori distribution of \( \mu \) conditional upon the observed values \([Z(s); t_0 \leq s \leq t]\) and \( Z(t_0) = z_0; Z(t) = z \) is \( N(z/t, 1/t) \).

By means of the transformation (3.61), it is possible to denote the optimal sequential procedures for our problem with respect to any aprriori Normal distribution by a single set of stopping and continuation regions in the \((z, t)\)-plane. If we set up a correspondence between the sequential sampling plan and a random walk in the \((z, t)\)-plane with two absorbing barriers, the situation can be further envisaged as follows: a random walk starts at a point \((z_0, t_0)\) and stops when \( t = 1 \) or earlier if it hits either of the two boundaries. We can also equivalently assume a dummy observation of the process starting from the point \( t = 0 \) up to a time \( t = t_0 \) giving rise to the value of the process \( Z(t_0) = z_0 \). For each aprriori Normal distribution, the starting point of the process will differ but the optimal boundaries will remain the same.

Let \( \tilde{B}(z; t) \) denote the conditional Bayes risk for the above decision problem given that the process has been observed up to time \( t \) and that the value of the process at this point \( Z(t) = z \); it is evident that this function is defined for points within the continuation set only.
We assume that the conditional Bayes risk $\bar{B}(z; t)$ is continuous everywhere and sufficiently differentiable within the continuation set.

Lemma 3.2

$\bar{B}(z; t)$ satisfies the differential equation (3.52) for all points $(z, t)$ within the continuation set.

Proof

Consider $\bar{B}(z; t)$ at any point in the continuation region and imagine that sampling is continued for a short period of length $\delta t$. It can be proved that the probability that $|Z(s) - z| > a$ for some $s$ in $(t, t + \delta t)$ is $o(\delta t)$. Thus we can write

\begin{equation}
\bar{B}(z; t) = c(z; t) \delta t + E[\bar{B}(z + \delta z; t + \delta t) | Z(t) = z] + o(\delta t)
\end{equation}

where

\[ \delta z = Z(t + \delta t) - Z(t). \]

Also, $\delta z$ is distributed $N(\mu \delta t, \delta t)$ and is independent of $Z(t)$.

Therefore,

\begin{equation}
\delta z = \mu \delta t + \sqrt{\delta t} \eta_1,
\end{equation}

where $\eta_1$ is distributed as $N(0, 1)$.

Further, $\mu$ is distributed $N(z/t, 1/t)$

\begin{equation}
\delta z = \left( \frac{z}{t} + \frac{1}{\sqrt{t}} \eta_2 \right) \delta t + \sqrt{\delta t} \eta_1,
\end{equation}

where $\eta_2$ is also distributed as $N(0, 1)$ and is independent of $\eta_1$.

We know that if $X_1$ and $X_2$ are independent random variables each having a Normal distribution with mean 0 and variance 1, then $c_1 X_1 + c_2 X_2$ has a Normal distribution with mean 0 and variance $c_1^2 + c_2^2$. 

We therefore have

\[(3.65) \quad \delta z = \frac{2}{t} \delta t + \left[ \delta t \left( 1 + \frac{\delta t}{t} \right) \right]^{\frac{1}{2}} \eta \cdot \]

Thus from (3.60), and (3.65) we can write (3.62) as

\[(3.66) \quad B(z; t) = E \left( B(z + \frac{2}{t} \delta t + \left[ \delta t \left( 1 + \frac{\delta t}{t} \right) \right]^{\frac{1}{2}} \eta; t + \delta t \right) + o(\delta t). \]

Expanding the right hand side of (3.66) in a Taylor series, we obtain

\[(3.67) \quad B(z; t) = B(z; t) + \delta t \bar{B}_t(z; t) + \left( \frac{2}{t} \delta t + \left[ \delta t \left( 1 + \frac{\delta t}{t} \right) \right]^{\frac{1}{2}} \bar{E}(\eta) \right) \bar{B}_z(z; t) + \frac{1}{2} \delta t \bar{B}_{zz}(z; t) \bar{E}(\eta^2) + o(\delta t), \]

where \(E(\eta) = 0\) and \(E(\eta^2) = 1\).

