LOW INTENSITY FOREST SAMPLING THROUGH USE OF STEM FREQUENCY DISTRIBUTION AND POPULATION PARAMETERS

By

MICHAEL JOHN DAHLEM

Bachelor of Science

Oklahoma State University

1970

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
MASTER OF SCIENCE
July, 1972

LOW INTENSITY FOREST SAMPLING THROUGH USE
OF STEM FREQUENCY DISTRIBUTION
AND POPULATION PARAMETERS

Thesis Approved:

Nat Walker
Thesis Adviser

Daniel D. Badger

Edward E. Sturgeon

N Durham
Dean of the Graduate College

ACKNOWLEDGEMENTS

Special thanks are due Dr. Nat Walker, my thesis adviser, who rendered great assistance and encouragement in this work. The assistance of the other members of my graduate committee, Dr. Daniel Badger and Dr. Edward E. Sturgeon, is also appreciated.

In addition I would also like to thank Floyd E. Bridgwater whose assistance in computer programming was appreciated.

Thanks are also due the Dierks Division of Weyerhaueser for the use of their inventory data.

I would also like to thank Linda Rolin for typing this manuscript.

Finally, I would like to thank my wife, Junie, whose understanding, moral support and many sacrifices made possible the completion of this work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES (Continued)

CHAPTER I

INTRODUCTION

In developing any forest management plan, a basic component is reliable inventory data. These data are used in making growth predictions, adopting cutting cycles and in formulating cutting budgets. Historically, such data have been obtained by conventional means, which, because of the requirement for sampling intensity and periodic field measurements, are time consuming and costly. Typical costs for a 1.8 million acre ownership reach $20,000.00 per year and represents a major cost of management.

The objective of any forest sampling project is to provide estimates on such forest population characteristics as volume, quality and species distribution. Limits of error are established and sampling intensities required to provide reliable estimates, within the present limits of error, are determined as the sampling proceeds. Timber stands, however, are not grown simply to be inventoried. The measurement of trees add no real value to the materials harvested. Therefore, in all sampling projects, cost factors assume an importance equal to, or greater than, the statistical accuracy desired. The forest manager then must seek out more efficient methods for obtaining forest inventories in order to cut costs to a minimum.

The purposes of this paper are twofold: (1) To determine if the observed size class distributions of the data obtained for this project fits well with the univariate negative binomial distribution which has been observed to be the case with most species and timber types and (2) If the data fits well with the negative binomial distribution, to make predictions based on low intensity samples and estimated population parameters. Emphasis is placed on reducing the intensity of the sampling so as to obtain a reduction of costs, while maintaining an adequate picture of the forest structure. In this study, quantification of savings will be made only by way of possible reductions in the sampling intensity. No assignment of dollar values will be attempted.

CHAPTER II

PREVIOUS WORK

A great deal of information on the distributions of various species in different plant associations has been available for a number of years. Early plant ecologists distinguished three types of plant distributions--regular, random and aggregate. Blackman (1935), in some of the first critical analysis of plant counts, attempted to fit the Poisson series to the observed distributions. Following his lead, others used this as a basis for their calculations. Clapham (1936), however, in studying various plant distributions, found only a few that could be termed "random". Feller (1943) suggested that plant species tended to be aggregated rather than randomly distributed, and suggested that it was due to habitat or reproductive characteristics. Cole (1946), in studying various forest dwelling biota, found only spiders to be distributed in a random manner. Archibald (1948) showed that most plant distributions could better be described as aggregated rather than random. Many of the earlier ecologists believed plants followed random distributions.

Robinson (1954), noting the significant differences obtained between the observed and expected frequencies reported by earlier workers, suggested that the groups of plants were distributed in a Poisson manner, but the individuals within the groups followed a logarithmic distribution. He then showed the resulting compound distribution of individuals to be the negative binomial distribution.

The negative binomial series, derived from the Poisson series, is very versatile in describing many naturally occurring populations. First derived in the early 18th century and used for many years with animals, the negative binomial series was developed in its present form by Bliss and Fisher (1953). Bates and Neyman (1952) derived the bivariate negative binomial to examine joint distributions.

The phenomenon of aggregation or clumpiness is termed "contagion". Simply stated, the presence of one individual in a sample unit influences in observing this clustering, termed the distribution "contagious". Grosenbaugh (1952) suggested the ratio $\frac{s^2}{\bar{x}}$ as a measure of the contagion. Pielou (1960) used the variance-mean ratio to identify regular, random, and aggregated populations.

