1970

Precision requirements for some information in timber management decisions

David Alexander Hamilton Jr.

Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Agriculture Commons, Animal Sciences Commons, Natural Resources and Conservation Commons, and the Natural Resources Management and Policy Commons

Recommended Citation

Hamilton, David Alexander Jr., "Precision requirements for some information in timber management decisions " (1970). Retrospective Theses and Dissertations. 4841.

https://lib.dr.iastate.edu/rtd/4841

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
HAMILTON, Jr., David Alexander, 1942-
PRECISION REQUIREMENTS FOR SOME INFORMATION
IN TIMBER MANAGEMENT DECISIONS.
Iowa State University, Ph.D., 1970
Agriculture, forestry & wildlife

University Microfilms, A XEROX Company, Ann Arbor, Michigan
PRECISION REQUIREMENTS FOR SOME INFORMATION
IN TIMBER MANAGEMENT DECISIONS

by

David Alexander Hamilton, Jr.

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major Subject: Forestry

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Head of Major Department

Signature was redacted for privacy.

Dean of Graduate College

Iowa State University
Of Science and Technology
Ames, Iowa

1970
PLEASE NOTE:

Some pages have indistinct print. Filmed as received.

UNIVERSITY MICROFILMS.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER 1. INTRODUCTION</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Previous Work</td>
<td>1</td>
</tr>
<tr>
<td>Scope and Objectives of This Work</td>
<td>6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 2. METHODS OF ANALYSIS</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum Utility</td>
<td>10</td>
</tr>
<tr>
<td>Marginal Utility</td>
<td>17</td>
</tr>
<tr>
<td>Inventory Control</td>
<td>19</td>
</tr>
<tr>
<td>Minimum Cost Plus Loss</td>
<td>25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 3. LOSS FUNCTIONS</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Development and Justification</td>
<td>30</td>
</tr>
<tr>
<td>Analytical Solutions</td>
<td>35</td>
</tr>
<tr>
<td>Bidding Strategy</td>
<td>40</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 4. TIMBER SALE APPRAISAL PROBLEM</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision Rules</td>
<td>48</td>
</tr>
<tr>
<td>Population Description</td>
<td>50</td>
</tr>
<tr>
<td>Procedures</td>
<td>51</td>
</tr>
<tr>
<td>Results and Discussion</td>
<td>64</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 5. ALLOWABLE CUT DETERMINATION PROBLEM</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision Rules</td>
<td>81</td>
</tr>
<tr>
<td>Population Description</td>
<td>85</td>
</tr>
<tr>
<td>Procedures</td>
<td>86</td>
</tr>
<tr>
<td>Results and Discussion</td>
<td>97</td>
</tr>
</tbody>
</table>
CHAPTER 6. SUMMARY  102

BIBLIOGRAPHY  111

ACKNOWLEDGMENT  114
CHAPTER 1. INTRODUCTION

Sampling is extremely important today in forestry in the United States. Almost all management decisions are based on information that either directly or indirectly comes from sampling. Whether the decision deals strictly with timber or with several forest uses combined in the multiple use concept, some form of sample survey very likely provides at least a portion of the information necessary for decision making.

In many of these decision making situations, the forest manager (i.e., decision maker) arbitrarily selects the level of precision to be required for any given sample on the basis of intuition or past experience. Avery (1967) stated that the precision required in a timber inventory is logically based on the planned use of the information to be collected. He then cited, as an example, that for land acquisition or timber sale purposes in areas where timber values are relatively high, the sample should produce estimates of mean value per acre with standard errors of ten percent or less.

Spurr (1952) went to some length to indicate that sample surveys that include as little as five to ten percent of the population should be expected to provide adequate results in many forestry situations. Several forestry examples were cited to indicate the sampling fraction necessary to obtain specified standard errors. For these examples the specified standard errors fell mostly between five and twenty percent of the value to be estimated by the sample survey (often total volume).

Such examples and rules of thumb may be very useful to the forest manager. However, some logical justification for selecting the precision
of information necessary for a given management decision would be of more use and importance and would, by necessity, precede the determination of the sample size required to assure compliance with this precision requirement.

It would be more reasonable to determine sample size on the basis of expected loss resulting from use of sample survey information. Most rules-of-thumb for specifying sample size undoubtedly are based to some extent on this concept. But each management situation and decision rule is different enough that no general rule-of-thumb will likely prove completely adequate.

Spurr's statement about sampling fractions of five to ten percent often being adequate in forestry situations demonstrates one of the commonly used rules-of-thumb to set precision levels. In this approach to setting precision levels, one plans to include a given percentage of the population in the sample survey. This percent sample is selected because it has been used in the past. The precision "desired" is set as the precision that can be obtained with the selected sampling fraction.

As one example of a rule-of-thumb used to specify precision levels, there is the relationship reported by Barton (1960). This relationship specifies the sample size needed for the continuous forest inventory system as a function of the forest area involved. This relationship was developed by observing the sample sizes and areas for continuous forest inventory installations already established for forests of $10^2$ to $3.5 \times 10^6$ acres. The resulting equation expressing sample size as a function of area then is used to specify sample size for any new installation. This specification of precision is thus based not on the specific uses of information or special management objectives for the given decision maker but instead
on the average subjective decisions of past planners of continuous forest inventory installations. Costs, rather than information use, were usually the major determinants of sample size and hence precision.

Stearns, Schweitzer, and Creed (1968) further illustrated the arbitrary or subjective methods by which desired levels of precision are often selected. For work they conducted on deer browse production, they concluded that because of the great variability of vegetation on plots and because of the great variability in the utilization of browse by deer, a reasonable precision for estimation would be a sampling error of about twenty percent of the mean. This puts the emphasis on the nature of the population rather than on the use of the information.

Spurr (1952) also pointed out that the more variable the population, the larger the required sample size, and the more valuable the timber the larger the sampling fraction. This statement appears rational and in accord with what Avery (1967) and Stearns, Schweitzer, and Creed (1968) concluded about precision requirements for a forest sample survey. However, these authors have not attempted to formalize any of these rules-of-thumb into specific decision rules that the individual decision maker might use as a guide where specific uses are planned for the sample information.

In this study I have investigated some of the relationships between forest management decisions, information needs, and forest sample survey objectives. The long run objective is to aid in the development of such formalized decision rules.

The decision maker who has this problem is, in most general terms, any management personnel faced with making decisions based on sample survey information. Ordinarily, the description of the decision maker for a
decision problem should be much more explicit. However, until a specific forest management decision is selected for study, this general description is adequate. The objective of this decision maker is to specify or prescribe that level of precision that is required for the forest sample survey information to be collected. We shall assume that this information is being gathered to help a forest manager to meet his objectives in a management situation of general interest. The alternatives open to our decision maker are the various levels of precision that could be prescribed for any given sample survey. The consequent sample sizes range from complete enumeration (one hundred percent sample) to no sampling at all.

Some examples will serve as both motivation and as possible areas of application for this study. The first of these comes from Hazard's (1969) work on sampling with partial replacement. It was necessary in the preliminary stages of that work for someone to specify the level of precision, in the form of a variance limit, which would be required for each of the five estimates that were to be made jointly. These precision levels were arbitrarily set with the experience and tradition in the U.S. Forest Service's nationwide Forest Survey serving as the principal guideline. The precision requirement for one or two of the estimates dominated in determination of the jointly optimal sample size. In a situation such as this, the question then arises as to whether the more stringent requirements for the dominating estimates are actually necessary to meet the overall objectives of the decision maker.

Forest resource appraisal provides several examples of applications where the study of information needs does not lead to a trivial precision specification procedure. Since the first of the forest management decisions
dealt with in detail is a timber sale appraisal problem, I will delay de­tailed discussion of this application. However, I will point out here that in specifying precision requirements in any timber appraisal situation, the most relevant question to be answered is: To what use will the appraisal information be put? This is a relevant question for any sample survey situation and thus is of major importance in the planning phase of all sample surveys.

A second specific area of application is damage appraisal (e.g., from insect epidemic, fire, weather, disease). Flora (1968) outlined the situation as follows:

"...damage should be appraised in terms that lend themselves to decisions that are actually pending or are anticipated. Some foresters who cast about for "better" appraisal techniques cannot say how they will use the appraisal information when it is obtained. One wonders how costly appraisals are justified when a decision between courses of action is not pending. An economist is left to ponder in such cases just how simple, how precise, how cheap, how understandable, and how definitive a system is really needed."

Large general purpose surveys such as the nationwide Forest Survey conducted by the U.S. Forest Service provide another example of areas of application for the kind of results sought in this study. It is extremely difficult to identify all the potential uses or users of the information provided by the Forest Survey. In fact, some decision makers feel that an important contribution of the Forest Survey is to provide a body of information that may be used to satisfy any unplanned needs for information that may arise in the management of our National Forests. Although these additional difficulties in defining all decision makers and their objectives make it undesirable to work with this Survey in developing procedures for precision specification, it does remain an area to which the results of
this study should be quite relevant. The planners of Forest Survey have considerable leeway and responsibility in specification of total budget, in internal allocation of their budget, and in writing precision specifications.

**Previous Work**

One of the earliest references to work specifying precision requirements and sample size in forestry is by Blythe (1945) who studied the optimum sample size for scaling logs in a timber sale. In this application, where the amount paid for timber is based on the scaled estimate of timber volume, the producer (say the Forest Service or any other large timber selling organization) is satisfied to operate on a long run basis (waiting for sales to "average out") and thus is relatively unconcerned if only a small number of logs is included in the sample scaled, while the purchaser is more concerned with the short run. This is especially true for purchasers of small quantities. The purchaser wants some assurance that there is only a reasonably small probability of his scale estimate of total volume and quality being a large overestimate. Thus the problem boils down to determining how much must be invested in scaling so as to give the purchaser this assurance.

Blythe expressed the tolerable limit of error in monetary units as a function of the coefficient of variation and the sample size. He suggested that a reasonable solution could be obtained by sampling to that point (i.e., increasing sample size to the point) where the marginal reduction in the limit of error expressed in monetary units is equal to the marginal increase in the cost of sampling. This is, in essence, a procedure of specifying that level of sampling that minimizes a cost plus loss function.
Yates (1960) and Cochran (1963) both treated the idea of determining the optimum level of precision (hence, the sample size) by the minimization of a cost plus loss function. According to Yates, it is often impossible to quantitatively place a value on survey information. Thus he concluded that the precision level specified for a survey must depend to a large degree on the judgment of the decision maker who is to use the survey information.

Cochran (1963) also pointed out that subjective judgment frequently must play a major role in the specification of desired precision levels. Two basic problems make a more formal specification of precision difficult. In addition the results of surveys are frequently applied to some use that was not or could not be foreseen at the time the survey was planned. The second problem is that very little is known, in many practical situations, about the consequences (i.e., losses) associated with various sizes of errors in estimates from survey information. The effects these errors have on the actual decisions to be made are frequently difficult to formalize.

Grundy, Healy, and Rees (1956) worked on a closely related problem that arises in many areas of management. When a new process becomes available, it is usually necessary to do some experimentation to determine whether the new process significantly outperforms the established one. These authors suggested a two stage sampling procedure in which sample size (i.e., precision) is optimized at the second stage by minimizing the risk associated with an incorrect decision as to the acceptance or the rejection of the new process.

Kempthorne (1952), Cochran and Cox (1957), and Steel and Torrie (1960) discussed similar difficulties in experimental design. In experi-
mental design the decision maker must determine the level of significance that he should use. This decision should be based on the use to which the experimental results will be put and on the importance of decisions that will be based on these results. As pointed out by Cochran and Cox (1957), the use of five percent and one percent significance levels is merely a useful convention. Lower probability levels should be used if the consequences of an erroneous rejection of the hypothesis are extremely serious. Steel and Torrie (1960) say that the magnitude of a real departure that is of practical significance can be determined only by a researcher with practical knowledge of the subject matter. This magnitude together with some measure of how desirable it is to detect it, serves to determine the required precision and, for example, the number of replicates for the experiment.

Scope and Objectives of This Work

As previously stated, in many decision making situations the required level of precision is often selected by some purely arbitrary method based on tradition or experience. One questions whether these levels of precision, as specified by tradition or experience, are usually most appropriate to satisfy the stated objectives of the decision maker. Therefore, the objectives of this study are:

(1) to investigate the levels of precision required for two fully specified management decisions;

(2) to determine whether one can derive a general approach to be followed in specifying precision requirements for a class of forest management decisions.
Several possible approaches are available for the analysis of this problem. Marginal value theory and minimization of a cost plus loss function have been applied in studies cited earlier. In Chapter 2 I discuss these approaches in detail. Two additional possible techniques for analysis are also discussed there. The difficulties encountered in developing loss functions for typical forestry situations are discussed in Chapter 3. The analysis is discussed in Chapter 4. A second example which utilizes a more sophisticated decision rule is discussed in Chapter 5. This example deals with allowable cut determination and uses, as a decision rule, the one used in SORAC, a computer program prepared by Sassaman, Chappelle, and Fritchman (1969) for determining allowable cut in Douglas-fir forests.

One of the objectives of this study is to attempt to develop a general approach for determining required precision levels in forest management decision making situations. The two management decision situations discussed in Chapters 4 and 5 serve as examples and as guides for the development of such an approach. Generalizations about such an approach are discussed in Chapter 6 which serves as an overall summary of this study.
CHAPTER 2. METHODS OF ANALYSIS

This problem of specifying precision requirements from information uses can be approached in many ways. Many similarities exist among these different approaches to analysis. These similarities as well as distinguishing differences are discussed for each of four specific approaches to the problem. The discussion in this chapter deals with the characteristics of each approach as these characteristics apply to management decisions in general. The analysis procedures applied in Chapters 4 and 5 are described in those chapters with specific reference to the management decisions under study there.

Maximum Utility

One approach to the analysis of the precision specification problem has an objective of specifying that level of precision for a sample survey that will maximize the utility function for the given decision situation. If the decision rule has not been fully specified, this form of analysis may be expanded to have as its objective to specify both the level of precision for sample survey information and the decision rule which will maximize the utility function associated with the given decision making situation.

The first requirement of this approach is to develop a utility function that adequately describes the given decision situation. This utility function should relate the various levels of precision at which the sample survey might be conducted to some measure of utility or relative value that results from running the survey at each respective precision level. As defined by Raiffa and Schlaifer (1961), a utility function, $u(e,z,a,a)$, assigns a measure of utility or value to performing a particular experiment,
e, observing a particular outcome, \( z \), taking a particular action (i.e., arriving at a particular decision), \( a \), and then finding that a particular state of nature, \( \Theta \), actually exists.

This definition needs to be expanded and set in the context of the general forest management situation to be analysed. Here one can view the family or set of experiments \( \{e\} \) from which the decision maker may choose as the set of alternative levels of precision at which the sample survey may be conducted. This set is equivalent to the class of alternatives open to the decision maker described in Chapter 1. The nature of the set of possible outcomes \( \{z\} \) will depend entirely upon the specific management decision situation to be analysed. In many forest sample survey situations, the set \( \{z\} \) is the possible range of volume estimates obtained from the sample survey.

The set of activities \( \{a\} \) represents the set of decisions that might be made based on any particular outcome, \( z \). Where the decision rule is fully specified, the set \( \{a\} \) is rigidly defined (i.e., for any outcome, \( z \), the decision rule specifies the action to be taken) and thus, only one member of \( \{a\} \) is considered for any specific \( z \). Where the decision rule has not been fully defined before analysis, each member of \( \{a\} \) is considered with each member of \( \{z\} \). Once the utility function has been fully evaluated, the decision rule is then defined as that strategy which prescribes taking the action, \( a \), for each outcome, \( z \), which maximizes the expected utility for that outcome, \( z \). Finally \( \{\Theta\} \) represents the set of true states of nature that might exist. In relation to the example used in describing \( \{z\} \), the set \( \{\Theta\} \) would represent the range of true volumes
that could exist for the stand being surveyed. In this case \( \{z\} \) is equivalent to \( \{0\} \).

A utility function of this form is needed for this analysis because of the difficulty of placing a value on the information that is provided by a sample survey. Each time a forest manager makes a decision as to how precise a given sample survey must be he either consciously or subconsciously implies a value for the information he plans to gather from the sample survey. This value assessment often is made by considering the fact that surveys designed for a given level of precision have provided adequate information in the past. Thus the information is assumed to be worth the cost of taking the survey at the given level of precision. This approach appears, at best, only to place an upper bound on the value of information.

To apply the analysis based on maximum utility in this study, we need a more formalized procedure to determine the value of information. Lave (1963) dealt with this kind of problem in his work to determine the value of better weather information to the raisin industry in California. He estimated the expected return both with and without improved weather information. The difference between these two values represented the value of the improved weather information.

Another approach to placing a value on information is based on the idea of avoidance of loss. This technique appears to be closely related to the procedure just described. It is dependent on the decision maker being able to define a loss function that accurately describes the losses incurred by using a specified decision rule. With this loss function the decision maker may evaluate the expected loss that results from basing decisions on any given level of information. This approach to placing a value on infor-
mation requires that the manager change his objective from seeking to maximize expected utility to seeking to minimize expected loss. Raiffa and Schlaifer (1961) have suggested that a procedure similar to this might be more easily applied in some situations than the maximum utility method of analysis.

The notation used in the definition of a utility function seems to imply that the measure of utility is directly dependent upon the observed outcome, \( z \), for its evaluation. In most forestry management decision situations, the observed outcome will affect the measure of utility only through the action that is taken as a result of the outcome. This is especially true for those situations for which the decision rule is completely defined. In these cases any given outcome, \( z \), immediately requires taking that action, \( a \), which is specified by the decision rule. A closer review of the discussion by Raiffa and Schlaifer (1961) indicates that their general notation was developed so that they could deal with the possibility of the dependence of the measure of utility on the observed outcome.

In addition to requiring one to place a value on information, the maximum utility approach requires the development of two measures of probability. The first of these is the probability of observing a specific outcome, \( z \), given the sample survey has been conducted at a given level of precision, \( e \). Standard notation for this measure of probability would be \( P(z|e) \). For many forest management situations, including those considered in Chapters 4 and 5, this measure of probability is readily estimated by using the Central Limit Theorem.

The mean, \( \bar{x}_n \), of a random sample of size \( n \) from an infinite distribution with mean, \( \mu \), and finite variance, \( \sigma^2 \), has a distribution that is approx-
approximately normal with mean, \( \mu \), and variance, \( \sigma^2/n \) (e.g., Mood and Graybill (1953)). Thus for those situations where the outcome, \( z \), is either a mean of the sample observations or a linear function of a mean, the outcomes are distributed approximately normal with the appropriately specified parameters. Once the distribution of outcomes is defined, the evaluation of \( P(z|e) \) is straightforward.