Simplifying we obtain

\[(3.68) \quad \delta t \bar{B}_t(z; t) + \frac{2}{t} \delta t \bar{B}_z(z; t) + \frac{1}{2} \delta t \bar{B}_{zz}(z; t) + o(\delta t) = 0. \]

Thus in the limit as \(\delta t \to 0\), we have

\[\frac{1}{2} \bar{B}_{zz} + \frac{2}{t} \bar{B}_z + \bar{B}_t = 0, \]

and the lemma is proved.

**Lemma 3.3**

\(\bar{B}(z; t)\) satisfies the boundary conditions (3.54) - (3.57).
Proof

The conditions (3.54) and (3.56) follow directly from the assumption that \( \bar{S}(z;t) \) is continuous. From (3.60), we have that \( L(\mu; d_1) = 0 \) and \( L(\mu; d_2) = 2(\mu - T\mu) \). Also from Corollary 3.1 we observe that the a posteriori distribution of \( \mu \) after observing \( Z(s) (t_0 \leq s \leq t) \) and \( Z(t) = z \) is \( N\left( \frac{z}{t}, \frac{1}{t} \right) \).

Let \( \phi(\mu; a, b) \) denote the Normal density with mean \( a \) and variance \( b \). Let \( D(z,t; d_i) \) denote the conditional stopping risk if \( d_i (i = 1, 2) \) is chosen when \( Z(t) = z \), and \( D(z; t) \) denote the minimum of these two quantities. We therefore have

\[
D(z,t; d_1) = 0
\]

\[
D(z,t; d_2) = - \int_{-\infty}^{\infty} 2(t\mu - \mu) \phi(\mu; \frac{z}{t}, \frac{1}{t}) \, d\mu
\]

\[
= - 2 \left( t \frac{z}{t} - \frac{z}{t} \right)
\]

\[
= - 2 \left( 1 - \frac{1}{t} \right) z,
\]

and

\[
D(z;t) = \min \{ 0, -2(1 - \frac{1}{t})z \}.
\]
Thus

\[(3.71)\] \[D(z; t) = 0 \text{ when } z > 0\]

\[D(z; t) = -2 \left(1 - \frac{1}{t} \right) z \text{ when } z < 0\]

The optimal procedure chooses the decision associated with the minimum stopping risk after ceasing to observe the process. Since we have assumed that \(B(z; t)\) is continuous, it follows that at the upper boundary \(B(z; t) = 0\) and at the lower boundary \(B(z; t) = -2 \left(1 - \frac{1}{t} \right) z\). These conditions are observed to be the same as (3.54) and (3.56).

The derivation of (3.55) and (3.57) is slightly more involved and depends on the same heuristic arguments as given in Section 6 of Chernoff [11]. As a matter of fact, his derivation of the extra boundary conditions with regard to the continuity of the partial derivatives of the Bayes risk does not depend on any particular form of the terminal loss function or costs. His arguments are equally applicable in our case without any alterations. We therefore omit the proof.

Now combining the results of Lemma 3.2 and Lemma 3.3, we obtain the proof of Theorem 3.1.

We will now show that for the decision problem described in the statement of Theorem 3.1, the optimum upper boundary lies at infinity and that all points with finite co-ordinates in the positive quadrant of the \((z, t)\)-plane are continuation points.
Lemma 3.4

When \( Z(t) = z \), it follows that \( \tilde{B}(z; t) \leq 0 \) for all \( (z, t) \) with \( z < \infty \) and \( 0 \leq t \leq 1 \).

Proof

From (3.70), we have

\[
D(z; t) = \min \{ 0, -2(1 - \frac{1}{t}) z \},
\]

therefore

\[
(3.72) \quad \tilde{B}(z; t) \leq D(z; t) \leq 0.
\]

Lemma 3.5

All points \( (z, t) \) with finite co-ordinates for which \( z > 0 \) and \( 0 \leq t \leq 1 \) lie within the optimal continuation set.