The work concerned with the application of stem frequency distribution information to the field of forestry, however, is quite limited. Walker (1970) has made the only effort in applying the negative binomial series to tree stem frequency distribution analysis. In 31 sampling projects in 19 different forest types he found all to be contagious distributions which fitted well with the negative binomial series. Thus far, however, no work has been attempted in applying the negative binomial distribution to sampling techniques.

CHAPTER III

PROCEDURE

The data for this project was obtained from the Dierks Division of Weyerhaeuser Company. Included are 1274 Continuous Forest Inventory (CFI) plots from the Oklahoma Block located in southeastern Oklahoma. These data represent an inventory of approximately 890,000 acres. The data were recorded in trees per acre by 2-inch diameter classes with the 4-, 6-, and 8-inch classes recorded on tenth-acre plots and the 10- through 30-inch classes recorded on quarter-acre plots. Plots devoid of pine were discarded for the purposes of this study. Pine was present on 1139 of the 1274 plots.

The data were converted to a trees-per-plot basis and punched on IBM cards (one card for each plot) and identified by plot number and district. From these cards, stem frequency tables were constructed. Each table represents, by district, the number of plots where the 4-inch count equaled 0, 1, 2, etc., the 6-inch count equaled 0, 1, 2, etc., and successively on up to the 30-inch class.

Stem frequency tables were constructed for each diameter class by district and for the entire block. A computer program was then used to obtain sets of negative binomial frequency values, means, variances and $\underline{k}$ values by the maximum likelihood method. The distributions were then tested with chi-square. Due to the estimating of the two parameters ($\bar{x}$ and $\hat{\underline{k}}$) of the negative binomial series, N - 3 degrees of freedom were used in testing. Distributions with fewer than four entries could not be tested. In general, the 22-inch and larger diameter classes lacked sufficient entries for testing to be accomplished. Negative binomial fittings were satisfactory for all distributions except that of the 6-inch class in District 2. Chi-square probabilities of .05 and greater are accepted as indicating satisfactory fittings.

Random samples were drawn from each district, using a series of six random numbers punched in the last six columns of each card and sorting columns at random.

The sampling proceeded by drawing 20 percent of the sample units in each district, repeating the procedure of calculating negative binomial distributions and testing the distributions with chi-square. The sample size was increased for successive samples and the procedure repeated until a level was reached that gave satisfactory results. When the satisfactory level of intensity was reached, additional samples of that size were drawn to verify the results. The level at which satisfactory results were obtained was 35 percent of the sample units.

A by-product of this project was an analysis of association between diameter classes. Joint frequency tables were constructed using a standard computer program. These tables represent the occurrence of each count of one diameter class to each count of another diameter class. Bivariate negative binomial distribution fittings were then made for each joint frequency table and tested with chi-square.

CHAPTER IV

DEVELOPMENT

Distribution Characteristics of Natural Populations

The three types of distribution patterns generally recognized in studying the spatial arrangement of biological populations are regular, random and aggregate. The population is said to have a regular distribution if the individuals of the population are spaced more evenly than they would be if they were distributed purely by chance. Random populations occur when the position of each of the individual plants is independent of that of all other plants. An aggregated population is one in which there is a tendency for the individuals of the species to occur in clumps.

Greig-Smith (1957) reports that regular distributions appear to be extremely rare in plant populations. Such distributions would occur only if the members of the population were so abundant that they were in competition with each other for the available space.

Early works of ecologists revealed that random distributions also seemed to be rare. Blackman (1935), in his analysis of plant counts, attempted to fit the Poisson series to the observed distributions. Others also used the Poisson series, but in the majority of cases, significant differences were obtained between observed results and the numbers expected. Clapham (1936) found only four plant species out of a total of 44 which showed distributions that could not be distinguished from random. Cole (1946), in the study of a variety of forest-dwelling organisms found only spiders were distributed in a random manner.

Aggregated populations, however, have been found to be very common in nature. Robinson (1954), points out that while groups of plant species in sample quadrats may exhibit a random distribution, the individuals in each group are aggregated, following a logarithmic distribution. The resulting compound distribution is the negative binomial type. Several causes for the occurrence of aggregated populations have been suggested.