The second measure of probability that must be evaluated is the probability that a specific state of nature, \( \theta \), exists given one has observed a specific outcome, \( z \). In the usual notation this is \( P(\theta|z) \). The most readily apparent approach to the evaluation of this measure of probability is through the use of simulation. This requires the use of a great number of populations. These populations should be representative of a wide range of values of \( \theta \). Ideally there should be at least one population for each member of the set of possible true states of nature \( \{\theta\} \). For each precision level to be included in the set \( \{e\} \), a large number of samples are drawn from each population. The resulting frequency of occurrence for each outcome in each population is recorded in a two-way table designed to tabulate the overall frequency of occurrence of each outcome for each true state of nature. This table of frequencies defines an estimate of \( P(\theta|z) \).

The determination of the maximum utility strategy, as described by Raiffa and Schlaifer (1961), is basically a game theory approach to analysis. It may be looked at as a two-person game between the decision maker and nature. The four moves of this game are:

1. the decision maker selects an \( e \) from \( \{e\} \);

2. nature chooses an outcome from \( \{z\} \) according to the probability measure \( P(z|e) \);
(3) the decision maker selects an action from \( \{a\} \);

(4) nature selects \( \theta \) from \( \{\theta\} \) according to the probability measure

\[ P(\theta|z). \]

The payoff to the decision maker is the utility \( u(e,z,a,\theta) \).

The evaluation of this game results in the complete specification of the strategy which yields maximum utility to the decision maker. For each move by nature, the expected value of the utility is determined, while for each move by the decision maker, the maximum action is determined. In equation form the maximum utility is defined as

\[ u = \max_e E(z|e) \max_a E(\theta|z) u(e,z,a,\theta) \]

As indicated previously, a slight modification of this procedure may prove advantageous in some situations. Raiffa and Schlaifer (1961) referred to this as minimizing opportunity loss in place of maximizing utility. In familiar notation, the opportunity loss function is expressed as

\[ l(e,z,a,\theta) = u(e_0,z_0,a_0,\theta) - u(e,z,a,\theta) \]

\( e_0 \) represents the case of conducting no experiment at all (i.e., taking no survey) but making a terminal decision immediately based on no information from experimentation. Hence \( z_0 \) is the dummy outcome associated with \( e_0 \), and \( a_0 \) is defined as that action that is optimal (i.e., maximizes utility) given \( \theta \) is the true state of nature.

The evaluation of the opportunity loss function is simplified if one can assume that the function can be broken down into additive parts. These two parts are the opportunity loss associated with taking the action \( a \), given \( \theta \) exists as the true state of nature, and the cost associated with conducting experiment \( e \) in order to observe outcome \( z \). Notationally this is

\[ l(e,z,a,\theta) = l_t(a,\theta) + C_s(e,z) \]
where $l_t(a, \theta) = u_t(a, \theta, \theta) - u_t(a, \theta)$ and the cost of experimentation for the null experiment, $e_0$, is clearly zero.

Minimization of this opportunity loss function is readily shown to be equivalent to maximization of the associated utility function. This is true because $u(e_o, z_o, a_o, \theta)$ is a function only of $\theta$ and not of the decision variables $e$ and $a$. Thus $u(e_o, z_o, a_o, \theta)$ is merely a constant, $C$, in terms of the variables to be minimized or maximized. The only additional fact needed is that $\min(x) = -\max(-x)$. The proof is then

$$\min_e (l) = \min_e (C - u)$$

$$= -\max_e (u - C)$$

$$= - (\max_e (u) - C)$$

$$= C - \max_e (u)$$

where $l$ represents the opportunity loss function and $u$ represents the associated utility function.

One final difficulty must be considered in adapting this method of analysis for use in this study. This deals with the nature of the set of outcomes $\{z\}$ and of the set of states of nature $\{\theta\}$. In many forest management situations both the set of outcomes $\{z\}$ and the set of states of nature $\{\theta\}$ are continuous rather than discrete. Much of the theory, however, concerning determination of the maximum utility strategy or the maximum opportunity loss strategy deals with discrete sets of both outcomes and states of nature. For those forest management decisions for which this method of analysis might be used, it is not unreasonable to group both the possible outcomes and the possible states of nature into appropriately sized classes. This allows the decision maker to apply the usual procedures.
of analysis, using the previously defined classes as discrete members of the sets of outcomes and of states of nature.

Marginal Utility

The objective of this method of analysis is to specify that level of precision for which the marginal unit cost of obtaining an extra unit of precision is equal to the expected marginal unit return (utility) derived from the additional unit of precision. Duerr (1960) described the procedures involved in applying marginal analysis to many common forest management decisions.

In this approach to analysis there are two basic data needs. First, one must evaluate the function that describes the marginal cost of each additional unit of precision. This is a rather straightforward process once a definition of precision is completely specified. Here one may define precision by use of the concepts of confidence interval estimation.

The upper limit, $L_u$, of a $(1 - \alpha)100$ percent confidence interval is

$$L_u = \bar{x} + t_{n-1, \alpha} S/\sqrt{n},$$

if one may ignore the finite population correction factor and assume that $\bar{x}$ is normally distributed about the population mean, $\mu$. Here $t_{n-1, \alpha}$ is the tabulated value of the "t" distribution with $n-1$ degrees of freedom at the $\alpha$ probability level, $S$ is the sample-based estimate of the standard deviation of the population of interest, $n$ is the sample size, and $\bar{x}$ is the estimate of the population mean based on the $n$ observations. A measure of precision, $D$, is thus defined as the half-width of the confidence interval expressed as

$$D = t_{n-1, \alpha} S/\sqrt{n}$$

By rearranging the terms of this relationship, the sample size required to attain a specified level of precision, $D$, is determined by the well-known
Thus, as precision increases (i.e., \( D \) decreases), the specified sample size, \( n \), increases quadratically. If one can assume that individual observations are obtained at a uniformly high level of technical precision and accuracy regardless of the level of precision required of the survey estimate, then the marginal cost of obtaining an extra unit of precision is directly proportional to some function of the number of additional sampling units required to attain that extra unit of precision. This function is closely dependent on the nature of the cost function that describes the cost of sampling for the forest management situation being studied.

The second data need for this method of analysis requires one to evaluate the function expressing the marginal unit return resulting from conducting the sample survey with an additional unit of precision. To accomplish this, one must look again at the problem of placing a value on information and thus, one must evaluate the appropriate utility function. The form of utility function to be used here is only slightly different from that described in the discussion of the maximum utility approach to analysis.

For the marginal utility approach to analysis, the utility function will evaluate the utility to be gained from taking a given action, \( a \), given \( \theta \) is the true state of nature. The cost of obtaining the sample survey information at a specified level of precision is not included in this utility function. This cost term is instead included in the marginal unit cost evaluation discussed previously.
For the evaluation of the marginal unit return function, it appears that the use of an idea similar to the opportunity loss concept might prove quite helpful. The actual evaluation of this function is clearly dependent upon a complete specification of the decision rule being used. Once the decision rule is specified, however, evaluation of the utility function, and thus the marginal unit return function, requires only the determination of the form of the loss function that most nearly expresses the losses resulting from the application of this decision rule.

For each level of precision it is necessary to determine the expected loss that results from conducting the sample survey at that level of precision. The differences between the values of expected loss for successive levels of precision indicate the value of the additional unit of precision and thus, describe the marginal unit return function.

Once these two basic data needs have been met, the analysis by this procedure is quite straightforward. That level of precision for which the marginal unit return equals the marginal unit cost of sampling represents the optimum level of precision specified by the marginal utility approach to analysis.

Inventory Control

One might consider applying the ideas of inventory control (for stock inventories) to this forest survey problem. The objective of the inventory control approach to analysis is to determine that combination of initial sample survey precision and time interval between resurveys that minimizes the cost of decision making. Inventory control is defined by Horowitz (1965) as that field of study which concerns regulating the size and composition of an inventory in order to minimize its costs.
This field of study has been developed fairly substantially for some specific types of inventory models. Although most of these models do not closely fit real-world situations, they do serve to illustrate some of the main factors that must be considered in developing an inventory control policy for any actual management situation. These factors include such things as the cost of placing an order, the cost of not having stock available on demand, the cost of storing stock in inventory, and the time lags between the time of placing an order and the time of receiving the same order. Though many other factors might be considered, these are the ones that seem to have application for the forest management decisions considered in this study.

In the application of inventory control as a method of analysis for the precision specification problem, the information gained from the forest sample survey is treated as the property or product that is to be kept in stock. In this light the objective of this method of analysis, which was previously stated as the determination of the answer to the question how often to resurvey a given forest and how precise each survey should be, may be restated in terms relating to inventory control as the evaluation of how often a given stock should be reordered and what quantity of stock should be ordered at each restocking.

The cost of placing an order, usually referred to as the setup cost in inventory control terminology, corresponds to the costs that are met in running the forest sample survey. This includes both sampling costs and the costs resulting from analysis of the sample observations. The concept of demand depleting the stock on hand can be readily compared to the deterioration of the precision of survey information over time. In this same
sense, the cost of not having stock available on demand is comparable to the cost incurred by making an incorrect decision based on survey information that has deteriorated over time. This could also be looked at as the cost of not being willing to make a decision on the basis of available information.

These last two statements indicate the two major data needs that must be fulfilled if this method of analysis is to be applied. The first of these is the determination of a function that expresses the relationship between the precision of information obtained from a forest sample survey and a variable measuring time since that information was collected. Several possible approaches to the derivation of this function are discussed in a possible application of inventory control procedures to the analysis of the precision specification problem that follows this general discussion.

The second data need that must be fulfilled is the evaluation of a loss function that will adequately express the loss that results from making decisions based on information that has decayed in precision over time. The actual evaluation of the loss function is highly dependent on both the forest management situation that is being considered and on the complete specification of the decision rule that is to be applied in this situation.

It is rather difficult to draw a comparison from forest management decision making situations for the cost of storing stock that is held in inventory. However, the lag time for receiving orders in inventory control analysis has a direct comparison in the lag time between the time at which it is decided to resurvey and the time the resurvey information actually becomes available to the decision maker.
In this discussion I have only briefly outlined some of the general factors included in inventory control procedures. I have also indicated how each of these factors might be related to corresponding factors necessary for the analysis of the precision specification problem in a forest management situation. This might best be summarized and clarified by a discussion of the application of a simplified inventory control model to the analysis of the general precision specification problem. The following relationship diagramatically represents the general form of the inventory control model that is to be considered. A general forest management situation that this model might fit is one in which a decision must be made periodically throughout a rotation of length T years. Sample surveys are to be conducted at regular intervals throughout the rotation. Each sample survey is to be conducted at the same level of precision. Thus the parameters which must be estimated are the time interval, t, between sample surveys and the precision level of each sample survey. The decision rule that is applied in this forest management situation specifies that decisions cannot be made on information with a precision level lower than some specified limit, P.

It is assumed that the lag time for this situation is zero. This means that the actual lag time (i.e., time required to obtain new sample survey information) is less than the period of time between decisions. The cost function associated with this model is a two-term function. The first term
is the setup cost. For this model the setup cost is the cost of obtaining sample survey information at a given precision level. Sample size is used in lieu of precision in this analysis. Thus, this first term of the cost function is of the form

\[ Q_1 = C_1 + C_2n \]

where \( Q_1 \) is the setup cost, \( C_1 \) is the fixed setup cost (i.e., those costs not depending on sample size), \( C_2 \) is the cost of observing a single sampling unit, and \( n \) is the sample size (i.e., precision level). This first term of the cost function may be modified by some management situations so that the factor measuring the cost of sampling is something other than a linear function of sample size. However, for illustrative purposes, this linear function is adequate.

The second term of the cost function measures the expected loss resulting from decisions made between sample surveys. The basic data need that must be fulfilled if this term is to be evaluated is the development of a function describing the decay of the precision of information over time. One approach to the development of this function is to consider the information available at any given time as a combination of the sample survey information obtained at the beginning of an interval and a measure of growth or change that is applied to this initial information. The precision of the resulting information is a function of both the precision of the initial sample survey information and the precision of the measure of change.

This appears to be somewhat similar to the procedure in regression theory by which confidence intervals are developed about a regression line, \( \hat{Y} = a + bx \). In the regression situation the confidence interval about the estimated line may be considered to be made up of two components. The first
of these deals only with variation in vertical placement of the estimated line and is based on the standard error of the estimated Y for a given $\bar{x}$.

The second component of the confidence interval deals in part with variation in the estimate of slope, $b$. The first component of the confidence interval may be compared to the precision of the initial sample survey information while the second component appears to be closely related to the precision associated with the measure of change.

A second approach that might be taken in evaluation of the function expressing the decay of information precision deletes the use of a measure of change. The survey information is collected at the beginning of each interval at that precision level specified by analysis. Rather than projecting changes in this information over time, it is assumed that this information will be used, as is, until a new sample survey is conducted. Thus, for this approach the function that describes the decay of the precision of survey information actually describes the effect of making decisions based on sample survey data collected in the past but used as if it were current.

Once this function that describes the decay of the precision of sample survey information is formulated, the actual analysis of the model may be completed. This function allows the decision maker to know the precision of the information available to him at each point in time that a decision is to be made. With this information and an appropriately defined loss function, the expected loss resulting from making a decision is readily evaluated.

The second term of the cost function is merely the sum of the expected losses for all decisions made in the interval, $t$. Thus the size of this term is a function of the length of the interval which in turn is a function of the precision level of the sample survey that is conducted at the begin-
ning of each interval. This second dependence is due to the fact that a lower limit of precision, \( P \), has been specified. Once the initial precision level of the sample survey is specified, the time, \( t \), required for this precision level to decay to the level \( P \) is readily obtainable from the precision decay function.

Thus the total cost associated with a given interval is

\[
C = C_1 + C_2 n + f(n)
\]

where \( f(n) \) is the second term of the cost function measuring expected loss resulting from decision making. In a rotation of \( T \) years, this sequence must be repeated \( T/t \) times. Thus the total cost of decision making over a rotation is

\[
TC = (T/t)(C_1 + C_2 n + f(n))
\]

\[
= (T/g(n))(C_1 + C_2 n + f(n))
\]

where \( g(n) \) is the function expressing \( t \) as a function of the precision level of each sample survey.

The inventory control method of analysis then specifies that the optimum precision level for each sample survey is that precision level which minimizes this total cost function. The interval between surveys is then specified by the function \( t = g(n) \).

**Minimum Cost Plus Loss**

The objective of this method of analysis is to specify that level of precision for sample survey information that minimizes a function that includes both a term expressing the cost of obtaining sample survey information at a given level of precision and a term expressing the expected cost of making a decision based on this information. The actual evaluation of this cost plus loss function depends on two data needs.
The first of these data needs is a function that expresses the cost of sampling as a function of sample size. The nature of this function will depend on both the management decision making situation involved and on the specific population being sampled. The second data need has been fully discussed in association with the previous proposed methods of analysis. This data need is the necessity of determining the nature of the loss function that adequately describes the loss that results from making a decision based on sample survey information collected at any given level of precision. Again the actual determination of this loss function is dependent on a complete specification of the forest management situation being analysed.

This method of analysis may be thought of as a special case of each of the three previously discussed methods. When the opportunity loss concept is applied in the maximum utility method of analysis, two further assumptions make the method almost exactly equivalent to the minimization of a cost plus loss function. It is first necessary to assume that for the management situation of interest there is only one true state of nature, which is known. Since this type of analysis would probably only be used in research where the objective of the research is to evaluate the precision level of information that would be required in similar situations, it is not unreasonable to expect that the population being used would be completely enumerated. Thus the assumption that the true state of nature is known is acceptable.

The second assumption that must be made is that the decision rule is completely specified. The only effect this has on the analysis is that those procedures used to determine the optimal strategy to be followed for each observed outcome may be omitted from the analysis. The minimum oppor-
tunity loss under these assumptions is evaluated by the equation

\[ l = \min_{e} E(z|e) l(e, z, a, e) \]

Since expected opportunity loss for any given precision level in fact measures the expected loss that results from making a decision based on sample survey information collected at that precision level, the precision level that minimizes opportunity loss is equivalent to the precision level that minimizes the associated cost plus loss function.

Although the procedures of analysis appear to be different for the marginal utility approach to analysis and the cost plus loss method, the results of analysis by these two methods are equivalent. The marginal utility specification of precision level is that level of precision for which an increase in precision results in an increased cost of sampling greater than the corresponding decrease in the cost of decision making and a decrease in precision results in an increase in the cost of decision making greater than the corresponding decrease in the cost of sampling. This precision specification is thus that precision level which minimizes the sum of the cost of sampling and the cost of decision making (i.e., expected loss of making a decision based on information collected at a specified level of precision). This is the same level of precision that is obtained when the analysis is done by minimization of a cost plus loss function.

The final comparison to be made is between the cost plus loss method of analysis and the inventory control method. Horowitz (1965) comments that few actual situations will ever fit the specified textbook models for inventory control. Several of the assumptions made in the analysis of these textbook models are made primarily for simplification or clarification of
the presentation and are not intended to be truly representative of any actual situation.

As a result of this lack of applicability, it is suggested that in developing an inventory control system for any specified management situation, the decision maker should develop a model representing the characteristics of that situation. This involves developing a cost function for the system of interest that considers all relevant factors. The optimal strategy for inventory control is then that strategy that minimizes this cost function.

The cost plus loss function described previously fits as a special case in this general discussion of inventory control analysis. In addition the model for the management situation outlined in the discussion of inventory control may be thought of as a form of a cost plus loss function. The first term of the model measured the cost of collecting sample survey information at a specified precision level, and the second term measured the sum of the expected losses resulting from making decisions in the interval between successive sample surveys.

Thus it appears that in some cases inventory control procedures may be considered to be equivalent to those specified for the cost plus loss method of analysis. However, the concepts of inventory control are very helpful in developing the appropriate cost plus loss function for certain types of management situations. In these situations the inventory control concepts provide a guide both to the factors to be included in the cost plus loss function and to the way in which these factors enter the function.

In other cases inventory control procedures appear to offer an intriguing new approach to the analysis of this type of a problem. Unfortunately at present the procedures necessary to formalize this method of analysis are
not clearly defined. Several alternatives exist which appear to offer promise. One of these deals with the application of Bayesian concepts to combine prior information with currently collected sample survey information to obtain a posterior distribution of estimates to be used in precision specification.

A second alternative for developing a model for the inventory control method of analysis deals with a two parameter situation similar to the first suggested method for developing a precision decay function. The two parameters of this model are \( \theta \), the true state of nature and \( \beta \), a measure of change in \( \theta \) over time. Using this concept, the distribution of estimates at any given time, \( t \), is dependent on both \( \theta \) and \( \beta \). The resulting precision of sample survey information at any time, \( t \), is then a function of the precision of the estimate of \( \theta \) and the precision of the estimate of \( \beta \). A major question here is how the precisions of these two estimates should be combined to evaluate the overall precision of sample survey information available at any time, \( t \).