Proof

Let \( \delta \) be any decision rule which chooses the decision \( d_2 \) with probability \( \lambda_\delta(z_s) \) given that it has stopped observing the process when \( Z(s) = z_s \). Let \( \psi_s(z_s) \) denote the probability density function of the stopping time \( S \) when the rule \( \delta \) is used. Let \( R(\mu; \delta) \) denote the risk associated with \( \delta \) when \( \mu \) is the true parameter. We then have

\[
(3.73) \quad R(\mu; \delta) = \int_0^1 \int_{-\infty}^\infty \psi_s(z_s) \left( [1 - \lambda_\delta(z_s)] L(\mu; d_1) + \lambda_\delta(z_s) L(\mu; d_2) \right) p_{\mu,s}(z_s) \, ds \, dz_s,
\]

where \( p_{\mu,s} \) denotes the probability density function of the chance variable \( Z(s) - Z(0) \).

If \( \mu \) has the prior distribution \( N(\mu_0, \sigma^2_0) \), then the expected value of \( R(\mu; \delta) \) taken with respect to this prior distribution is given by
(3.74) \[ R(\xi; \delta) \equiv \int_{-\infty}^{\infty} R(\mu; \delta) \phi(\mu; \mu_0, \sigma_0^2) \, d\mu \]

\[ = \int_{-\infty}^{\infty} \int_{0}^{1} \int_{-\infty}^{\infty} \psi_s(z_s) \left[ 1 - \lambda_s(z_s) \right] L(\mu; d_{\perp}) \]

\[ + \lambda_s(z_s) L(\mu; d_{\perp}) \mathcal{P}_{\mu,s}(z_s) \phi(\mu; \mu_0, \sigma_0^2) \, d\mu \, ds \, dz_s. \]

From Lemma 3.1, we therefore have

(3.75) \[ R(\xi; \delta) = \int_{-\infty}^{\infty} \int_{0}^{1} \int_{-\infty}^{\infty} \psi_s(z_s) \left[ 1 - \lambda_s(z_s) \right] L(\mu; d_{\perp}) \]

\[ + \lambda_s(z_s) L(\mu; d_{\perp}) \phi\left(\mu; \frac{z_s + \frac{\mu_0}{\sigma_0^2}}{s + \frac{1}{\sigma_0^2}}, \frac{1}{s + \frac{1}{\sigma_0^2}}\right) \mathcal{P}_{\mu_0, s^2}(z_s) \, d\mu \, ds \, dz_s. \]

where

\[ \mathcal{P}_{\mu_0, s^2}(z_s) = \int_{-\infty}^{\infty} \mathcal{P}_{\mu,s}(z_s) \phi(\mu; \mu_0, \sigma_0^2) \, d\mu. \]

Now let

(3.76a) \[ z_0 = \frac{\mu_0}{\sigma_0^2} \]

(3.76b) \[ t_0 = \frac{1}{\sigma_0^2} \]

(3.76c) \[ t = s + \frac{1}{\sigma_0^2} \]

(3.76d) \[ z_t = z_s + z_0 \]

(3.76e) \[ \psi_t'(z_t') = \psi_s(z_s) \]

(3.76f) \[ \lambda_t'(z_t') = \lambda_s(z_s). \]
Therefore we have

\[ R (\xi; \delta) = \int_{-\infty}^{\infty} \int_{t_0}^{1} \int_{-\infty}^{\infty} \psi_t' (z_{t'}) \left( [1 - \lambda_t' (z_{t'})] L (\mu; d_1) \right. \]
\[ + \lambda_t' (z_{t'}) L (\mu; d_2) \phi (\mu; \frac{z_{t'}}{t}, \frac{1}{t}) p_{\mu_0}, \sigma_0^2 (z_{t'} - z_0) \mu \, dt \, dz_{t'} \.
\]

Now changing the order of integration, we obtain

\[ R (\xi; \delta) = \int_{t_0}^{1} \int_{-\infty}^{\infty} \psi_t' (z_{t'}) \left( [1 - \lambda_t' (z_{t'})] 0 \right. \]
\[ + \lambda_t' (z_{t'}) \left[ - 2 (1 - \frac{1}{t}) z_{t'} \right] p_{\mu_0}, \sigma_0^2 (z_{t'} - z_0) \mu \, dt \, dz_{t'} \]
\[ \geq \int_{t_0}^{1} \int_{-\infty}^{\infty} \psi_t' (z_{t'}) \left( \min [0, - 2 (1 - \frac{1}{t}) z_{t'}] \right) p_{\mu_0}, \sigma_0^2 (z_{t'} - z_0) \mu \, dt \, dz_{t'} \]
\[ = \int_{t_0}^{1} \left( \int_{-\infty}^{\infty} \psi_t' (z_{t'}) \right. \left. p_{\mu_0}, \sigma_0^2 (z_{t'} - z_0) \right) \mu \, dz_{t'} \]
\[ + \int_{-\infty}^{0} \psi_t' (z_{t'}) \left[ - 2 (1 - \frac{1}{t}) z_{t'} \right] p_{\mu_0}, \sigma_0^2 (z_{t'} - z_0) \, dz_{t'} \] dt.

Let

\[ (3.79) \int_{t_0}^{1} \int_{0}^{\infty} \psi_t' (z_{t'}) \, dz_{t'} = \psi_0' \]

Thus \( \psi_0' \) denotes the probability of stopping in the positive quadrant.
Now let $\delta_0$ be a decision procedure which has the same sampling rule as $\delta$ but which chooses the decision associated with the minimum stopping risk, i.e.,

\[(3.80) \quad R(\xi; A) \geq R(\xi; \delta_0)\]

\[= \int_{t_0}^1 \left( \int_{t_0}^0 \psi_t'(z_t') \left[ - 2 \left( 1 - \frac{1}{t} \right) z_t' \right] \mu_0' \circ \sigma_t^2 \right) dt.\]

We shall now show that $\delta_0$ can never be the optimal procedure if $\psi_0' > 0$.

Assume that $\delta_0$ is the optimal procedure and $\psi_0' = \varepsilon > 0$. Then changing $\psi_0$ to 0 and increasing $\psi_t'(z_t')$ for any particular $z_t$ by $\varepsilon$, we can always obtain another procedure which has a lower average risk than $\delta_0$. We thus obtain a contradiction. We therefore conclude that for the optimal procedure, $\psi_0'$ has to be necessarily equal to 0.

The proof of the lemma is now complete.

It is thus seen that for the decision problem described in the statement of Theorem 3.1., the upper boundary lies at infinity and all points in the $(z, t)$-plane for which $z > 0$ and $0 \leq t < 1$ are continuation points. This is equivalent to stating that, for the free boundary problem described in (3.52) to (3.57), the conditions (3.54) and (3.55) can be ignored. We therefore arrive at essentially the same free boundary problem considered by Chernoff and Ray [13], for which solutions in the form of asymptotic power series expansions have been given in Sections 4 and 5 of their paper.
Remark

The above results which we derived from arguments based on statistical decision theory probably could also be obtained from the theory of partial differential equations. A rigorous derivation of Lemma 3.3 by purely analytical techniques, however, appears to be very complicated, and we shall not attempt to do this here. Presumably the techniques used by Aronszajn [3] in connection with the Cauchy problem involving a homogeneous elliptic equation can be applied with some modification to our case as well.

We now give some heuristic arguments which are not based on any statistical theory to indicate why Lemma 3.3 should be true.

From equation (3.54) we have

\[ B(\tilde{z}^+(t); t) = 0. \]

Differentiating the left and right hand sides of (3.54) with respect to \( t \) along \( \tilde{z}^+(t) \) and assuming that \( \tilde{z}^+(t) \) has finite slope, we obtain

\[ \frac{\partial B(\tilde{z}^+(t); t)}{\partial z} \frac{d \tilde{z}^+(t)}{dt} + \frac{\partial B(\tilde{z}^+(t); t)}{\partial t} = 0. \]  

(3.81)

Since from (3.55), \( \vec{B}_z = 0 \) on \( \tilde{z}^+(t) \), we obtain from (3.81) that

\[ \vec{B}_t = 0 \]  

on the upper boundary \( \tilde{z}^+(t) \).
Substituting (3.52) and (3.55) in (3.52), we now get

\[(3.83) \quad \overline{\mathbb{b}}_{zz} = 0 \text{ on the upper boundary } \hat{z}^+(t).\]

Next differentiating \( \overline{\mathbb{b}}_z = 0 \) along the upper boundary with respect to \( t \), we obtain

\[(3.84) \quad \overline{\mathbb{b}}_{zz} \frac{d\hat{z}^+(t)}{dt} + \overline{\mathbb{b}}_{zt} = 0 \text{ on } \hat{z}^+(t).\]