These alternatives offer some interesting possibilities for future research studies. However, this research will not be included as a part of this study.
CHAPTER 5. LOSS FUNCTIONS

Development and Justification

In each of the proposed methods of analysis discussed in the previous chapter, the development of a loss function that adequately describes the losses incurred by the decision maker when decisions are made based on sample survey information is one of the primary data needs that must be fulfilled. Therefore, it is important to consider how such a loss function would be developed for a forest management situation.

Since each specific type of forest management situation will have a unique loss function, I will use a timber sale appraisal situation as an example. The decision rule to be applied in this situation is the same one that is to be applied in the analysis described in Chapter 4. Very briefly, this decision rule specifies that the timber on a tract of land will be purchased on the basis of a sample survey of the tract. The price per thousand board feet to be paid for the timber is set before the sale.

In this discussion I will describe the various sources from which losses can arise in this situation. This discussion of sources of losses will be used to attempt to justify the use of the commonly applied loss functions (e.g., squared error loss or absolute value loss). If these loss functions may be justified as adequate approximations, analytical solutions to the precision specification problem are available and will be reported. If more complex loss functions appear to be called for, exact analytical solutions become unavailable and the numerical methods of analysis reported in Chapters 4 and 5 are more practical.

The most obvious source of loss in this forest management situation is that incurred because the volume paid for is not the true volume on the
stand. This source of loss is measured by the difference between the actual value of timber on the tract and the amount paid for by the purchaser. This source of loss is clearly perfectly described by a straight linear loss function.

The other sources of loss that must be considered are either associated directly with this deviation of estimated volume from true volume or occur as a result of it. In looking at these other sources of loss, it is probably easiest to consider losses resulting from overestimates of volume separately from losses resulting from underestimates of volume.

I will first consider losses occurring when the estimate of total volume proves to be an overestimate. The first source of loss that occurs when the estimate of total volume is an overestimate is a result of the fact that the mill or processing plant for which the timber is being purchased has a specified volume of timber that must be procured if the mill is to operate at full capacity and efficiency. Thus if the estimate of total volume on the stand is an overestimate, the mill will either have to operate at less than full capacity or additional timber will have to be purchased from other sources. In either case the loss that occurs would appear to be linearly related to the magnitude of the overestimate.

A second possible source of loss that might occur when the estimate of total volume is an overestimate occurs because harvesting crews were hired to harvest the estimated volume. This source of loss could take two possible forms. First, if the harvesting crews were paid a fixed amount for harvesting based on the estimate of total volume, they would have been paid too much per unit volume of timber actually harvested. This form of loss also appears to be linearly related to the magnitude of the overestimate of
total volume. The second form this source of loss might take is rather
difficult to evaluate. This form of loss occurs when the harvesting crew
is paid on the basis of timber actually harvested. Loss occurs here in the
form of dissatisfaction of harvesting crews. These crews were hired to har­
vest the estimated volume but were paid only for the amount harvested.

This loss could be evaluated in two possible ways. A clause could be
written into the contract with harvesting crews that provides for a penalty
to be paid to the harvesting crew if the estimate of volume on the tract
proves to be an overestimate. For this method of evaluation, it would
appear that the loss is linearly related to the magnitude of the overesti­
mate. The second possible method of evaluation would be valid when no such
clause is included in the contract. In this situation the dissatisfaction
of harvesting crews would represent a loss to the timber purchaser because
the harvesting crews would seek higher wages in the future as protection
against the possibility of an overestimate of timber volume available. It
is difficult to determine the relationship between loss and the magnitude
of the overestimate in this situation.

An underestimate of total volume available results in a different set
of sources for loss. In many ways, however, this set of losses is very
similar to the sources of loss for overestimates of volume. Again the first
source of loss is a result of the fact that the mill or processing plant has
a specified volume of timber which must be obtained if the mill is to oper­
ate at peak efficiency (i.e., full capacity). When the total volume of
timber available on a given tract is underestimated, additional supplies of
timber will have been purchased in order to fulfill this volume requirement.
Thus loss occurs as the amount paid for the unnecessary additional timber purchased.

It would appear that this excess supply of timber could be stored until it was needed. However, this would involve storage costs associated with losses due to decay or other reductions in timber quality. In addition, it must be considered that this cost was an additional cost incurred in the given operating year, and that although it might be recovered to a large degree in the following year, it must be accounted for to some extent in the given year.

The second source of loss occurring as a result of an underestimate of timber volume is associated with the harvesting costs that are experienced in the given year. Clearly since original estimates of harvesting costs are based on the estimate of volume available, the costs of harvesting the timber available will be greater than had been estimated. Both this source of loss and the first source of loss occurring as a result of an underestimate of volume appear to be linear functions of the magnitude of the underestimate.

Other sources of loss may exist for this situation. However, those sources listed above appear to be the most obvious and readily explained. On the basis of these losses it would appear that some form of an absolute value loss function should provide an adequate approximation to the overall loss that could be expected to be experienced when operating with this decision rule. It is quite possible, however, that the price coefficients appropriate for overestimates and underestimates may not be equal.

As the above discussion indicates, the proper form of loss function for this situation is
where \( a \) and \( b \) indicate the relationship between the magnitude of overestimate or underestimate and the size of loss incurred by the decision maker. Similar parameters are needed in all loss functions used to measure loss in this decision making situation. For the timber sale appraisal problem discussed in Chapter 4 it is assumed that \( a \) and \( b \) are equal to one. While this assumption is clearly unrealistic, it serves a purpose in this study by removing another source of variation from the precision specification problem. As the analytical solution will show, the sensitivity of precision specification to variations in this parameter is similar to that experienced due to variations in \( \Pr \).

Using these same forms of loss, it is also possible to construct a set of conditions for which a squared error type loss function could provide an adequate approximation to losses incurred. This could occur if some of the sources are insignificant until the magnitude of the overestimate or underestimate reached a given size. At this point the slope of the absolute value loss function is increased. Then when this magnitude reaches a second level, other sources of loss may become significant or the magnitude of the previous forms of loss may increase. If this trend continues as the magnitude of overestimate or underestimate increase the general form of loss would appear as follows.
While in fact this loss function is made up of a series of segments of absolute value loss, a squared error loss function clearly could be used to approximate the form of the loss function.

In view of this discussion, it is first necessary to evaluate the loss that is incurred by the decision maker for each magnitude of overestimate or underestimate of volume. This information would have to be obtained by interviews with potential timber purchasers. This would involve the development of a questionnaire to be used in the interview procedure. In such a situation it is quite possible that some firms might hesitate to reveal some of the types of information necessary for a complete description of the appropriate loss function. This hesitation would be motivated by a desire to avoid the publication of operating procedures to competitors. Thus this questionnaire would have to be designed in such a way that the necessary information could be obtained without antagonizing the individual or firm from which information is sought.

Analytical Solutions

If the loss function appears to be adequately approximated by a very simple loss function such as the absolute value loss function or the squared error loss function, an analytical solution to the optimal sample size problem is readily available. This solution is developed by the use of the principles of differentiation. Although I stated previously that the cost coefficients for either of these loss functions might be expected to be unequal for overestimates and underestimates, for the development of an analytical solution to the optimal sample size problem I will assume that cost coefficients are equal for overestimates and underestimates of volume.
The expected value of the absolute value loss function may be expressed as

$$E(|Prz - Pr@f|) = 2PrN\sigma^2/\sqrt{2m\pi}$$

where \(Pr\) is the appropriate cost coefficient, \(z\) is the estimate of volume, \(q\) is the true volume, \(N\) is the population size, \(\sigma^2\) is the population variance of the sampling units, and \(n\) is the sample size. The development of this expected value and the expected value of the squared error loss function will be discussed in more detail in Chapter 4. However, I should indicate here that for this evaluation of expected loss it is assumed that estimates of total volume, \(z\), are approximately distributed as \(N(0, N^2, \sigma^2/n)\).

In order to provide a tie between the results of this chapter and the results of Chapter 4, I will also use the cost function that is used in Chapter 4. This cost function has the form

$$Cost = C(420T_1 + T_2n)/60$$

where \(C\) is the total salary paid the sampling crew in dollars per hour, \(T_1\) is the time, in minutes, required to travel a chain between plots, \(T_2\) is the time, in minutes, required to measure a 1/5-acre plot, and \(n\) is the sample size. Again, the discussion of the development of this cost function will be included in Chapter 4.

Thus the cost plus loss function that must be minimized for the solution to the optimal sample size problem is of the form

$$f(n) = 2PrN\sigma^2/\sqrt{2m\pi} + C(420T_1 + T_2n)/60$$

where \(f(n)\) is the notation used for the cost plus loss function.

A minimum, a maximum, or an inflection point of \(f(n)\) exists at the value of \(n\) for which the first derivative of \(f(n)\) with respect to \(n\) equals zero. Application of the basic principles of differentiation indicate that
\[ \frac{\partial f(n)}{\partial n} = -\frac{PrN0}{(2\pi n^{3/2})} - C_{T_2}/60 \]

The minimum or maximum value of \( f(n) \) is reached when this function is set equal to zero and solved for \( n \). Thus the specified \( n \) is defined by

\[ n = \left( \frac{1800Pr^2N^2\sigma^2}{C_{T_2}} \right)^{1/3} \]

In order to assure that this solution provides a minimum for the cost plus loss function, it is necessary to consider the second derivative of \( f(n) \) with respect to \( n \). This is

\[ \frac{\partial^2 f(n)}{\partial n^2} = \frac{3PrN0}{(2\pi n^{5/2})} \]

Since this function of \( n \), when evaluated at the value of \( n \) derived above, is positive, the equality

\[ n = \left( \frac{1800Pr^2N^2\sigma^2}{C_{T_2}} \right)^{1/3} \]

defines the value of \( n \) which minimizes the cost plus loss function.

Since \( N \) and \( \sigma^2 \) are population parameters which cannot be manipulated by the decision maker for a given population, this analytical solution of the optimal sample size problem indicates that for the absolute value loss function, the required sample size is dependent only on stumpage price, salary paid the survey crew, and time required to measure an individual 1/5-acre plot. However, if one is considering all factors that affect the determination of required sample size, both \( N \) and \( \sigma^2 \) must be included.

The general conclusions that can be drawn from this solution are not particularly startling. As would be expected an increase in stumpage price (i.e., an increase in value of timber) results in an increase in optimal sample size. Similarly an increase in the cost of sampling results in a decrease in optimal sample size. Also, a population of larger size or larger variance will require a larger sample size than a smaller population or a population with smaller variance.
If the squared error loss function is assumed to provide an adequate approximation to the losses that are experienced by the decision maker, this same procedure may be followed to provide a solution to the optimal sample size problem. For this loss function, the expected value of loss may be expressed as

\[ E(\text{Pr}_{z} - \text{Pr}_{0})^2 = \text{Pr}^2 \frac{\sigma^2}{n} \]

where the notation is the same as was used for the absolute value loss function. Since the same cost function is used here as was used with absolute value loss, the cost plus loss function, \( g(n) \), can be expressed as

\[ g(n) = \text{Pr}^2 \frac{\sigma^2}{n} + C(420T + T^2n)/60 \]

The first derivative of the cost plus loss function is not unlike the similar result for the absolute value loss function. The first derivative of \( g(n) \) with respect to \( n \) is expressed as

\[ g'(n) = -\text{Pr}^2 \frac{\sigma^2}{n^2} + CT_2/60 \]

Again setting this first derivative equal to zero and solving for \( n \) results in the evaluation of that value for \( n \) which minimizes or maximizes \( g(n) \). This evaluation results in \( n \) being specified as

\[ n = \sqrt{60\text{Pr}\sigma^2/CT_2} \]

In order to determine whether a minimum or maximum value of \( g(n) \) exists at this point, the second derivative of \( g(n) \) with respect to \( n \) must again be evaluated. This differentiation yields

\[ g''(n) = 2\text{Pr}^2 \frac{\sigma^2}{n^3} \]

Since this function, when evaluated at

\[ n = \sqrt{60\text{Pr}\sigma^2/CT_2} \]

is positive, the existence of a minimum value for \( g(n) \) at this point is confirmed.
This solution indicates that the same set of parameters influence the specification of sample size when a squared error loss function is used as is the case when the absolute value loss function is used. However, when the squared error loss function is used, the response to variations in these parameters is greater than is the case when the absolute value loss function is used. This results because, for the absolute value loss function, \( n \) is related to each of the parameters raised to the \( 2/3 \) power while for the squared error loss function the relationship is with the parameters raised to the first power.

Similar procedures may be applied to a loss function that can be readily expressed in strict functional form. However, as the loss functions become more complex, the first derivatives of the associated cost plus loss functions also become more complex. The exact solution of these functions for the evaluation of the optimal sample size soon becomes unavailable.

This can readily be demonstrated by observing the first derivative of the third loss function that will be discussed in Chapter 4. For this loss function, loss is described as absolute value loss for underestimates of volume and as squared error loss for overestimates of volume. The expected value of the cost plus loss function is

\[
f(n) = \Pr N \frac{\bar{d}^2}{2n} + \Pr N \frac{\bar{d}}{2n}^2 + C\left(4\pi T_1 + T_2 n\right)/60
\]

The first derivative of \( f(n) \) with respect to \( n \) may be expressed as

\[
\frac{df(n)}{dn} = -\Pr N \frac{\bar{d}^2}{2n^2} - \Pr N \frac{\bar{d}}{2n}^2 + CT_2/60
\]

Setting this function of \( n \) equal to zero and collecting terms results in the following equation.

\[
CT_2 n^2 - (30\Pr N \frac{\bar{d}}{2\pi}) n^{1/2} = 30\Pr N \frac{\bar{d}^2}{2}
\]
For this loss function and for the more complex functions that are reported in Chapter 4, the numerical techniques described in Chapter 4 appear to provide a more manageable solution to the optimal sample size problem.

Bidding Strategy

In most actual timber sale situations the stumpage price paid for timber is not set in advance but is determined by the process of competitive bidding. For sales made by the U.S. Forest Service, this procedure is commonly followed. However, a timber appraisal is made to determine a minimum acceptable price. If the competition among buyers does not result in this published minimum price, the sale is not made. A brief review of Forest Service procedures used to determine this minimum price provides an introduction to the factors that must be considered in the development of a sound bidding strategy.

According to the Forest Service Manual (1966),

"Fair market value or appraisal value as used by the Forest Service is based on the operator of average efficiency and is aimed at a market value which will interest sufficient purchasers to harvest the allowable cut under multiple use and sustained yield principles. In accomplishing this objective consideration must be given to providing an adequate margin for profit and risk which will be sufficient to maintain operations over the long run and thus provide a stable market for National Forest Timber."

One of the major costs that must be considered in this determination is the cost associated with constructing permanent access roads for the tract of timber being sold. The Forest Service feels that a large portion of the cost of such permanent lots should be paid by the public. Thus, the minimum appraised value is adjusted for these costs.

Three other basic factors are included in the appraisal procedure. The first of these is an estimate of the ultimate value that the purchaser
may be expected to realize from the sales of timber products harvested from
the tract. A second factor to be considered is the costs of harvesting,
transporting, and processing timber that the purchaser will experience.
The final factor to be considered is an allowance for the risks involved
in investing in the purchase of standing timber.

When a potential timber purchaser develops his bidding strategy, he
must consider all of these factors. A good deal of information is needed
as input to the determination of the bid stumpage price. Included in this
information are estimates of harvesting, transportation, and processing
costs; an investigation of potential markets for final product; and estimates
of volume and quality of timber available. In this discussion I will con­sider only the effect of volume of timber on the bid stumpage prices. How­ever, volume of timber available may be considered to have some effect on
the evaluation of the other factors as well.

In this forest management situation, the decision maker (i.e., poten­tial timber purchaser) is faced with a more complex loss function than is
the case when stumpage price is fixed. For this situation loss results from
two possible sources. Both sources of loss result when the estimated volume
differs from the actual volume on the stand. The first source of loss occurs because plans made for operations based on the estimate of volume
available must be revised when the true volume becomes known. The second
source of loss is a result of the fact that the stumpage price paid for the
timber is based on the estimate of volume rather than on the true volume.
Clearly these two sources of loss are closely related. It is this rela­tionship that results in the more complex form for the resulting loss
function.
The nature of the first source of loss has been discussed in detail. However, the second source of loss that occurs in this forest management situation needs further explanation. In order to discuss the losses that result from this source, it is first necessary to consider the influence of the volume estimate on stumpage price specification. The basic concept to be considered is the fact that the greater the estimate of volume in the proposed sale the greater the bid stumpage price (on a per thousand board foot basis) may be. This is true because for a larger sale, the fixed costs (i.e., those unaffected by the amount of available timber) associated with the purchase, harvesting, transporting, and processing of the timber may be spread over a larger base.

In light of this fact, an overestimate of volume in the sale results in a bid stumpage price greater than what should have been bid had the true volume been known. Thus the resulting loss appears to be a function of the magnitude of the overestimate of the bid stumpage price. This loss is magnified by the fact that in addition to paying too high a price for the timber on a per unit volume basis, the decision maker is also paying this higher stumpage price for more timber than is actually available.

By this same reasoning, an underestimate of volume available results in a lower bid stumpage price than should have been bid had the true volume been known. When this occurs, the major source of loss to the potential purchaser results from the fact that his low bid lowers the probability that his bid will be the high bid for the timber. Thus, he increases the chances that he will have to look elsewhere for a source of harvestable timber.

It seems reasonable to assume that since there clearly is a degree of risk involved in using a bidding strategy based on the estimated volume of
timber available, the bidding strategy should also include a consideration of some measure of the reliability of the estimated volume. This may be accomplished by use of the coefficient of variation of the volume estimate where coefficient of variation is defined as

$$c.v. = \frac{S}{\bar{x}}$$

$S_{\bar{x}}$ is the standard deviation of the estimated mean volume per acre, $\bar{x}_n$, of timber available for harvest.

This information may be used to develop a potential bidding strategy for timber purchasers. By evaluating the expected costs of harvesting, transporting, and processing timber; the potential return from sales of timber products; and the risks involved, the decision maker must determine the appropriate stumpage price for each class of possible volume estimates. The reliability of the estimate of volume available may be included in the bidding strategy by evaluating the coefficient of variation of the volume estimate. If the coefficient of variation of the estimate is greater than some specified value, say seven tenths, the bid stumpage price is lowered from the price associated with the volume estimate to the price that would be specified for the next lower volume class.

This portion of the bidding strategy implies that the loss function that applies in this situation is somewhat one sided. The loss associated with making a bid that is too high and thus paying too much for the timber is greater than the loss associated with making a bid that is too low and thus reducing the probability of the decision maker's bid being the high bid.

I offer this proposed bidding strategy not as an example of an actual bidding strategy used in timber sale appraisals but merely as an example of potential bidding strategy. This strategy includes some of the major com-
siderations that must be included in the development of an adequate bidding strategy and provides a model which may be used to indicate the procedures necessary for the development of an associated loss function.

The use of a squared error loss function to approximate the losses incurred by the decision maker when the estimate of total volume on the stand is an overestimate appears to be a sound procedure. Clearly this component of loss is not linearly related to the magnitude of the overestimate. Since, as I have indicated, the stumpage price paid for the timber is based on the estimate of volume available, an overestimate of volume results in both payment for more timber than is available and in payment of a price per unit volume that is greater than should have been paid. For each increase in the magnitude of the overestimate the stumpage price per unit volume increases and this increased price is paid not only for the increase in magnitude of the overestimate but for the entire estimate of total volume. Thus a squared error type (i.e., quadratic) loss function appears to be appropriate for this measure of loss.

The use of this form of loss function for overestimates of estimated stand volume implies that the second decision variable, the coefficient of variation, is less than the specified limit (seven tenths here). If the coefficient of variation is greater than this limit, the general shape of the loss function is not altered. However, the price coefficient that is applied is reduced from what it would have been if the coefficient of variation had not exceeded the limit.

As was the case with the simple loss functions discussed previously in this chapter, interviews would have to be conducted with individual decision makers (i.e., potential bidders) to evaluate the appropriate price
coefficients for the loss function. These interviews would also serve to provide the necessary information to test the applicability of this form of loss function.

When the estimate of total volume available is an underestimate, the form of the appropriate loss function appears to vary from the form appropriate for an overestimate. The actual form is difficult to predict without the use of empirical data that would have to be collected by interviewing potential bidders. However, either an absolute value type loss function or a squared error type loss function with the appropriate price coefficients would appear to be logical forms of approximation. It is also possible that the actual form of loss for this component of the loss function does not lend itself to approximation by any functional form but must be evaluated individually for each class of underestimate. Each underestimate in this situation would be assigned to a class based on the magnitude of the underestimate and on the size of the coefficient of variation.

In addition to problems concerning the form of the appropriate loss function to be utilized when this form of bidding strategy is used in decision making, the evaluation of the expected value of loss requires the development of a probability measure more complex than the probability measure used with the simpler loss functions used when a prespecified stumpage price is utilized. Since the price paid for timber is a function of both total volume available and the coefficient of variation of this estimate of total volume, the evaluation of the expected value of loss requires the development of the joint probability of the estimate of total volume and its coefficient of variation.
Let
\[ \bar{x} = \text{estimate of mean volume per acre of timber available based on a} \]
\[ \text{sample of size } n \]
\[ S^2_x = \text{variance of the estimated mean} \]
\[ N = \text{population size in acres} \]

Then the necessary joint density function of \( N \bar{x} \) and \( S^2_x / \bar{x} \) may be denoted as \( f(N \bar{x}, S^2_x / \bar{x}) \). Since \( \bar{x} \) and \( S^2_x \) are distributed independently, their joint density function \( f(\bar{x}, S^2_x) \), may be expressed as the product of their individual density functions. \( \bar{x} \) is approximately normal with mean \( \mu \) and variance \( \sigma^2 / n \). \( n(n-1)S^2_x / \sigma^2 \) has an approximate chi-square distribution with \( n-1 \) degrees of freedom. Thus

\[ f(\bar{x}) = \left( \frac{1}{\sqrt{n} / \sqrt{2\pi}} \right) \exp\left(-n(\bar{x} - \mu)^2 / 2\sigma^2\right) \]

and

\[ f(S^2_x) = \left(1 / \Gamma(n/2)\right)(n(n-1)/2\sigma^2)^{n/2} \left(S^2_x / \sigma^2\right)^{(n-2)/2} \exp\left(-n(n-1)S^2_x / 2\sigma^2\right) \]

Therefore,

\[ f(\bar{x}, S^2_x) = \left(1 / \Gamma(n/2)\right)(n(n-1)/2\sigma^2)^{n/2} \left(S^2_x / \sigma^2\right)^{(n-2)/2} \exp\left(-n(n-1)S^2_x / 2\sigma^2\right) \]

\[ \times \left(1 / \sqrt{n} / \sqrt{2\pi}\right) \exp\left(-n(\bar{x} - \mu)^2 / 2\sigma^2\right) \]

Let \( y = N \bar{x} \) and \( z = S^2_x / \bar{x} \). Then by algebraic manipulation, \( \bar{x} = y/N \) and \( S^2_x = y^2z^2/N^3 \). The proper Jacobian for performing this transformation is

\[ J = 2y^2z^2/N^3 = (2(N \bar{x}^2) S^2_x \sqrt{\bar{x}} )/N^3 \]

and the transformed density function is

\[ f(y, z) = \left(1 / \Gamma(n/2)\right)(n(n-1)/2\sigma^2)^{n/2} \left( y^2z^2 / N^3 \right)^{(n-2)/2} \exp\left(-n(n-1)y^2z^2 / 2\sigma^2\right) \]

\[ \times \left(1 / \sqrt{n} / \sqrt{2\pi}\right) \exp\left(-n(y/N - \mu)^2 / 2\sigma^2\right) 2yz^2 / N^3 \]

Thus the joint density function of \( N \bar{x} \) and \( S^2_x / \bar{x} \) is

\[ f(N \bar{x}, S^2_x / \bar{x}) = \left(2 / \Gamma(n/2)\right)(n(n-1)/2\sigma^2)^{n/2} \left( N \bar{x}^2 \right)^{(n-2)/2} \exp\left(-n(n-1)(N \bar{x}^2) / 2\sigma^2\right) \]
\[ \exp(-n(x - \mu)^2/2\sigma^2) \]

While the development of the appropriate density function is straightforward, the application of this density function to obtain the expected value of the loss will very likely pose major difficulties. It appears that the necessary integration would have to be done by means of numerical approximation procedures.
CHAPTER 4. TIMBER SALE APPRAISAL PROBLEM

Decision Rules

Timber sale appraisal was mentioned in Chapter 1 as an example of a forest management decision making situation to which the precision specification problem could appropriately be applied. A large number of decision rules exist which might be applied to this general management situation.

One of the simplest of these decision rules might be applied when a given tract of timber is being offered for sale at a given total price. In this situation the decision maker is faced with two alternatives. These are to either purchase the timber or to not purchase the timber. The decision maker must first determine what volume must exist on the tract if he is to purchase the timber at the offered price. This volume would be determined by an economic evaluation of the forest industry in the geographic area of interest. Once this value is determined, it is then only necessary to obtain an estimate of the volume on the tract. If this estimate is greater than the level specified by the economic evaluation, the decision is to buy the timber at the specified price. Otherwise, the timber should not be purchased.

A modification of this simple decision rule is proposed that makes it somewhat more realistic. The specification of the level above which the purchaser will buy the timber should depend, to some extent, on the variance of the estimate of volume that the decision is to be based on. This has the effect of bringing the risks associated with decision making more closely into the actual decision making process.

This modification might be accomplished by looking at the confidence interval around the estimate of volume rather than at the point estimate.
itself. In this case the decision rule would specify that if the lower limit of the, say, ninety-five percent confidence interval is greater than or equal to some purchase limit, then the timber should be purchased. Under this decision rule the purchase limit represents that volume below which the decision maker would experience greater losses by purchasing the timber than he would experience if he did not purchase the timber. This rule is equivalent to saying that the decision maker should buy the timber if, at the ninety-five percent level of confidence, the true volume is greater than the purchase limit.

This decision rule is dependent upon the variance of the estimate and thus provides the buyer some protection from the risk of an incorrect decision. This protection is provided by the fact that the amount by which the volume estimate must exceed the purchase limit is directly proportional to the square root of the variance of the volume estimate. Thus for a population with a relatively high variance estimate, the decision rule specifies that a proportionally more severe test must be satisfied by the volume estimate if the purchase is to be made than would be the case for a second population with equal observed volume but with a smaller variance.

The actual decision rule that I am using in this analysis is somewhat different than the two possibilities I have just discussed. The decision rule specifies that the tract of timber will be purchased on the basis of the estimate of total volume derived from a sample survey of the tract. The total amount paid for the timber is obtained by multiplying the estimate of total volume by the appropriate stumpage price. This stumpage price is determined before the estimate of total volume is obtained from the sample survey information.
A more complex decision rule is not used for this first application for two reasons. First, it appears advantageous to be able to initially study the precision specification problem in a simple management situation where outside sources of variation may be held to a minimum. The second advantage of this decision rule is that its use allows me to look at several possible forms of loss functions and investigate the sensitivity of precision specifications to changes in the loss structure that is applied in analysis.

Population Description

Johnson and Nixon (1952) provided a square forty-acre population of old growth Douglas-fir. This forty-acre population is located on the Blue River Experimental Forest within the Willamette National Forest in Lane County, Oregon. The average volume per acre on the stand is approximately 91,000 board feet. The average diameter of individual trees is about forty-five inches d.b.h. The stand includes a small amount of hemlock, sugar pine, and cedar in addition to the Douglas-fir.

The population data consists of an enumeration of the 1600, 1/40-acre plots that exist in a forty-acre area. Observations are recorded as the total timber volume in Scribner Decimal C on each plot. Thus the population may be expressed as a forty by forty matrix of values, each representing the total volume, utilizing the Scribner Decimal C log rule, on a 1/40-acre plot.

In addition, the results of a time study conducted on this population are available. This study indicates the average time required per chain of travel between plot boundaries is 1.84 minutes. Also the average time required to measure various sized plots is recorded. Each plot size included in the tabulation represents some combination of the original 1/40-acre plots.
The population is easily transferred to eighty-column computer cards. Since individual plot values, in Scribner Decimal C, range from zero to 2721, a four-column field is set aside for each value. This necessitates the use of two eighty-column cards to record each row of the population matrix.

This basic population is modified in two ways in order to generate two additional arbitrary populations which have less volume per acre and thus, which have less value. The first modification is accomplished by dividing the observed volume on each 1/40-acre plot by ten. This results in the variance of the observations being 1/100 the variance of the original observations.

The second generated population is formed by dividing each observed volume by one hundred. The corresponding variance is equal to the original variance divided by 10,000. For each of these two new populations it is assumed that the results of the original time study conducted on the Douglas-fir population are still valid. The volumes on these two populations correspond fairly closely to those that might be expected, respectively, on well stocked and average to below average stocked stands of some eastern hardwoods.

Procedures

The method of analysis that is applied to this forest management decision making situation is the general maximum utility method described in Chapter 2. This is one of those situations for which the use of an opportunity loss function proves advantageous. Thus the optimum precision level will be that precision level that minimizes the expected opportunity loss.
The true state of nature, S, is known to be equal to 3,652,640 board feet for this population. This is one of the conditions that must be met if the maximum utility method of analysis is to be considered equivalent to a cost plus loss approach to analysis. Since the other assumption necessary for this equivalence to exist is also met (i.e., the decision rule is completely specified), the method of analysis that is applied to this problem could also be considered to be the minimization of a cost plus loss function.

For this analysis the 1600, 1/40-acre plots have been combined to form 200, 1/5-acre plots. This combination is accomplished by arbitrarily grouping the existing 1/40-acre plots into four plot by two plot 1/5-acre plots. The resulting population matrix is a twenty by ten matrix of values, each of which represents the total volume on a 1/5-acre plot. The mean volume on these plots is 18,263.2 board feet and the variance of the population of observations is 108,490,800 board feet squared.

Three basic data needs must be met for this modified maximum utility method of analysis to be carried out. These are the development of a cost function that describes the cost of sampling associated with each precision level, the evaluation of a measure of probability that expresses the probability of observing any outcome, z, given the sample survey was conducted at precision level, e, and finally the development of a loss function that adequately describes the losses to the decision maker (i.e., timber purchaser) resulting from making an incorrect decision based on sample survey information collected at any given level of precision.

The first of these data needs to be considered is the development of a cost of sampling function. In order to use the time studies conducted by
Johnson and Hixon (1952) I assume that sample survey information is collected by a two-man crew. It will further be assumed that each member of the crew will be paid at the rate of six dollars per hour. While this may appear somewhat high, I attempt to justify it by suggesting that this higher rate might be considered to include other costs of data handling such as those costs associated with preparing the data for analysis and obtaining the sample survey estimates of total volume.

The plots to be included in the sample survey will be selected by a systematic sampling scheme with a random start. It will be assumed that for physical reasons, travel across the tract can only be in an east or west direction. Therefore, if the sample size is less than twenty, the crew will be required to walk one length of the stand for each plot taken. In addition, they will travel one width of the stand in moving from row to row. Thus the total distance travelled between plots, given that the sample size is less than or equal to twenty, is (n + 1)20 chains where n is the sample size. If the sample size is greater than twenty, the total distance travelled between plots is 420 chains.

If one defines C as the total salary paid the crew in dollars per hour, \( T_1 \) as the time, in minutes, required to travel a chain between plots, \( T_2 \) as the time, in minutes, required to measure a 1/5-acre plot, and n as the sample size, then the cost of sampling may be expressed as

\[
\text{Cost} = C \left( T_1 (n + 1)20 + T_2 n \right) / 60 \quad 2 \leq n \leq 20
\]

\[
= C \left( 420T_1 + 2T_2 n \right) / 60 \quad 20 < n \leq 200
\]

Johnson and Hixon (1952) specified that a good estimate for \( T_1 \) in this type of forest is 1.84 minutes per chain of travel. This assumes a two-man crew using staff compass and trailer tape for direction. In addition, their
time study indicates that an average time of 8.20 minutes are required to measure a rectangular 1/5-acre plot. Since the salary of each member of the crew has been specified as six dollars per hour, C is defined to be twelve dollars per hour.

In order to investigate the sensitivity of precision specification to variations in these parameters of the cost function, an additional set of arbitrarily selected time parameters was used in the analysis. This set assumes the time required to travel a chain between plots is \( T_1 = 4.00 \) minutes and the time required to measure a 1/5-acre plot is \( T_2 = 13.00 \) minutes. This set of time parameters will be referred to as the modified set of time parameters.

The second data need for this modified maximum utility method of analysis is the development of the probability measure \( P(z|e) \). As I indicated in the general discussion of the maximum utility method of analysis, the set of possible outcomes, \( \{z\} \), is frequently continuous rather than discrete in forest management situations. The suggestion was made at the time to group the possible outcomes into appropriately sized classes and then use these classes as a set of discrete outcomes in the analysis. For the analysis of this timber appraisal problem, I suggest an alternative approach for handling the continuous nature of the set of possible outcomes.

This alternative is to omit the actual analysis of \( P(z|e) \) and to instead evaluate directly the expected value of the loss resulting from making decisions based on sample survey information collected at any specified level of precision. This direct evaluation is possible for two reasons. First, the true state of nature, \( \theta \), is known for this population and second, the decision rule is fully specified. Therefore, each possible
outcome, $z$, has associated with it a unique action, $a$, which, if the appropriate loss function is defined, results in the complete definition of the loss that results from taking this action given $\theta$ exists as the true state of nature. In addition, since the set of outcomes, $\{z\}$, is made up of possible volume estimates for the total volume on the stand, evaluated as a linear function of the sample mean, the distribution of the members of the set $\{z\}$ may be approximated by use of the Central Limit Theorem.

Since the population being used in this analysis is a finite population, there arises a question as to whether the normal approximation is valid for small sample sizes (i.e., near zero) and for large sample sizes that approach a complete enumeration of the population. The following frequency distribution indicates the general shape of the distribution of the population of 200, 1/5-acre plots.

A computer program was written to generate 2,000 repeated samples of specified sample size from this population. The sampling scheme followed in selecting sampling units to be included in each of the 2,000 samples was simple random sampling without replacement. Sample sizes included were 180 and 134, ..., 198 and their compliments (i.e., 20 and 2, ..., 16).
For each sample size a frequency distribution of estimates was reported. The discrete classes of the distribution were similar to those indicated for the frequency distribution of the parent population. Statistics for skewness and kurtosis and a chi-square for goodness of fit were calculated. The chi-square goodness of fit statistic indicates that for \( n \) greater than 188 the distribution of estimates differs from a normal distribution. Investigation of the skewness and kurtosis statistics indicates that this deviation from normality is due to positive skewness of the population of estimates. By arguments of symmetry, analogous conclusions can be drawn for the small sample sizes, \( n < 12 \). The normality assumption appears adequate for \( n \) between 12 and 188.

Some of the loss functions used in this analysis resulted in specified optimal sample sizes less than 12 or greater than 188, sizes for which the normality assumption is invalid. Therefore, the analysis of the precision specification problem for one of these loss functions, \( L_3 \), was conducted using the actual frequency distribution obtained from the computer rather than the normal approximation. \( L_3 \) defines loss as absolute value loss for underestimates of volume and as squared-error loss for overestimates of volume.

For the modified population generated by dividing each observed plot volume by one hundred, the optimal sample size when stumpage price is set at twenty dollars per thousand board feet is 187 plots. When the observed frequency distribution is used in place of the normal distribution, the optimal sample size is 186 plots. For this same population, when stumpage price is set at two dollars per thousand board feet, the optimal sample size is seven plots. The use of the observed frequency distribution also
results in an optimal sample size of seven plots. These calculations were performed using the second set of time parameters (i.e., $T_1 = 4.00$ and $T_2 = 13.00$). Thus it appears that the assumption of normality for this population produces no inconsistencies in the results of the analyses performed.

In order to characterize the loss associated with making decisions based on sample survey information, it is necessary to fulfill the third data need mentioned for this method of analysis. This is the development of a loss function that adequately describes the losses to the decision maker that result from making decisions based on sample survey information collected at any given level of precision. Since the form of the loss function that fits this management situation does not appear to be completely defined in the literature, several alternative forms of loss function are discussed and investigated in this study.

The simplest form of loss function that might be considered is suggested by the form of the first term of the opportunity loss function. I am referring here to a straight linear loss function of the form

$$L = (z - \theta)Pr$$

where $Pr$ is the appropriate constant for expressing loss in monetary terms. In this study $Pr$ is the stumpage price that one might expect to pay for the timber being purchased. Austin (1968) stated that the average stumpage price for west side Douglas-fir was $59.90 per thousand board feet. The corresponding value for east side Douglas-fir was $22.40. These two values are used in the analysis of the original Douglas-fir population. For the modified populations a series of stumpage prices are used in the analysis. These are twenty dollars, ten dollars, and two dollars, per thousand board feet.
Several of the assumptions necessary for the use of this form of loss function are rather difficult to justify. Primary among these is the fact that if the distribution of $z$ is symmetric around $\theta$, as is the case for this timber appraisal problem as well as for any other management situation in which the Central Limit Theorem is utilized to approximate the distribution of the members of $\{z\}$, then the expected value of the loss is zero. The implication here is that regardless of the level of precision used, in the long run losses resulting from overestimates of volume on one stand would be counteracted by equivalent gains from underestimates of volume on other stands.