From (3.83) we see that \( \overline{\mathbb{b}}_{zz} = 0 \) on \( \hat{z}^+(t) \); hence we conclude that

\[(3.85) \quad \overline{\mathbb{b}}_{zt} = 0 \text{ on } \hat{z}^+(t).\]

Next differentiating the diffusion equation (3.52) (which we note is also satisfied on the boundary) along \( \hat{z}^+(t) \) with respect to \( z \), we obtain

\[\frac{1}{3} \overline{\mathbb{b}}_{zzz} (\hat{z}^+(t); t) + \frac{1}{t} \overline{\mathbb{b}}_z + \frac{3}{t} \overline{\mathbb{b}}_{zz} + \overline{\mathbb{b}}_{zt} = 0\]

which implies that

\[(3.86) \quad \overline{\mathbb{b}}_{zzz} = 0 \text{ on } \hat{z}^+(t).\]

In this way, if \( \overline{\mathbb{b}} \) is assumed to be real analytic, we can prove by induction that partial derivatives of all orders of \( \overline{\mathbb{b}} (z; t) \) are zero. Then assuming that the boundary curve is analytic, for any point \((z', t')\) within the continuation region, we can expand \( \overline{\mathbb{b}} (z', t) \) about \( (\hat{z}^+(t); t) \) in a Taylor series and show that \( \overline{\mathbb{b}} (z', t') = 0 \) for all points \((z', t')\) inside the continuation region. This result will contradict the other set of boundary conditions (3.56) and (3.57), and it is intuitively clear that for the boundary conditions to be consistent, the upper boundary should be at infinity in which case condition (3.55) can be disregarded.
The results which we obtain from Lemma 3.3 are really very strange inasmuch that starting from a sequential decision problem with two stopping regions and two optimal boundaries, we obtain a free boundary problem which leads to a single optimal boundary. The question therefore naturally arises why we are getting such curious results. Some of the assumptions of Chernoff and Ray, e.g., sufficient differentiability of the function $\tilde{B}(z; t)$ which we have adapted for our problem cannot be easily justified. Nevertheless, it does not appear that the failure of this assumption to hold good is the reason for obtaining such curious conclusions. As a matter of fact, such regularity assumptions are frequently made in many branches of applied mathematics and physics without getting such strange answers. The arguments with the help of which we obtain the extra boundary conditions (3.55) and (3.57), although not sufficiently rigorous, appear to be reasonable. Then, why are we getting such curious results? In the next few paragraphs we give what we believe to be a plausible explanation for obtaining such answers.

At the end of Chapter II, we included some diagrams showing the optimal stopping regions for certain different values of $M$ and $k$. A casual glance at these graphs is enough to indicate that for each set of values for these parameters, the two boundaries are extremely asymmetric with respect to the neutral line. While the upper boundary is more or less a straight line and is relatively close to the neutral line, the lower boundary is a convex curve with a high curvature and is much farther away from the neutral line.
From equation (3.9a) we observe, that the z co-ordinate of any point \((z', t')\) on the boundary in the transformed \((z, t)\)-plane measures the normalized distance of the corresponding point \((\alpha', \beta')\) in the \((\alpha, \beta)\)-plane from the neutral boundary along the line \(\alpha + \beta = \alpha' + \beta'\). We have just remarked in the previous paragraph that for our problem the lower boundary in the \((\alpha, \beta)\)-plane appears to be at a much greater distance from the neutral line as compared to the upper boundary. Therefore it is possible that as \(M\) becomes large, the distance of the lower boundary from the neutral line increases at a much faster rate as compared to the corresponding distance of the upper boundary from the same line. Thus even after applying normalization (3.9a) and (3.9b), we do not get finite co-ordinates for points on the upper boundary although it is possible to reduce the lower boundary to non-degenerate limits by the use of the above set of transformations. This appears to be the reason why we get \(z^+(t) = \infty\) in the transformed free boundary problem described in (3.52) - (3.57).

Existing literature in sequential analysis describing such free boundary problems which are obtained in the limit in large samples for discrete distributions have been confined entirely to the case where the two boundaries are symmetric or where there is only a single boundary. The situation described above, which we suspect may occur frequently in extreme asymmetric cases, has not apparently been encountered previously.