For the small timber operator this form of loss function would clearly not be adequate. His limited capital would not let him sustain too large a loss even with the expectation that the future should eventually bring compensating gains and that in the long run, his expected losses would be zero. The applicability of this loss function is, thus, dependent on the ability of the timber purchaser to absorb losses on any given timber purchase with the expectation that long run losses will be zero.

An additional difficulty of this form of loss function is the fact that it requires the assumption that an overestimate of the volume represents a loss to the timber purchaser and that an equivalent underestimate of volume represents a gain to the timber purchaser of equal size. This assumption ignores any loss that results from planning harvesting and processing operations on the basis of the estimate of volume present and then discovering during the operations that the true volume on the stand varies from this estimate.
These losses occur for both overestimates and underestimates of volume. The loss due to an overestimate of volume is fairly clear. Timber that is paid for is not available. Thus, not only is the per unit price of the timber increased but also additional timber must be purchased elsewhere to meet preset demands of the processing operation. The losses resulting from underestimates of volume need more explanation. An underestimate of volume results in more timber being available than was planned. While this causes a decrease in the per unit price of timber, it may result in increased total harvesting costs. Also, to meet the prespecified demands of the processing operation, additional unnecessary timber would have been purchased from other sources. For these reasons, this form of loss function does not appear to adequately describe the losses associated with this decision making situation and thus is not included in the analysis. However, an interesting modification of this loss function is discussed later and is included in the analysis.

The second form of loss function to be discussed is an absolute value loss function. This loss function has the general form

$$L_1 = |z - \delta| \Pr$$

For this loss function to be valid, it is necessary to make two basic assumptions about the nature of the loss associated with an incorrect decision based on sample survey information. The first assumption is that the loss resulting from an overestimate of volume is exactly equal to the loss resulting from an underestimate of volume of equal magnitude. This assumption is necessary because of the symmetry of the absolute value loss function around the true volume. The second necessary assumption is that the loss is a strictly linear function of the deviation of the estimate of volume from
the true volume on the stand. The validity of either of these assumptions for this timber appraisal situation is open to some question. However, these assumptions appear to be more realistic than those required for the application of the straight linear loss function and thus, I have included this form of loss function in the analysis.

By simple integration the expected value of this loss function may be expressed as

$$ E(|Pr(z - \theta)|) = \int_{0}^{a} Pr(z - \theta) f(z) dz $$

$$ = \left[-2Pr\frac{N}{\sqrt{2\pi n}}\exp\left(-\frac{(z - \theta)^2}{2N^2}\right)\right]_{0}^{a} $$

$$ = (2Pr\frac{N}{\sqrt{2\pi n}}(1 - \exp(-\frac{(a - \theta)^2}{2N^2})) a > \theta $$

where $f(z)$ as approximated by the Central Limit Theorem is $N(\theta, N^2/2n)$.

This is the expected value of the loss in the range $(2\theta - a, a)$. For the interval $(0, \infty)$

$$ E(|Pr(z - \theta)|) = 2Pr\frac{N}{\sqrt{2\pi n}} $$

The question as to whether the loss is a strictly linear function of the departure of the sample based volume estimates from the true stand volume suggests a third form that the loss function might take. This is the often used squared error loss function and has the general form

$$ L_2 = (Prz - Pr\theta)^2 $$

This loss function is still symmetric about the true volume. Thus it is necessary again to make the assumption that an underestimate of total volume and an overestimate of equal magnitude result in equal losses to the decision maker. The second assumption necessary for this form of loss function to be valid is that loss is proportional to a quadratic function of the deviation of the estimate of total volume from the true total volume on the stand rather than being proportional to a linear function of this deviation as
is true in the previous two forms of loss function. This assumption appears to be more easily supported because it implies that small deviations result in relatively minor losses while large deviations result in much more significant losses.

The expected value of this loss function is well known. It is expressed as

$$E(Prz - Pr\theta)^2 = \int_0^\infty (Prz - Pr\theta)^2 f(z)dz$$

$$= Pr^2 \frac{2}{n} \frac{\phi^2}{n} 0 \leq z \leq \infty$$

where $f(z)$ is again approximated as $N(\phi, N^2 \sigma^2/n)$. For cases where the range of $z$ is limited to, say, $a \leq z \leq b$

$$E(Prz - Pr\theta)^2 = \int_a^b (Prz - Pr\theta)^2 f(z)dz$$

$$= (-Pr^2 N(z - \phi)\phi/2n)\exp(-n(z - \phi)^2/2N^2 \sigma^2)\bigg|_a^b$$

$$+ (Pr^2 Nz^2 \sigma^2/n)\int_a^b f(z)dz$$

The next five loss functions to be considered consist of combinations of the squared error loss function and the absolute value loss function. The first of those combinations, designated as $L_3$, is a loss function in which loss is described as absolute value loss for underestimates of volume and as squared error loss for overestimates of total volume on the stand.

The arguments used to justify the exclusion of the straight linear loss function from this analysis seem to imply that losses associated with overestimates of stand volume may not be equal to losses associated with underestimates of equal magnitude. This form of loss function represents an attempt to express this fact in the evaluation of expected loss. Since this function is not symmetric about the true stand volume, the assumption of symmetry of losses, necessary for the previous loss functions, is
unnecessary hero. However, it is still necessary to assume that the losses associated with underestimates are proportional to the square of the deviation of the estimate of total volume from the true volume and that losses associated with overestimates are proportional to a linear function of those deviations.

The remaining four loss functions formed by combinations of the absolute value and the squared error loss functions utilize the absolute value loss function in an interval about the true mean and the squared error loss function outside this interval. For two of these loss functions the stumpage price is increased to 1.2 Pr for any estimate of total stand volume falling outside the interval. These four loss functions are identified as:

- $L_4$, interval is true volume plus and minus 1/10 true volume
- $L_5$, interval is true volume plus and minus 1/10 true volume with penalty added to stumpage price
- $L_6$, interval is true volume plus and minus 1/100 true volume
- $L_7$, interval is true volume plus and minus 1/100 true volume with penalty added to stumpage price.

These loss functions are not continuous at the point of change from absolute value loss to squared error loss. This implies that at the point of change an additional source of loss, for example, one such as a penalty for excessive error in the estimate, is encountered.
In addition to the loss functions that I have been discussing, I also consider one-sided loss functions of a similar nature in all cases where this procedure is reasonable. The combination loss function for which loss is absolute value loss for underestimates of volume and squared error loss for overestimates of total stand volume clearly could not be included here. However, the remaining loss functions that I have discussed, with the exception of the straight linear loss function, are considered. These functions represent the case where loss is incurred if an overestimate of volume results but no loss occurs from underestimates of total stand volume. This general form of loss function is suggested in a problem by Cochran (1963).

These loss functions are designated as $L_8$, $L_9$, $L_{10}$, $L_{11}$, $L_{12}$, and $L_{13}$. $L_8$ is the one-sided absolute value loss function and $L_9$ is the one-sided squared error loss function. $L_{10}$, $L_{11}$, $L_{12}$, and $L_{13}$ are, respectively, the one-sided loss functions corresponding to $L_q$, $L_5$, $L_6$, and $L_t$.

A final approach to the formation of a loss function is considered to handle the situation in which the decision maker cannot sustain losses beyond a certain limit. This form of loss function is a strictly linear loss function in a specified range about the true total stand volume. Outside this specified range when an overestimate occurs, the purchaser cannot absorb the loss that results and thus is forced out of business. The expected value of this loss function within the specified limits is zero. However, outside the limits the expected loss equals the probability of an estimate of total stand volume exceeding the upper limit of the specified range multiplied by the loss associated with being forced out of business. This loss function is labeled $L_{14}$. 
Several sets of limits for the specified range and several values to associate with the loss of being forced out of business are considered in this analysis. The limits that are included in the analysis are overestimates of 500, 1000, 5000, 13500, and 18000 board feet. These correspond approximately to five-tenths percent, one percent, five percent, ten percent, fifteen percent, and twenty percent overestimates. The values associated with the loss of being forced out of business are arbitrarily set as 50,000 dollars, 100,000 dollars, and 150,000 dollars.

As I indicated at the beginning of this discussion, each of these proposed forms of loss function is included in the analysis of the precision specification problem in the timber sale appraisal problem. This is done for two basic reasons. First, the form of loss function which is most appropriate for this situation is not fully specified. Second, according to Mood and Graybill (1965),

"It is impossible, for instance, to specify accurately the loss incurred in making an erroneous decision in asserting a scientific hypothesis... Experience with statistical problems shows that "good" procedures are insensitive to small changes in the loss function, especially when considerable data are available. Thus precise values of the loss are not absolutely necessary."

Thus, I would like to consider several forms of loss function in this analysis in order to determine the sensitivity of this problem to various forms of loss function.

Results and Discussion

While the U.S. Forest Service does make some lump-sum sales based on the appraisal volume, this is not the most common form of timber sale. In addition I would assume that sales such as these would be made in cases where a one hundred percent sample is made at the same time the trees are marked for
cutting. In other timber sale procedures, the purchase price may depend on a one hundred percent scaling of all logs harvested from the stand. This is equivalent to purchasing the stand on the basis of a one hundred percent sample, if one can assume equivalence between scaled volume of harvested logs and estimated volume from standing timber. Thus it appears that in cases where timber is purchased on the basis of volume present on the stand, this volume is often estimated by a one hundred percent sample of the stand.

The cost function and the loss functions included in this analysis are listed in Table 1. These loss functions may be divided into two large categories according to whether they describe losses to a decision maker capable of operating in the long run or losses incurred by a decision maker forced to operate on a short run basis. The majority of the loss functions that are considered appear to be most applicable to the short run situation. These include the absolute value loss function, the squared error loss function, and the five combinations of these two basic loss functions that have been discussed. In addition the set of one-sided loss functions fall in this classification. The remaining loss functions considered in this study belong to the class of loss functions applicable to long run decision making situations.

The first set of loss functions to be discussed is the set made up of the absolute value loss function, the squared error loss function, and the five two-sided loss functions that are formed by combinations of these two basic loss functions. The results of computer runs to determine the level of precision (i.e., sample size) that minimizes the opportunity loss function for each general form of loss function in this set are presented in
Table 1. Notation of cost and loss functions

C = total salary paid the crew in dollars per hour

T₁ = time in minutes required to travel a chain between plots

T₂ = time in minutes required to measure a 1/5-acre plot

n = sample size

Cost = \( C(T_1(n + 1)20 + T_2n)/60 \) \( 2 \leq n \leq 20 \)

= \( C(4.20T_1 + T_2n)/60 \) \( 20 < n \leq 200 \)

Pr = stumpage price in dollars per thousand board feet

\( \theta \) = true stand volume in board feet

z = observed stand volume in board feet

\( L_1 = |z - \theta|/Pr \) \( \theta \leq z \leq \infty \)

\( L_2 = (Prz - Pr\theta)^2 \) \( \theta \leq z \leq \infty \)

\( L_3 = |z - \theta|/Pr \) \( \theta \leq z \leq \theta \)

= \( (Prz - Pr\theta)^2 \) \( \theta < z \leq \infty \)

\( L_4 = |z - \theta|/Pr \) \( (9/10)\theta \leq z \leq (11/10)\theta \)

= \( (Prz - Pr\theta)^2 \) \( z < (9/10)\theta \) or \( z > (11/10)\theta \)

\( L_5 = |z - \theta|/1.2Pr \) \( (9/10)\theta \leq z \leq (11/10)\theta \)

= \( (1.2Prz - 1.2Pr\theta)^2 \) \( z < (9/10)\theta \) or \( z > (11/10)\theta \)

\( L_6 = |z - \theta|/Pr \) \( (99/100)\theta \leq z \leq (101/100)\theta \)

= \( (Prz - Pr\theta)^2 \) \( z < (99/100)\theta \) or \( z > (101/100)\theta \)

\( L_7 = |z - \theta|/1.2Pr \) \( (99/100)\theta \leq z \leq (101/100)\theta \)

= \( (1.2Prz - 1.2Pr\theta)^2 \) \( z < (99/100)\theta \) or \( z > (101/100)\theta \)

\( L_8 = |z - \theta|/Pr \) \( z \geq \theta \)

\( = C \) otherwise

\( L_9 = (Prz - Pr\theta)^2 \) \( z \geq \theta \)

\( = 0 \) otherwise
Table 1. (Continued)

\begin{align*}
L_{10} &= |z - \theta|/Pr \quad \theta \leq z \leq (11/10)\theta \\
&= (Prz - Pr\theta)^2 \quad z > (11/10)\theta \\
&= 0 \quad \text{otherwise}
\end{align*}

\begin{align*}
L_{11} &= |z - \theta|/1.2Pr \quad \theta \leq z \leq (11/10)\theta \\
&= (1.2Prz - 1.2Pr\theta)^2 \quad z > (11/10)\theta \\
&= 0 \quad \text{otherwise}
\end{align*}

\begin{align*}
L_{12} &= |z - \theta|/Pr \quad \theta \leq z \leq (101/100)\theta \\
&= (Prz - Pr\theta)^2 \quad z > (101/100)\theta \\
&= 0 \quad \text{otherwise}
\end{align*}

\begin{align*}
L_{13} &= |z - \theta|/1.2Pr \quad \theta \leq z \leq (101/100)\theta \\
&= (1.2Prz - 1.2Pr\theta)^2 \quad z > (101/100)\theta \\
&= 0 \quad \text{otherwise}
\end{align*}

\begin{align*}
C_B &= \text{cost associated with being forced out of business} \\
Er &= \text{limit of acceptable loss}
\end{align*}

\begin{align*}
L_{14} &= C_B P(z > Er) \quad z > Er \\
&= 0 \quad \text{otherwise}
\end{align*}
Tables 2 and 3. Both the optimal sample size and the minimum value of the opportunity loss function are recorded.

When this set of loss functions is applied to the original population of 200, 1/5-acre plots of old growth Douglas-fir with the original time parameters presented by Johnson and Hixon (1952) and with a stumpage price of $59.90 per thousand board feet, the computer analysis indicates that it is necessary to take a one hundred percent sample of the stand in order to minimize the opportunity loss function. This result holds for all loss functions included in this set. Since I have hypothesized that present practice may require a one hundred percent sample in this forest management decision making situation, this set of loss functions may quite closely approximate the actual form of loss function utilized in establishing this requirement.

When the stumpage price is lowered to $22.40 per thousand board feet, the analysis still indicates that a one hundred percent sample is necessary for each of the seven loss functions being considered in this set. In each case where a one hundred percent sample is indicated the minimum opportunity loss is $482.56, the cost of obtaining a complete enumeration of the population. The second term of the opportunity loss function is zero because the total stand volume is known without error (i.e., sampling error) and thus, the expected loss of making a decision based on this information is zero.

The modified population obtained by dividing each observed volume by ten responds only slightly differently to analysis with this set of loss functions. With a stumpage price of twenty dollars per thousand board foot and the original time parameters, analysis indicates that a one hundred percent sample is necessary. Reducing the stumpage price to ten dollars per
Table 2. Analysis of two-sided loss functions for populations of 200, 1/5-acre plots using the original time parameters

<table>
<thead>
<tr>
<th>Population</th>
<th>$L_1$</th>
<th>$L_2$</th>
<th>$L_3$</th>
<th>$L_4$</th>
<th>$L_5$</th>
<th>$L_6$</th>
<th>$L_7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda = 13,263$</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
</tr>
<tr>
<td>Pr = 559.90</td>
<td>482.56</td>
<td>482.56</td>
<td>482.56</td>
<td>482.56</td>
<td>482.56</td>
<td>482.56</td>
<td>482.56</td>
</tr>
<tr>
<td>$\lambda = 13,263$</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
</tr>
<tr>
<td>Pr = 522.40</td>
<td>482.56</td>
<td>482.56</td>
<td>482.56</td>
<td>482.56</td>
<td>482.56</td>
<td>482.56</td>
<td>482.56</td>
</tr>
<tr>
<td>$\lambda = 1,826$</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
</tr>
<tr>
<td>Pr = 320.00</td>
<td>482.56</td>
<td>482.56</td>
<td>482.56</td>
<td>482.56</td>
<td>482.56</td>
<td>482.56</td>
<td>482.56</td>
</tr>
<tr>
<td>$\lambda = 1,826$</td>
<td>n=74</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
</tr>
<tr>
<td>Pr = 310.00</td>
<td>429.69</td>
<td>482.56</td>
<td>482.56</td>
<td>482.56</td>
<td>482.56</td>
<td>482.56</td>
<td>482.56</td>
</tr>
<tr>
<td>$\lambda = 1,826$</td>
<td>n=7</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
</tr>
<tr>
<td>Pr = 52.00</td>
<td>194.11</td>
<td>482.56</td>
<td>482.56</td>
<td>482.56</td>
<td>482.56</td>
<td>482.56</td>
<td>482.56</td>
</tr>
<tr>
<td>$\lambda = 183$</td>
<td>n=7</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
</tr>
<tr>
<td>Pr = 52.00</td>
<td>194.11</td>
<td>482.56</td>
<td>482.56</td>
<td>482.56</td>
<td>482.56</td>
<td>482.56</td>
<td>482.56</td>
</tr>
<tr>
<td>$\lambda = 183$</td>
<td>n=4</td>
<td>n=163</td>
<td>n=177</td>
<td>n=124</td>
<td>n=131</td>
<td>n=167</td>
<td>n=200</td>
</tr>
<tr>
<td>Pr = $10.00</td>
<td>125.85</td>
<td>471.13</td>
<td>428.37</td>
<td>422.10</td>
<td>432.43</td>
<td>482.36</td>
<td>482.56</td>
</tr>
<tr>
<td>$\lambda = 133$</td>
<td>n=2</td>
<td>n=14</td>
<td>n=10</td>
<td>n=14</td>
<td>n=36</td>
<td>n=14</td>
<td>n=36</td>
</tr>
<tr>
<td>Pr = $2.00</td>
<td>48.81</td>
<td>248.67</td>
<td>184.95</td>
<td>244.95</td>
<td>254.84</td>
<td>268.89</td>
<td>251.47</td>
</tr>
</tbody>
</table>
Table 3. Analysis of two-sided loss functions for populations of 200, 1/5-acre plots using the modified time parameters

<table>
<thead>
<tr>
<th>Population</th>
<th>( L_1 )</th>
<th>( L_2 )</th>
<th>( L_3 )</th>
<th>( L_4 )</th>
<th>( L_5 )</th>
<th>( L_6 )</th>
<th>( L_7 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \kappa = 18,263 )</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
</tr>
<tr>
<td>Pr = $359.90</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
</tr>
<tr>
<td>( \kappa = 18,263 )</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
</tr>
<tr>
<td>Pr = $22.40</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
</tr>
<tr>
<td>( \kappa = 1,826 )</td>
<td>n=91</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
</tr>
<tr>
<td>Pr = $20.00</td>
<td>830.50</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
</tr>
<tr>
<td>( \kappa = 1,826 )</td>
<td>n=52</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
</tr>
<tr>
<td>Pr = $10.00</td>
<td>670.00</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
</tr>
<tr>
<td>( \kappa = 1,826 )</td>
<td>n=4</td>
<td>n=200</td>
<td>n=187</td>
<td>n=162</td>
<td>n=170</td>
<td>n=200</td>
<td>n=200</td>
</tr>
<tr>
<td>Pr = $2.00</td>
<td>255.40</td>
<td>856.00</td>
<td>855.50</td>
<td>841.80</td>
<td>854.70</td>
<td>856.00</td>
<td>856.00</td>
</tr>
<tr>
<td>( \kappa = 183 )</td>
<td>n=3</td>
<td>n=200</td>
<td>n=187</td>
<td>n=162</td>
<td>n=170</td>
<td>n=200</td>
<td>n=200</td>
</tr>
<tr>
<td>Pr = $20.00</td>
<td>255.40</td>
<td>856.00</td>
<td>855.50</td>
<td>841.10</td>
<td>854.00</td>
<td>856.00</td>
<td>856.00</td>
</tr>
<tr>
<td>( \kappa = 183 )</td>
<td>n=3</td>
<td>n=129</td>
<td>n=92</td>
<td>n=106</td>
<td>n=113</td>
<td>n=129</td>
<td>n=142</td>
</tr>
<tr>
<td>Pr = $10.00</td>
<td>167.30</td>
<td>790.80</td>
<td>708.90</td>
<td>712.90</td>
<td>729.80</td>
<td>803.30</td>
<td>826.30</td>
</tr>
<tr>
<td>( \kappa = 183 )</td>
<td>n=2</td>
<td>n=10</td>
<td>n=7</td>
<td>n=10</td>
<td>n=11</td>
<td>n=10</td>
<td>n=11</td>
</tr>
<tr>
<td>Pr = $2.00</td>
<td>76.70</td>
<td>366.90</td>
<td>272.00</td>
<td>363.00</td>
<td>394.50</td>
<td>367.10</td>
<td>399.80</td>
</tr>
</tbody>
</table>
thousand board feet or two dollars per thousand board feet results in a change in the optimum sample size only for the absolute value loss function. Analysis indicates that the remaining six loss functions still require a one hundred percent sample.

Use of the absolute value loss function in the analysis with the stumpage price set at ten dollars per thousand board feet results in a minimum opportunity loss of $429.69 occurring at an optimal sample size of seventy-four plots. This represents a sampling fraction of thirty-seven percent. The shape of the opportunity loss function is rather flat in the region about this optimum. The range of sample sizes from fifty-nine to ninety-four represents the interval within which the value of the opportunity loss function varies from the optimum value by less than one percent of the optimum. A similar interval using ten percent deviations as limits includes the range of sample sizes from thirty-four to 156.

Reducing the stumpage price to two dollars results in the minimum opportunity loss of $194.11 occurring at a sample size of seven plots when the absolute value loss function is applied. The shape of the opportunity loss function is not as flat about the optimum as is the case with stumpage price of ten dollars. However, the width of the interval around the optimum sample size, expressed as a percent of the optimum, is only slightly smaller.

The second modified population, developed by dividing each observed volume by one hundred, responds quite differently to analysis. As would be expected from consideration of the nature by which this population is developed, the results of analysis with stumpage price set at twenty dollars per thousand board feet are equivalent to the results obtained from the
analysis of the first modified population with stumpage price set at two dollars per thousand board feet. In the determination of expected loss, this increase in stumpage price from two dollars to twenty dollars is cancelled by the decrease in the value of the population variance by a power of one hundred.

When the stumpage price is lowered to ten dollars per thousand board feet the optimum sample size is lowered to four plots. Further reduction of stumpage price to two dollars per thousand board feet results in the minimum opportunity loss occurring at a sample size of two plots for the absolute value loss function. These results of analysis with the absolute value loss function seem to imply that for low value timber where this form of loss function might be considered applicable, very little sample survey information is required for decision making. The value of the timber is not great enough to justify any more than an estimate of volume of relatively low precision.

At a stumpage price of ten dollars per thousand board feet, the remaining six loss functions fall into two subclasses. The squared error loss and the combination loss functions, designated as $L_6$ and $L_7$, have a minimum opportunity loss at sample sizes falling in the range of eighty-one percent sample to one hundred percent sample. The shape of the opportunity loss functions is very flat. As was indicated earlier, the cost of a one hundred percent sample is $482.56. For these three loss functions the minimum opportunity loss for the squared loss function provides the largest decrease in expected cost. This decrease is only 2.3 percent of the cost of a one hundred percent sample. The remaining three loss functions in this set have minimum opportunity loss occurring in the range of fifty-eight percent sampl-
ing to sixty-six percent sampling. Again the shapes of the opportunity loss functions are flat enough that a ten percent deviation from the minimum opportunity loss includes sampling fractions of over ninety percent. Further reduction of the stumpage price to two dollars per thousand board feet results in the minimum opportunity loss occurring at sampling fractions less than eighteen percent for all loss functions in this set.

The results of this analysis of the precision specification problem for the timber sale appraisal problem indicate that if this set of loss functions adequately describe the losses that result from making decisions based on sample survey information collected in high value, old growth Douglas-fir stands, then it is necessary to measure one hundred percent of the population. As I indicated earlier, these loss functions seem to best describe losses that occur in situations where the timber operator is forced to operate on a short run basis. Thus these results are what might be expected as necessary for relatively small scale timber operators dealing with high valued timber.

If I omit the absolute value loss function, this same conclusion may be reached for the two modified populations which are representative of more moderately valued stands. Except for the extremely low valued, below average stocked stand, the analysis implies that sampling fractions of close to one hundred percent are still required.

Table 3 presents the results of analysis of this same set of loss functions with the modified time parameters used in place of the parameters provided by Johnson and Hixon (1952). The increase in time parameters has no effect on the precision specification for the original Douglas-fir population with either of the two stumpage prices that are used.
Again the absolute value loss function appears to respond somewhat different to analysis than do the remaining six loss functions in this set. For all stumpage prices used with the two modified populations, the minimum opportunity loss for the absolute value loss function occurs at a sample size somewhat smaller than was the case when the original time parameters were used in the analysis. When one considers the effect this increase in time parameters has on the opportunity loss function, this result is readily explained. The increase in time parameters causes an increase in the cost of obtaining each observation. Thus the decrease in cost resulting from decreasing the sample size by one unit is greater than was the case when the original time parameters were used. The corresponding increase in the expected loss of making decisions based on information from one less sampling unit is unaffected by the increase in time parameters. Thus when the increased time parameters are used, the analysis must indicate that the minimum opportunity loss occurs at a sample size less than or equal to the optimal sample size when the original parameters were used. The effect of increasing the time parameters was shown analytically for the absolute value loss function and the squared error loss function in Chapter 3.

For the other six loss functions the increase in time parameters has very little effect. The changes in precision specification that do result are small decreases in the optimal sample size. This is most noticeable in the lowest density population where, for stumpage price ten dollars per thousand board feet, the average decrease in the optimal sample size is approximately fifteen percent of the total population of sampling units.

The second set of loss functions that are included in the analysis is the set of one-sided loss functions. The results of the analysis using
these loss functions are presented in Tables 4 and 5. As should be expected, the results of the analysis of this set of loss functions utilizing the original time parameters correspond closely to the results obtained from the analysis of the two-sided loss functions using the increased time parameters.

This is expected because the use of one-sided loss functions in effect merely divides the second component of the opportunity loss function in half, while leaving the cost of sampling term unchanged. The increase in time parameters in the analysis of two-sided loss functions caused the cost of sampling term of the opportunity loss function to be almost doubled while the second term of the function remains unchanged. Thus an argument similar to that used in explaining the decrease in optimum sample size associated with an increase in time parameters may be used to explain the similar result occurring from the use of one-sided loss functions.

In the analysis of this set of one-sided loss functions, when the time parameters are increased from the original set, the resulting minimum opportunity loss, as would be expected, occurs at sample sizes that are somewhat smaller than the optima for the original set of time parameters. However, the conclusions that were drawn for the set of two-sided loss functions still appear to hold for this set of one-sided loss functions.

For \( L_9, L_{10}, L_{11}, L_{12}, \) and \( L_{13} \) the optimal sampling fraction is close to one hundred percent for all populations except for the low value, lowest density stands where the optimal sampling fraction varies from approximately five percent for the lowest valued stands to approximately fifty percent sampling for more moderately valued stands. The results of analysis with the one-sided absolute value loss function, \( L_8 \), correspond so closely to
Table 4. Analysis of one-sided loss functions for populations of 200, 1/5-acre plots using the original time parameters

<table>
<thead>
<tr>
<th>Population</th>
<th>$L_8$</th>
<th>$L_9$</th>
<th>$L_{10}$</th>
<th>$L_{11}$</th>
<th>$L_{12}$</th>
<th>$L_{13}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu = 18,263$</td>
<td>$n=200$</td>
<td>$n=200$</td>
<td>$n=200$</td>
<td>$n=200$</td>
<td>$n=200$</td>
<td>$n=200$</td>
</tr>
<tr>
<td>$Pr = 59.90$</td>
<td>$482.56$</td>
<td>$482.56$</td>
<td>$482.56$</td>
<td>$482.56$</td>
<td>$482.56$</td>
<td>$482.56$</td>
</tr>
<tr>
<td>$\mu = 18,263$</td>
<td>$n=200$</td>
<td>$n=200$</td>
<td>$n=200$</td>
<td>$n=200$</td>
<td>$n=200$</td>
<td>$n=200$</td>
</tr>
<tr>
<td>$Pr = 22.40$</td>
<td>$482.56$</td>
<td>$482.56$</td>
<td>$482.56$</td>
<td>$482.56$</td>
<td>$482.56$</td>
<td>$482.56$</td>
</tr>
<tr>
<td>$\mu = 1,826$</td>
<td>$n=74$</td>
<td>$n=200$</td>
<td>$n=200$</td>
<td>$n=200$</td>
<td>$n=200$</td>
<td>$n=200$</td>
</tr>
<tr>
<td>$Pr = 20.00$</td>
<td>$429.10$</td>
<td>$482.56$</td>
<td>$482.56$</td>
<td>$482.56$</td>
<td>$482.56$</td>
<td>$482.56$</td>
</tr>
<tr>
<td>$\mu = 1,826$</td>
<td>$n=43$</td>
<td>$n=200$</td>
<td>$n=200$</td>
<td>$n=200$</td>
<td>$n=200$</td>
<td>$n=200$</td>
</tr>
<tr>
<td>$Pr = 10.00$</td>
<td>$337.20$</td>
<td>$482.56$</td>
<td>$482.56$</td>
<td>$482.56$</td>
<td>$482.56$</td>
<td>$482.56$</td>
</tr>
<tr>
<td>$\mu = 1,826$</td>
<td>$n=4$</td>
<td>$n=200$</td>
<td>$n=151$</td>
<td>$n=159$</td>
<td>$n=200$</td>
<td>$n=200$</td>
</tr>
<tr>
<td>$Pr = 2.00$</td>
<td>$125.50$</td>
<td>$482.56$</td>
<td>$460.60$</td>
<td>$469.78$</td>
<td>$482.56$</td>
<td>$482.56$</td>
</tr>
<tr>
<td>$\mu = 183$</td>
<td>$n=4$</td>
<td>$n=200$</td>
<td>$n=151$</td>
<td>$n=159$</td>
<td>$n=200$</td>
<td>$n=200$</td>
</tr>
<tr>
<td>$Pr = 20.00$</td>
<td>$125.50$</td>
<td>$482.56$</td>
<td>$460.20$</td>
<td>$469.40$</td>
<td>$482.56$</td>
<td>$482.56$</td>
</tr>
<tr>
<td>$\mu = 183$</td>
<td>$n=3$</td>
<td>$n=115$</td>
<td>$n=97$</td>
<td>$n=104$</td>
<td>$n=115$</td>
<td>$n=126$</td>
</tr>
<tr>
<td>$Pr = 10.00$</td>
<td>$81.90$</td>
<td>$423.30$</td>
<td>$377.50$</td>
<td>$388.20$</td>
<td>$429.60$</td>
<td>$445.20$</td>
</tr>
<tr>
<td>$\mu = 183$</td>
<td>$n=2$</td>
<td>$n=10$</td>
<td>$n=10$</td>
<td>$n=11$</td>
<td>$n=10$</td>
<td>$n=11$</td>
</tr>
<tr>
<td>$Pr = 2.00$</td>
<td>$37.00$</td>
<td>$179.80$</td>
<td>$177.90$</td>
<td>$193.30$</td>
<td>$179.90$</td>
<td>$196.00$</td>
</tr>
</tbody>
</table>
Table 5. Analysis of one-sided loss functions for populations of 200, 1/5-acre plots using the modified time parameters

<table>
<thead>
<tr>
<th>Population</th>
<th>$L_9$</th>
<th>$L_9$</th>
<th>$L_{10}$</th>
<th>$L_{11}$</th>
<th>$L_{12}$</th>
<th>$L_{13}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu = 18,263$</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
</tr>
<tr>
<td>Pr = 359.90</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
</tr>
<tr>
<td>$\mu = 18,263$</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
</tr>
<tr>
<td>Pr = 322.40</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
<td>356.00</td>
<td>856.00</td>
</tr>
<tr>
<td>$\mu = 18,263$</td>
<td>n=52</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
</tr>
<tr>
<td>Pr = 320.00</td>
<td>356.00</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
<td>356.00</td>
</tr>
<tr>
<td>$\mu = 18,263$</td>
<td>n=31</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
<td>n=200</td>
</tr>
<tr>
<td>Pr = 310.00</td>
<td>310.00</td>
<td>553.60</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
<td>856.00</td>
</tr>
<tr>
<td>$\mu = 18,263$</td>
<td>n=3</td>
<td>n=183</td>
<td>n=132</td>
<td>n=200</td>
<td>n=3</td>
<td>n=91</td>
</tr>
<tr>
<td>Pr = 320.00</td>
<td>320.00</td>
<td>166.90</td>
<td>852.10</td>
<td>777.30</td>
<td>793.60</td>
<td>856.00</td>
</tr>
<tr>
<td>$\mu = 18,263$</td>
<td>n=3</td>
<td>n=183</td>
<td>n=132</td>
<td>n=200</td>
<td>n=3</td>
<td>n=172</td>
</tr>
<tr>
<td>Pr = 620.00</td>
<td>620.00</td>
<td>166.90</td>
<td>852.10</td>
<td>776.90</td>
<td>792.90</td>
<td>856.00</td>
</tr>
<tr>
<td>$\mu = 18,263$</td>
<td>n=2</td>
<td>n=91</td>
<td>n=82</td>
<td>n=88</td>
<td>n=91</td>
<td>n=100</td>
</tr>
<tr>
<td>Pr = 810.00</td>
<td>810.00</td>
<td>111.60</td>
<td>702.60</td>
<td>644.40</td>
<td>660.90</td>
<td>706.50</td>
</tr>
<tr>
<td>$\mu = 18,263$</td>
<td>n=2</td>
<td>n=7</td>
<td>n=7</td>
<td>n=8</td>
<td>n=7</td>
<td>n=7</td>
</tr>
<tr>
<td>Pr = 82.00</td>
<td>82.00</td>
<td>64.90</td>
<td>265.90</td>
<td>264.10</td>
<td>287.50</td>
<td>265.90</td>
</tr>
</tbody>
</table>
the results of analysis with the two-sided absolute value loss function
that no further discussion is necessary.

In general I would conclude that if any of the loss functions in
either of those two sets were utilized in the analysis of this type of
precision specification problem, the prospective timber purchaser would be
a small scale timber purchaser forced to operate on a short run basis by
a lack of the large amounts of capital necessary for the operator to absorb
occasional losses on individual sales. In this situation the timber pur­
chaser would place a large value on knowing the true total stand volume
and thus, except for very low valued timber, the purchaser would require
almost a one hundred percent sample of the population if he was to purchase
the timber on the basis of sample survey information.

The third and final set of loss functions to be considered in this
analysis has only one member. The results of analysis with this loss
function, $L_{14}$, are recorded in Tables 6 and 7. As is stated earlier, this
loss function is a modified linear loss function which assumes that, in
the long run, losses on one sale will be balanced by future sales as long
as the loss on any given sale does not exceed some limit point beyond which
the timber operator cannot absorb the loss. This loss function is applied
only to the unmodified Douglas-fir population, with the stumpage price set
at $59.90 per thousand board feet.

When the limit of acceptable loss is set at approximately five-tenths
percent of the population mean, a sample size of about 170 plots out of a
total population of 200 plots is specified as optimal. This implies that
if this much accuracy is required, then it is almost necessary to take a
one hundred percent sample. As would be expected, as the acceptable limit
Table 6. Analysis of a long run loss function, $L_{14}$, for populations of 200, 1/5-acre plots using the original time parameters

<table>
<thead>
<tr>
<th>Loss of Being Forced Out of Business</th>
<th>Limit of Acceptable Loss (bd. ft.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>1000</td>
</tr>
<tr>
<td>$50,000</td>
<td>$n=168</td>
</tr>
<tr>
<td></td>
<td>432.59</td>
</tr>
<tr>
<td>$100,000</td>
<td>$n=169</td>
</tr>
<tr>
<td></td>
<td>434.54</td>
</tr>
<tr>
<td>$150,000</td>
<td>$n=170</td>
</tr>
<tr>
<td></td>
<td>436.61</td>
</tr>
</tbody>
</table>

Table 7. Analysis of a long run loss function, $L_{14}$, for populations of 200, 1/5-acre plots using the modified time parameters

<table>
<thead>
<tr>
<th>Loss of Being Forced Out of Business</th>
<th>Limit of Acceptable Loss (bd. ft.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>1000</td>
</tr>
<tr>
<td>$50,000</td>
<td>$n=167</td>
</tr>
<tr>
<td></td>
<td>774.44</td>
</tr>
<tr>
<td>$100,000</td>
<td>$n=168</td>
</tr>
<tr>
<td></td>
<td>777.82</td>
</tr>
<tr>
<td>$150,000</td>
<td>$n=169</td>
</tr>
<tr>
<td></td>
<td>779.64</td>
</tr>
</tbody>
</table>
of loss increases (say to approximately twenty percent of the population mean), the required sample size decreases to only five plots out of the total population of 200 plots. This represents only a two and one-half percent sample of the population.