It seems, therefore, that in many asymmetric situations, it will not be possible to obtain sensible asymptotic results by the
application of Chernoff and Ray's techniques. This observation, although perhaps a little unfortunate, does not appear to be very surprising in view of the arguments given in the previous paragraphs.

We mentioned earlier that the free boundary problem stated in (3.52) - (3.57) turns out to be essentially the same as that considered by Chernoff and Ray. In this connection, we point out another interesting fact. The value of conditional Bayes risks for different points in the \((\alpha, \beta)\)-plane for a given value of \(k\) for our two-action problem is the same as those for the one-action problem considered by Chernoff and Ray [13]. We can therefore utilize the series expansions given by them in order to compare the theoretical values of \(\bar{B}(z; t)\) obtained in the limit and the actual \(\bar{B}(z; t)\) we get by considering a very large \(M\). In the following table, we make such a comparison for the case when \(M = 1700\), \(K = .02\) for points on the neutral line, i.e., for those values for which \(z = 0\). For simplicity, we confine our consideration to only those values for which \(t\) is sufficiently close to 1. From Section 4 of Chernoff and Ray [13], we have that for \(t\) sufficiently near 1 and for \(z = 0\), the asymptotic expansion for \(\bar{B}(0; t)\) is given as follows:

\[
\frac{1}{2} \bar{B}(0; t) = B(0; t) = a_3 \left( \frac{1 - t}{t} \right)^{3/2} \frac{1}{3} \varphi(0) + a_5 \left( \frac{1 - t}{t} \right)^{5/2} \frac{1}{15} \varphi(0)
\]

\[
+ a_7 \left( \frac{1 - t}{t} \right)^{7/2} \frac{1}{105} \varphi(0) + \ldots.
\]
where

\[ \phi(0) = \frac{1}{\sqrt{2 \pi}} = 0.3989423 \]

\[ a_3 = -1.6029 \]
\[ a_5 = 5.9278 \]
\[ a_7 = 3.4094. \]

The values obtained by a direct substitution in the above formula have to be multiplied by 2 in order to obtain \( \bar{B}(z; t) \) because our boundary values differ from those of Chernoff and Ray [13] by a factor of 2.

**TABLE 3.1**

Comparison of Normalized Actual Bayes Risk with those Obtained in the Limit as a Solution to the FBP for \( z = 0 \) and \( t \) near 1

\( M = 1700; \ k = 0.02 \)

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>( \beta )</th>
<th>( \rho^*(\alpha; \beta) )</th>
<th>( t )</th>
<th>( \bar{B}(0; t) ) (Actual)</th>
<th>Limiting value of ( \bar{B}(0; t) ) (FBP solution)</th>
</tr>
</thead>
<tbody>
<tr>
<td>33</td>
<td>1617</td>
<td>0.99518</td>
<td>0.97058</td>
<td>-0.001669</td>
<td>-0.002198</td>
</tr>
<tr>
<td>32</td>
<td>1598</td>
<td>1.98480</td>
<td>0.94117</td>
<td>-0.005265</td>
<td>-0.006353</td>
</tr>
<tr>
<td>31</td>
<td>1519</td>
<td>2.97033</td>
<td>0.91176</td>
<td>-0.01028</td>
<td>-0.01194</td>
</tr>
<tr>
<td>30</td>
<td>1470</td>
<td>3.95216</td>
<td>0.88235</td>
<td>-0.01658</td>
<td>-0.01870</td>
</tr>
<tr>
<td>29</td>
<td>1421</td>
<td>4.93041</td>
<td>0.85294</td>
<td>-0.02411</td>
<td>-0.02662</td>
</tr>
<tr>
<td>28</td>
<td>1372</td>
<td>5.90507</td>
<td>0.82353</td>
<td>-0.03289</td>
<td>-0.03558</td>
</tr>
</tbody>
</table>
The actual and the limiting values obtained as solutions to the free boundary problem do not agree too closely. However, in any such comparison, we have to remember that the limiting values of $\tilde{\bar{B}}(0; t)$ given here are only approximate having been computed from the asymptotic expansions given by Chernoff and Ray [13]. Also, we do not know how rapidly the transformed Bayes risks approach the given limiting values as $M$ tends to infinity.