Surprisingly, to me, the loss associated with being forced out of business has practically no effect on the specification of the optimal sample size. Doubling or even tripling this value results in at most an increase of five sampling units in the optimum. In addition, over the range of values I consider, the absolute size of this value has an almost negligible effect on the size of the minimum value of the opportunity loss function.

When the analysis is performed with the increased set of time parameters, the specification of optimal sample size remains unaffected. However, the size of the minimum opportunity loss occurring at these optimal sample sizes is considerably increased from the corresponding minimum opportunity loss values that occurred when analysis with the original set of time parameters is evaluated.

These last two facts seem to imply that at the optimal sample size the term of the cost plus loss function which expresses the cost of sampling dominates the function. A determination of the cost of sampling at the optimal sample size verifies this contention. This situation offers some explanation for the lack of effect variations in the cost of being forced out of business have on the optimal sample size. The probability of an estimate of total volume exceeding the limits of acceptable error, for this population, is so small that variations in the cost of being forced out of business are effectively masked. In other populations with larger variances this situation might not hold true.
For this loss function to be applied, it is necessary to accurately evaluate the limit of acceptable loss beyond which the timber purchaser is forced out of business. In addition, it must be assumed that the given timber purchaser is indeed able to operate on a long run basis within these limits. However, the ability of the timber purchaser to operate on a long run basis would probably be expressed in the determination of the limit of acceptable loss beyond which the purchaser is forced out of business.

In conclusion I might say that this final form of loss function could probably be used in a wide range of situations. The small scale timber purchaser who is forced to operate on a short run basis by a lack of available capital could use the loss function by setting the limit of acceptable loss close to zero (say less than five-tenths percent of the population mean). As the ability of the timber purchaser to absorb losses increases, the corresponding limits of acceptable loss are also increased. Thus the degree of ability of any given timber purchaser to operate on a long run basis is readily expressed in setting the limits of acceptable loss for this form of loss function.
CHAPTER 5. ALLOWABLE CUT DETERMINATION PROBLEM

Decision Rules

The SORAC (Short Run Allowable Cut) computer program written by Sassa-
man, Chappelle, and Fritchman (1969), provides a fully developed decision
rule that may be applied to the problem of the determination of annual
allowable cut faced by many forest managers. The program is written with
several management options available. The choice of the specific management
practices to be used in a given run is included as a portion of the input
data required for the operation of the program.

The first option available is a choice between either area or volume
control for use as the basis of the allowable cut determination. According
to Davis (1966)

"The principle of area control is very simple; it means that the
volume to be harvested is defined by timber to be removed on the area
allocated for cutting. Cutting is scheduled on a forest under manage-
ment, so that each year a certain area of timber is available for
harvest."

For volume control, however, the determination of allowable cut is
based on the volume and distribution of growing stock and on projected
increment values. The procedure used in SORAC for determination of allowable
cut by this method is very similar to the procedure proposed by Davis (1966).
The first step is to obtain an estimate of allowable annual cut. This may
be accomplished by means of one of the many formulae available, by using
the annual cut for the previous year, or even by making an intelligent guess.
A first estimate of the time required to cut a stand is obtained by making
use of the yield that could be expected from the stand if it were harvested
at its present age. This yield is calculated by use of a yield equation
with age when cut as the only independent variable. The initial estimate of allowable cut is divided into this yield to determine the time required to cut the stand. With this value, a new average age of the stand when cut is evaluated and the resulting yield is calculated. Again using the initial allowable annual cut estimate, a new estimate of time required to cut the stand is calculated. When the difference between two successive estimates of time to cut the stand falls within preset tolerance limits, the estimate is accepted and analysis proceeds to the next stand. Once all stands have been analysed, checks are made to see if the cumulated time to cut the forest is equal to the specified rotation length and if the area cut per year is reasonable. If these criteria are not met, the initial estimate of allowable annual cut is systematically changed and the procedure is repeated. Once the criteria are met, the volume-regulation allowable cut specification is completed.

As an additional option it is possible to combine these two methods of allowable cut determination by using one for some of the planning periods in the rotation and the other for the remaining periods. It is also possible to include the effects of expected management and environmental changes in the allowable cut decision process within a rotation by specifying individual yield functions for each stand and for each planning period. A final method by which changes in management strategy and utilization trends may be introduced into the allowable cut determination is through the specification of stumpage prices. These prices may be unique for each stand and each planning period or may be constants over any period up to a complete rotation.
For the SORAC program to be used as the decision rule for this study, it is necessary to make some modifications and additions to the program. As I indicate above, the yield equation used in SORAC to predict volume yield has as its only independent variable, average age when the stand is cut.

Yield equations used in forest management situations frequently include other independent variables which measure in some way the variations in productivity of individual stands. Examples of such variables are site index, average stand diameter, or stand basal area. Since SORAC provides for the use of individual yield equations for each stand in the population, variables indicating the variation in productivity of individual stands need not be included in the yield equations.

For the purposes of this study, it is desirable to use a yield equation that has as its independent variables, measures of productivity which must be estimated by use of sample survey information. This is not always the case with the variable, average age of stand when cut. For a stand that has been under management since its origination, the average age of the stand is known without any sample survey information.

An equation form such as that proposed by Myers (1967, 1968) and by Myers and Godsey (1968) would more nearly fulfill this need. These equations predict stand volume as a function of average stand d.b.h. and the product of stand basal area and the average height of dominant and codominant trees on the stand. These independent variables would be evaluated from sample survey information. Thus, it would be necessary to determine the level of precision required in the estimation of these variables for use in the decision making process of allowable cut determination.
While Myers' equations are applicable specifically to ponderosa pine and lodgepole pine, similar yield equations are available for the Douglas-fir population that I am dealing with. Since the allowable cut determination involves making decisions over time, it is also necessary to add to the SORAC program a subroutine that expresses the growth or change of the independent variables of the yield equation as a function of time.

The actual changes that have been made in the SORAC program are all a result of the conversion of the yield equation used in the program from one that is a function of age to a yield equation that is a function of diameter. I have worked previously with the Hartman Douglas-fir population which is described by Schreuder, Sedransk, and Ware (1968). Thus the yield equation I have inserted in SORAC is one that was developed by regression analysis for this population. The general form of this equation is

\[ \text{Volume} = a + b(d.b.h.)^2 + c(d.b.h.)^4 \]

The use of this yield equation requires the addition of a subroutine that will simulate diameter growth over time. Such a subroutine has been written, based on a diameter-age relationship described in tabular form by McArdle, Meyer, and Bruce (1961) for Douglas-fir stands. This relationship was not developed from the Hartman Douglas-fir population or from the arbitrary population to be used in the analysis for this study. Thus it probably does not perfectly describe the diameter growth of the population under analysis. However, since diameter growth is strictly a prediction of future events, no model can be expected to predict diameter growth with certainty. For the purposes of this study, this diameter-age relationship adequately describes the growth in diameter of the Douglas-fir stands under analysis.
Population Description

The population used in this study is purely hypothetical. It is developed in a form to meet the input demands of the SORAC computer program. It consists of two, 40,000-acre Douglas-fir stands, one with an average age of 200 years and an average diameter of 40.3 inches d.b.h. and the second with an average age of 190 years and an average diameter of 39.1 inches d.b.h.

In order to determine the expected loss that results when the sample survey is conducted at any specified level of precision, it is necessary to know the variance that exists among individual diameters in a stand. For simplicity I assume that the variances for the two stands are equal. The numerical value used in this study is $S^2 = 166,7455$. This value is obtained by use of the standard variance formula

$$S^2 = \frac{\sum_{i=1}^{N} (x_i - \bar{x})^2}{N - 1}$$

applied to a population of 126 trees in a predominantly Douglas-fir stand. The data on this stand were collected as part of a 3-P sampling procedure.

While the population used in this study is hypothetical, it is desirable for it to be similar to actual populations that might exist in nature. Therefore, an attempt has been made to specify population parameters which are representative of actual populations. The average diameters specified above for the two stands are justified as reasonable by an extension of the diameter-age relationship described by McArdle, Meyer, and Bruce (1961).

The small population used to evaluate the variance that exists among individual diameters in a stand has an average diameter only slightly less than the two stands being utilized in this analysis. Since I have assumed
that variances for the two stands are equal, it is not unreasonable to further assume that their variance is equal to the variance of this small, completely enumerated population. Further evidence of the "reasonableness" of this value for the variance lies in the fact that the Hartman Douglas-fir population (size $N = 579$) has a variance among individual diameters only slightly less than the value used in this study.

Procedures

The objective of this analysis is to determine the optimum precision level at which a sample survey should be run to provide the necessary information for allowable cut determination. The actual determination of this desired level of precision for the sample survey is made by the approach suggested by Cochran (1965). This method of analysis is the minimization of a cost plus loss function.

The loss function utilized in this phase of the study is made up of two components. The first of these components measures the value of the timber cut under inaccurate estimates of average stand diameter relative to the value of the timber that would have been cut if perfect information were supplied as input to SORAC. The form of this component of the loss function is that of a straight linear loss function.

The term, value of timber cut, refers to the total value of timber harvested over the entire rotation. Thus, this value is dependent on the schedule of allowable cuts for the rotation provided as output by SCRAC. Three alternative measures of this value of timber harvested are considered. For the first alternative, the value of timber scheduled to be harvested in each planning period is determined and the total value of timber harvested
is the sum of the values for each planning period. For the remaining two alternatives, the value of timber harvested in each planning period is determined as in the first alternative. However, the total value of timber harvested over the rotation is determined by discounting the value for each planning period back to the present. The difference between these last two alternatives is in the discount rate used. These three methods of determining the value of timber harvested will be discussed further after the operation of SORAC is more fully described.

Because of the form of this component of the loss function and the nature of the SORAC program, this component of loss can be either positive or negative. A diameter specification which results in a total harvest over a rotation of a greater value than that expected from perfect information would clearly be viewed as a gain or negative loss by the decision maker. Similarly a diameter specification which results in a total harvest of value smaller than that resulting from perfect information is viewed as a loss by the decision maker. In most cases a negative loss is more than absorbed by increases in the second component of loss.

In addition to the possible losses in monetary value of timber cut, variations in cutting patterns resulting from inaccurate diameter estimation will also affect the final makeup of the new stand that will exist at the end of the first rotation. If the decision maker is utilizing the volume regulation option of SORAC, it might be assumed that one of his management goals is to cut the available timber in such a manner that at the completion of the first rotation the age-area distribution of the timber will be such that equal volumes of timber will be available for harvest each year or period of years in the following rotation.
The nature of SORAC is such that this ideal status will not be reached at the end of the first rotation. In fact, the age-area distributions that result following each successive rotation will only asymptotically converge to this ideal status. If the primary management goal during the first rotation is to obtain a perfect age-area distribution of timber regardless of the volume cut each year of the rotation, then a strict area control approach to determining allowable cut would be more reasonable. Usually the forest management decision maker is interested in maintaining a reasonably equal flow of volume during the first rotation as well as in developing a desirable age-area distribution. Thus, volume control or some combination of area and volume control can usually be justified.

Therefore, the second component of the loss function measures the cost that results when the composition of the forest following the completion of the first rotation varies from a perfect age-area distribution. This perfect age-area distribution exists when the forest is in the completely managed status mentioned above (i.e., when the total area in the forest is equally divided among the age classes that exist following the completion of the first rotation).

This definition of a forest in a completely managed status implies the acceptance of an assumption concerning equal productivity of each acre in the forest. In nature this assumption is almost never valid. Thus a more correct definition of a forest in a completely managed status is that given previously in this discussion. By this definition the forest reaches a completely managed status when the age-area distribution is such that equal volumes of timber become available for harvest each year or period of years. This same assumption of equal productivity is also implied when a
single diameter-age relationship is used to estimate diameter growth over the entire forest. Although this assumption is unlikely to be valid in nature, for the purposes of this study it is only necessary to make some arbitrary assumption about the variability of productivity over the forest. Therefore, an assumption of equal volume productivity per unit area over the entire forest is acceptable.

The measurement of this component of loss is by a loss function with the general form of an absolute value type loss function. There exists an area, A, which represents the acreage in each age class when this perfect age-area distribution is achieved. Any deviation from A in any age class represents a loss to the decision maker. I have evaluated this loss by determining the value of the timber that would be present on an acre of land at harvest age. Each acre of deviation from A in any age class represents a unit of loss. The loss for each age class is determined by multiplying the acres of deviation from A by the value of the timber on an acre at harvest age. The sum of these losses over all age classes is the second component of the loss function.

The method used to measure this component of loss needs to be discussed further in order to provide some justification for its use. This can perhaps best be done by giving an example of a specific situation to which the allowable cut problem is applied and by indicating the nature of the loss that is expressed by this component of the loss function.

A good example is developed when one assumes that this population of two, 40,000-acre stands is owned by a private forestry firm which plans to develop a mill which will depend on this population to provide the timber necessary to keep it running at full capacity. In this situation it is
desirable to have equal volumes of timber become available for harvest yearly. Since I have assumed equal productivity per acre over the entire forest, this is equivalent to requiring equal acres of harvestable timber yearly.

The acres of timber that become available for harvest in any year or period of years may fall in any one of three possible classes. The simplest case to consider occurs when the acres of timber that become available for harvest in any year or period of years are equal to the acres that should be available if the stand is in a perfectly managed status. When this is true, this component of loss is zero.

The second case to be considered occurs when the acres of timber available for harvest are less than those that should be available if the stand is in a perfectly managed status. When this occurs the firm is faced with two alternatives. They may either operate the mill at less than full capacity or they may purchase from other sources the additional timber necessary to operate the mill at full capacity. In either case a loss proportional to the reduction in timber available for harvest is incurred by the firm. If the mill is run at less than full capacity the loss is made up of potential output by the mill and by the possible reductions in efficiency of mill operations that result from running at less than full capacity. If additional timber is purchased, the loss is made up of the costs of setting up purchases of timber and the difference between the price paid to outside sources and the price assumed to be paid for the firm's own timber.

The final possible case occurs when the timber available for harvest is greater than the amount which should be available if the forest is in a perfectly managed status. Since the mill is designed to run at full capacity
with the volume of timber that should be available if the forest were in a perfectly managed status, this additional timber that becomes available for harvest cannot be utilized. If one can assume that the rotation age has been set so that timber is harvested at that point when the net return to management is optimal, then holding timber beyond the rotation age results in a loss that is proportional to the amount of timber uncut at the optimal age.

In this study I have considered the loss in the two cases where loss occurs to be equal to one hundred percent of the value of the excess or shortage of timber available for harvest. In fact, this proportion should be something less than one hundred percent. By looking at the most severe possible measure for this component of loss, the results of analysis should result in the maximum possible specified sample size. By reducing these proportions, the optimal sample size should be reduced. To illustrate this point the analysis was also run with this component of loss set at a more realistic value of forty percent of the value of the excess or shortage of timber available for harvest in each planning period.

The final component of the cost plus loss function is the function that expresses the cost of sampling. The cost of sampling function that I will use in this study is a two-term function. The first term of the function measures the cost of measuring the sample plot. The second term measures the cost of traveling between plots. Jessen (1942) suggests that a function of the form

\[ D = d/n \]

estimates the total distance traveled between randomly located plots. Hazard (1969) has used this form of function to describe the distance
traveled between randomly located sample survey plots in west coast Douglas-fir.

The cost coefficients for the cost function are again evaluated from data presented by Johnson and Hixon (1952). The time study conducted by these authors indicated that the average time per chain of travel between plots in old-growth Douglas-fir is 1.84 minutes. The results of a second time study showed that circular one-quarter acre plots require 10.6 minutes for measurement. Each of these times is for the field work of a two-man crew. Travel between plots is guided by the use of staff compass and trailer tape. It is assumed that personnel used to run this type of sample survey would be paid at a rate of approximately three dollars per hour.

The cost of sampling function that results from these components is of the form

\[ C_n = 6(10.6n + 1.84d/Vn)/60 \]

where \( C_n \) is the total cost of sampling, \( n \) is the sample size, and \( d \) is a constant of proportionality as defined by Jessen (1942), Hansen, Hurwitz, and Kadow (1953), and Hazard (1969). For a large west coast Douglas-fir population included in the Forest Survey, Hazard (1969) set \( d \) equal to 146.98 miles. However, a simple example quickly shows this value to be inappropriate for the population being considered in this study. Each of the two, 40,000-acre stands is to be sampled independently. Thus the population to be sampled is actually only 62.5 square miles in size. When the constant of proportionality suggested by Hazard is utilized for this population, the estimate of total distance traveled between four randomly located plots is

\[ D = 146.98/4 = 294 \text{ miles} \]
This is clearly an unreasonable result.

Hansen, Hurwitz, and Madow (1955) suggest that

\[ d_1 = \sqrt{A/n} \]

where \( A \) is the total area in the population to be sampled, \( n \) is the sample size, and \( d_1 \) is the average distance between sample points. In this terminology

\[ D = d_1 n = (\sqrt{A/n})n = \sqrt{A}Vn \]

Thus the constant of proportionality, \( d \), is equal to the square root of the area of the population to be sampled. For this 40,000-acre population the appropriate constant of proportionality would then be

\[ d = \sqrt{62.5} = 7.9057 \text{ miles} \]

Since the time coefficients are in terms of minutes per chain, \( d \) must be converted to chains. Thus, for this study,

\[ d = 7.9057(80) = 632.456 \text{ chains} \]

The actual data used in the determination of required precision is generated by running the SORAC computer program over a specified range of diameter estimates for each of the two stands in the population. The range of diameter estimates for each stand is bounded by the true average stand diameter plus and minus three inches. These diameter estimates serve as midpoints for eleven subintervals. Nine of these intervals are one-half inch in width while the final two are one and one-half inches wide. The nine, half-inch intervals are arranged symmetrically around the true average stand diameter. The two, inch and one-half width intervals occur at the extremes of the range. This combination of intervals for the two stands requires 121 runs of the SORAC program (i.e., one run for each combination of diameter estimates for the two stands).
Each run of SORAC yields the two basic forms of information needed to evaluate the loss function. These two forms of information are the cutting schedule for a complete rotation and the structure of the age-area distribution for the forest at the completion of the first rotation.

For this study I am using a one hundred-year rotation with ten-year planning periods. Planning periods are defined as that interval of time in which parameters of the planning model do not change in value (Sassaman, Chappelle, and Fritchman, 1969). This closely corresponds to the interval between regular management plan revisions.

The SORAC program allows for considerable flexibility in the management scheme that is to be applied to the forest. For this study a specified scheme will be followed throughout in order to control management practices as a possible source of variation. The lengths of the rotation and of the planning periods have already been defined. The yield equation coefficients will remain constant throughout the rotation. A regeneration period of five years is assumed. Since all stands will be extensively managed, it is only necessary to use a single empirical yield equation. Finally, volume control is used for the first five planning periods and area control is used for the remaining five planning periods. This specification has the effect of assuring that the majority of the forest will be cut in the first rotation. These assumptions satisfy the requirements of SORAC and thus define the management scheme that will be assumed throughout this study.

The determination of the value of the timber harvested under the allowable cut allocation requires some further discussion. The SORAC program determines the value of the timber harvested from each stand. This value is based on the average age of the stand when it is cut. I am interested,
however, in the value of timber actually cut in each planning period. In order to obtain an accurate estimate of this value, it is necessary to perform additional calculations on the raw data provided by a run of SCRAC.

The procedure for evaluation varies depending on whether the harvest for a given planning period comes from one stand or from two stands. If the harvest is entirely from one stand, the procedure is fairly straightforward. The average age of the timber harvested is taken as the age of the stand at the midpoint of the planning period. This average age is inserted in the diameter growth model to determine the diameter growth that has occurred since the beginning of the rotation and thus, the average diameter of the timber being cut. The average yield per acre of the timber harvested is then readily calculated.

This estimate of average yield per acre of timber harvested in a given planning period is more accurate than the estimate provided by the SORAC program. This is true because the SORAC program evaluates average yield per acre on the basis of the average age when harvested for all timber left uncut in the stand at the beginning of the given planning period, rather than on the basis of the average age of timber actually harvested in the given planning period. This procedure is especially inaccurate for those cases where several planning periods are required to harvest a stand. For these cases the SORAC estimate of yield per acre for the first planning periods would clearly be an overestimate. As the harvest in each planning period is completed, the estimate of yield per acre for the following planning period overestimates the actual yield per acre by a smaller amount. However, only for the final planning period needed to harvest a given stand will the SORAC estimate of yield per acre not be an overestimate.
Once the estimate of average yield per acre is obtained as described above, an estimate of timber harvested is readily determined by multiplying the yield per acre for a planning period by the number of acres specified to be cut in that planning period. The value of timber harvested for a planning period is then estimated by multiplying this result by a value coefficient for the timber. For this study a constant value of fifty dollars per thousand board feet is assumed. The sum of these values for each planning period in the rotation yields the total value of timber cut for the rotation.

If the harvest for a given planning period comes from two stands, the acreage cut in each stand is obtained from the SORAC program. The above procedure for analysis is then carried out separately for the timber cut in each stand. The only modification to this is that the average age of timber cut in each stand is taken as the age of the stand at the midpoint of the planning period for each stand.

Two alternative approaches are also considered as ways of determining the value of the timber harvested. Both of the alternatives use the idea of applying a discount rate in order to determine the value of the timber harvested over a rotation in terms of present day worth. This has some appeal since sampling costs will have to be incurred now rather than being spread out over an entire rotation. The first alternative uses a discount rate of two and one-half percent. The second uses five percent. These values were arbitrarily selected to examine a range of values. In addition, I have investigated the effect of applying these discount rates to the component of loss measuring deviations from a perfect age-area distribution.
For any given sample size, the probability of occurrence for any combination of diameter estimates is readily determined by use of the Central Limit Theorem and the variance estimate for diameters in a stand. The expected value of the loss corresponding to any given sample size is then

\[ E(\text{loss } n) = \sum_{i} \sum_{j} P_{ij} L_{ij} \]

where \( P_{ij} \) is the joint probability of the diameter estimate for the first stand falling in interval \( i \) and the diameter estimate for the second stand falling in interval \( j \) and \( L_{ij} \) is the loss resulting from making decisions based on this combination of diameter estimates.

I have defined the population as two, 40,000-acre stands. Therefore, if independent samples are taken in each stand to estimate average stand diameter, the distribution of possible diameter estimates from one stand should be independent of the distribution of possible diameter estimates from the second stand. Thus, \( P_{ij} \) is actually just the product of the independent measures of probability for diameter estimates in each stand.

A short computer program was written to determine the expected value of the cost plus loss function for any specified sample size. The sample size that results in the minimum value of the cost plus loss function will be the sample size that corresponds to the optimal precision level.

Results and Discussion

The major results of this study may be summarized very quickly. For the most severe loss function (i.e., that where the second component of loss is set equal to the value of the excess or shortage of timber available for harvest in each planning period), the optimal sample size for the undiscounted data is 750, one-quarter-acre circular plots for each of the
40,000-acre stands. The effect of applying a discount rate in the determination of the value of timber harvested is almost undistinguishable. The optimal specification for the data discounted at two and one-half percent is 744 plots. When a five percent discount rate is applied, the optimal specification is 741 plots.

Very little variation exists between allowable cut specifications for combinations of diameter estimates close to the true stand diameters. Thus it would appear that at these sample sizes the expected loss is so dominated by the losses associated with these combinations of diameter estimates close to the true stand diameters that the effect of the use of discount rates in the determination of the value of timber harvested is masked.

When the second component of the loss function is altered to include only forty percent of the value of the excess or shortage of timber available for harvest in each planning period of the second rotation, the precision specifications are reduced considerably. The optimal sample size for the undiscounted data is 342 plots. When a discount rate of two and one-half percent is applied, the optimal sample size is reduced to 306 plots. The use of a five percent discount rate results in an optimal sample size of 168 plots.

The optimal sample size for the more severe loss function corresponds to sampling fractions of about one-half percent. For the more realistic second loss function the sampling fraction for the two larger optimal sample sizes is approximately one-fifth percent while the smaller sample size represents a one-tenth percent sample.
All of these specifications appear to be rather low. However, readings in Davis (1966) and the form of the original yield equation in SORAC lead me to believe that the decision rule would be rather insensitive to errors in diameter estimation. Davis pointed out that in any volume control method of allowable cut determination, the reliability of the allowable cut specification is dependent on the accuracy of both the volume and increment data used. In addition to the estimation of the independent variables to be used in the yield equation, inaccuracies are also introduced in the development of the yield equation being utilized and in the development of the equation that is used to express increment in the independent variables over time.

These factors are undoubtedly some of the major reasons that usual forest management practice calls for a re-evaluation of allowable cut at the beginning of each planning period. This allows for the inclusion of any change in management practices in the allowable cut specification. In addition it provides a means of correcting any obvious errors in the allowable cut specification rather than allowing them to be carried through the entire rotation as is the case in the specific management situation used in this study.

The results of analysis using discount rates in the determination of the value of timber harvested over a rotation can probably be explained by the fact that the application of these discount rates decreases the absolute size of the loss associated with each of the 121 possible combinations of diameter estimates. Thus the decrease in expected loss associated with the addition of each additional sampling unit is smaller than is the case when discount rates are not used. Since the cost of sampling component of the
cost plus loss function remains unaffected by the application of discount rates, the precision level (i.e., sample size) at which the minimum cost plus loss occurs will be reduced. This statement may be clarified by pointing out again that the minimum cost plus loss occurs at that sample size where the increase in cost that results from adding the last sampling unit equals the decrease in expected loss that results from the added unit of information.

The decrease in the specified sample size from 342 plots to 306 plots is practically insignificant when one remembers that the population being sampled is a 40,000-acre stand. On this basis the decrease is only a decrease of about two-hundredths percent in the portion of the population being sampled. This fact leads me to offer another possible result of using discount rates that may also have some effect on the precision specification.

The application of discount rates to the determination of the value of timber harvested results in the last three or four planning periods of a one hundred-year rotation contributing very little to the total value. This has the effect of magnifying variations in timber harvested near the beginning of the rotation. This could have the effect of increasing the decrease in expected loss that results from adding an additional unit of information. This would result in the minimum cost plus loss occurring at a larger sample size. It appears that the actual effect of the use of discount rates is probably some combination of these two effects that results, for this population and for this decision rule, in a slight overall decrease in the specification of the optimal sample size.
When discount rates are applied to the component of loss measuring deviations from a perfect age-area distribution, the resulting evaluation of the loss function yields total losses that are negative for the majority of the combinations of possible diameter estimates. This leads to two possible conclusions. The first is that by the use of incorrect estimates of the diameters of the stands, the SORAC program may prescribe a course of action that is more profitable than if the correct estimates are used. This would imply that the decision rule is invalid. The second conclusion is that the measure of loss is invalid. Since the procedures of allowable cut determination used by SORAC have been in use for some time, I am forced to accept this second conclusion. More specifically, I assume that the addition of a discount rate to the previously described procedure for evaluating that component of loss dealing with deviations from a perfect age-area distribution of the forest results in an invalid measure of this component of loss.
CHAPTER 6. SUMMARY

The objectives of this study are:

(1) to investigate the levels of precision required for two fully specified forest management decisions.

(2) to suggest a general approach to be followed in specifying precision requirements for any forest management situation.

Chapters 4 and 5 deal with a discussion of the first objective. A brief review of the results of the analyses performed in these two chapters should serve as an adequate introduction to a discussion of the second stated objective.

The most evident result of the analysis is that precision specification is heavily dependent on the forest management situation being considered, on the decision rule being applied, and on the particular population being studied. This strong dependence is in close agreement with what was to be anticipated.

In the planning stage of any sample survey it is necessary to fully specify the proposed uses to be made of the sample survey information. A complete specification of both the management situation of interest and the decision rule being utilized provides a major portion of this definition of proposed uses. Since each management situation and the associated decision rule define a unique use or set of uses for sample survey information, it is not unexpected that the precision specification for the two management situations and decision rules investigated here are quite different.

In this study the forest management situation of interest and the decision rule being utilized are closely related. For each of the two manage-
ment situations considered, only a single decision rule is utilized. Thus the combined effect of the management situation and of the decision rule cannot be broken down into the effect attributed to each individually. However, since the combination of management situation and decision rule fully specify the use that is to be made of the sample survey information, this is no major drawback.

The particular population under consideration also affects the value placed on sample survey information. The effect of variations in the population characteristics are demonstrated in the timber sale appraisal problem of Chapter 4. The three populations considered were the original Douglas-fir population reported by Johnson and Hixon (1952) and two modified populations generated by dividing each observed volume by ten and by one hundred, respectively. The effect of these modifications is to reduce the value of the timber on a per unit area basis. This has the effect of reducing the value of sample survey information and thus, reduces the specified optimal sample size.

In addition each modified population has a variance that is 1/100 and 1/10,000, respectively, of the variance of the original Douglas-fir population. This has the effect of reducing the probability that large deviations from the truth occur in volume estimations. Although the size of these deviations relative to the population mean remain unaffected, the absolute value of the expected loss resulting from making decisions based on sample survey information collected at any given level of precision is reduced. This causes a decrease in the optimal sample size. This relationship of optimal sample size to population variance is demonstrated analytically for squared error and absolute value type loss functions in Chapter 3.
For the timber sale appraisal problem, the sensitivity of the precision specification to changes in cost and price parameters was also investigated. These changes might also be considered as changes or variations in the population. The cost of sampling is directly related to population characteristics that control the difficulty of traveling through the stand and of measuring individual plots. At any given time the stumpage price is a parameter of the population based on species, timber quality, potential markets, and other similar population characteristics.

The sensitivity of precision specification to changes in cost and stumpage price parameters appears generally to be dependent on the loss function being used and on the population being analysed. For high value, high density, old growth Douglas-fir, moderate changes in cost and stumpage price parameters appear to have no effect on precision specification. When the absolute value loss function is used for analysis with either of the two modified populations, this sensitivity becomes more apparent. For these populations an increase in cost parameters results in a decrease in the optimal sample size while an increase in stumpage price results in an increase in the optimal sample size. Similar results occur for the other members of the first two sets of loss functions when they are applied to the lowest density modified population. These results are demonstrated analytically for the squared error loss function and the absolute value loss function in Chapter 3.

The selection of the proper loss function also has an effect on the precision specification. For some general forms of loss function, small variations in the individual loss function result in marked differences in the specification of optimal precision. This is clearly indicated by the
loss function labeled $L_{14}$. Varying the limit of acceptable error from approximately one-half percent to approximately twenty percent of the population mean results in precision specifications ranging from eighty-five percent to two and one-half percent, respectively. However, for this same loss function, increasing the loss associated with being forced out of business from 50,000 dollars to 150,000 dollars has practically no effect on the precision specification.

The loss function used with allowable cut determination also clearly exhibits this sensitivity. In the first component of this loss function, varying the discount rate used in determining the value of timber harvested over the rotation from zero to two and one-half percent to five percent resulted in precision specifications of 342 plots, 306 plots, and 158 plots, respectively. Variations in the second component of this loss function also resulted in marked differences in the specification of optimal precision.

Those loss functions formed by combinations of the absolute value and squared error loss functions represent a class of loss functions for which the precision specification is affected only slightly by small changes in the structure of individual loss functions. In addition, with the exception of low density, low value timber, the precision specification, when these loss functions are utilized, is almost unaffected by small variations in sampling costs or stumpage prices.

In general it appears that modifications of loss functions which leave the basic loss structure of the function relatively unaltered have little effect on the precision specification procedure. However, those modifications of the loss function that result in a major change in the basic loss structure may have a marked effect on precision specification. In addition,
major changes in cost and price parameters may be expected to usually result in variations in the specified optimal precision level.

Clearly each management decision making situation and each decision rule presents a unique case for the analysis of the precision specification problem. Therefore, in this study I have offered no general rule of thumb for determining the optimal sample size. Instead, I suggest a general approach or procedure to be followed in developing the analysis of the precision specification problem for any given forest management decision making situation. Rather than specifying the required sample size for a sample survey on the basis of a desired level of precision, I suggest evaluating the sample size directly as a function of the value of the information that is to be provided by the sample survey.

As is shown in Chapter 2 all the methods of analysis discussed in this study may be considered modifications of the basic minimization of a cost plus loss function procedure suggested by Cochran (1963). Thus for any given forest management situation, there are only three data needs that must be filled for the analysis of the precision specification problem. These are the development of an appropriate cost of sampling function, the development of a loss function that adequately describes the losses that occur when decisions are made based on sample survey information collected at any given level of precision, and the determination of the appropriate distribution of the possible outcomes of the sample survey.

For most of the forest management situations that might be encountered, the population of interest and the set of possible outcomes will fulfill the assumptions necessary for the Central Limit Theorem to provide an acceptable estimate of the distribution of the possible outcomes. For
these cases the distribution of outcomes may be assumed to be normal. If this population is a completely enumerated research population, the mean and variance of the normal distribution will be known. For other populations it will be necessary to estimate the mean and variance of the population either from a small preliminary sample of the population or from experience in working with similar populations. Frequently this second approach to estimating the population mean and variance will be adequate. Some sampling designs and estimation procedures may lead to sets of possible outcomes of the sample survey which do not fulfill these assumptions. For these cases the normal distribution suggested by the Central Limit Theorem may not provide an adequate estimate of the distribution of outcomes and alternative distributions may have to be investigated.

Much is written about the development of cost of sampling functions. The basic form of this function is closely related to the sampling design used to specify which sampling units in the population are to be measured. However, except for some situations in which the sampling units must be visited in some prespecified sequence, a cost of sampling function similar to that suggested for the allowable cut determination problem should prove to be adequate. The appropriate cost coefficients for the cost of sampling function are unique for each forest management situation. They are developed by conducting time studies on the population of interest similar to those conducted by Johnson and Nixon (1952) and by determining the average wage of personnel likely to be conducting the survey.

The development of the appropriate loss function for the management situation of interest should be based on a thorough study of the economic environment surrounding this management situation. This study should include
an investigation of all factors that might influence the losses that result from making a decision based on the outcome of the sample survey given \( \theta \) is the true state of nature. For various forest management decision making situations this might include, among other things, stumpage price expected for the timber, possible markets for timber, efficiency of the available labor force, and capacity of the decision maker to alter management plans when errors become apparent.

A large part of the information necessary to develop an appropriate loss function would have to be obtained from interviews with the decision maker. The loss function utilized in any given situation is a formalized statement of the losses the decision maker can expect when he takes a specified action given \( \theta \) is the true state of nature. Thus the actual form of the function cannot be specified in advance by someone unfamiliar with the forest management decision being analysed. Instead this function must be developed by the decision maker himself with a full awareness of all factors that might affect the nature of losses he can expect.

For some cases where the possible actions and possible states of nature are discrete and limited, it may be possible to evaluate the loss that can be expected for each combination of action and state of nature. In other cases where either or both the set of possible actions and the set of possible states of nature are continuous, the loss function will have to be expressed in a functional form. However, this functional form would be determined by the decision maker considering the losses expected for specified combinations of actions and states of nature. By considering a range of such combinations, the decision maker may decide either that one of the commonly used loss functions (e.g., squared error loss or absolute value
loss) adequately describes the losses incurred or that a unique loss function is described by the losses incurred. If this second decision is reached the form of the loss function is determined by fitting the appropriate function to the set of losses evaluated for the range of combinations of actions and states of nature considered above. A discussion in Chapter 5 illustrates in some detail the factors that must be considered in developing a loss function for a given forest management situation and describes the type of rationale that might be used to arrive at an appropriate approximation of the functional form of this loss function.

Clearly this method of analysis cannot be performed each time sample survey information is required in a decision making process. Instead, for a given forest management situation in a given area, where similar decisions will have to be made for many populations or perhaps at several times for the same population, this precision specification analysis should be conducted on a "representative" population. For this analysis considerable effort should be put into satisfying completely and thoroughly the three basic data needs which I have just discussed.

This analysis should include a study of the sensitivity of the precision specification process to changes in cost of sampling and loss function parameters. Once this analysis has been performed for a given forest management decision making situation, the forest manager has available the necessary information required to specify the optimal sample size for his sample survey. He needs only to evaluate the cost of sampling parameters and the loss function parameters for his population. By comparing these values with the parameters used in the sensitivity analysis, the desired sample size usually is readily estimated.
The largest expense that is incurred in this general analysis procedure lies in fully defining the nature of the cost plus loss function. Thus, once this function is fully defined, the actual determination of the optimal precision specification may be handled quite easily, even when parameters vary from those specified in the original analysis.
Austin, John W. 1968. Production, prices, employment, and trade in north­west forest industries, third quarter 1968. Pacific Northwest Forest and Range Experiment Station, U. S. Forest Service.


Raiffa, Howard and Schlaifer, Robert. 1961. Applied statistical decision theory. Boston, Massachusetts, Division of Research, Graduate School of Business Administration, Harvard University.


Schreuder, Hans P., Sedransk, Joseph, and Ware, Kenneth D. 1968. 3-P sampling and some alternatives, I. Forest Science 14: 429-453.


Acknowledgment

I would like to thank the members of my graduate committee and in particular, Dr. Kenneth Ware, for the guidance and encouragement offered during the preparation of this dissertation. In addition I would like to express appreciation to fellow graduate students who frequently provided a critical first audience for the discussion of ideas or concepts during the developmental stages of this dissertation.