Feller (1943) suggests that while seeds may fall at random over an area, the habitat may not be homogeneous and the proportions germinating and surviving will vary from site to site so that the density is high on some sites and low on others. In contrast, the habitat may be homogeneous but due to reproductive characteristics, such as heavy seeds or vegetative reproduction, the individuals may occur in family groups. Pielou (1960), and Robinson (1954), suggest that quadrat size influences the type of distribution. Trees so large that their shade inhibits the reproduction and growth of other individuals may appear to be random, or uniformly distributed but would be aggregated if a larger sample unit size were used. The same would hold true for small trees if the sample size were small enough to conceal the identity of the distribution.

Regardless of the cause for the departure from evenness or random types of distributions, aggregated distributions belong to the type termed "contagious distributions" (Ploya, 1931). The implication is that the presence of one or more organisms within a sample unit influences the probability of other like organisms occuring in the sample unit. Organisms become concentrated in relatively few of the sample units, leaving by contrast with random distributions, an excess of unoccupied units and of densely occupied units, and a deficiency in the number of sample units yielding isolated individuals or small numbers of individuals (Cole 1946).

The departure from randomness or degree of contagion can be measured by the ration $\frac{s^2}{\bar{x}}$. Grosenbaugh (1952) suggested the use of this ratio as a basal area clustering coefficient. Pielou (1960) used the variance-mean ratio in identifying regular, random and aggregate populations in an attempt to identify a mechanism to account for these types of distributions. When the value of the contagion factor is greater than 1, contagious distributions are indicated, that is, the individuals are clustered more tightly than would occur by chance. With the variance-mean ratio equal to 1, a random or Poisson distribution is suggested. Values less than 1 indicate an approach toward even, or regular, distributions.

Computation of the contagion factor in the analysis of the data for this study revealed that the individual diameter classes exhibited aggregation. Contagion varied from very high for the smaller diameter classes, to quite mild in the larger diameter classes. The high values of the contagion factor for the smaller diameter classes indicates very tight clustering and a high degree of variation in the stem count per plot. The reverse would be true for the larger diameter classes.

Use of the Negative Binomial to Forecast Distributions within Size Classes

In studying the occurrence of plants and animals in nature, a number of distributions series has been devised. The individuals occurring in any unit of time or space can be transformed into a frequency distribution indicating the number of units containing x = 0, 1, 2,...n, individuals. If every unit in a series were exposed equally to the chance of containing the individual, the distribution would follow the Poisson series (Bliss 1953). The expected variance of the Poisson distribution is equal to the mean. Forest stand structures, however, exhibit the characteristic of having variances significantly larger than the mean. This characteristic is called "overdispersion", or contagion.

The negative binomial series, developed in its present form by Bliss and Fisher (1953), is a curve-fitting equation used in describing distributions which exhibit contagion. It is defined by the two parameters of the equation, the mean (μ), and the exponent $\underline{k}$. The two parameters ($\bar{x}$ and $\hat{k}$), of the negative binomial series, are estimated from the field plot data. The hat symbol is normally reserved for the maximum likelihood estimate of $\underline{k}$, which will be explained later in this paper.

The mean ($\bar{x}$) is estimated efficiently by the arithmetic mean of the sample:

$$\bar{x} = \frac{\Sigma fx}{N}$$

where N = the number of sample units.

The exponent $\underline{k}$ is estimated, as described by Bliss (1953), in three ways, two of which will be discussed here. The simplest solution, the moment estimate, for determining k, is derived from the mean and variance of the sample by:

$$k_1 = \frac{(\bar{x})^2}{s^2 - \bar{x}}$$

where the variance (s^2) is:

$$s^2 = \frac{\Sigma\,(fx^2) - \frac{(\Sigma fx)^2}{N}}{N - 1}$$

The moment estimate is a suitable estimate for most practical purposes.

The more cumbersome maximum likelihood estimate requires that $z_i \tilde{=} 0$ in the equation:

$$z_i = \Sigma\left(\frac{Ax}{k_i + x}\right) - N \cdot \ln\left(1 + \frac{x}{k_i}\right)$$

where z_i....j are scores that are computed from trial values of k, selected so that they bracket the required estimate, $\hat{k}$.

The expected frequencies (Ø) of the negative binomial series are derived from the expansion of $(Q - P)^{-k}$, where $P = \frac{\mu}{k}$ and $Q = 1 + P$. The probability that a plot will contain 0, 1, 2,...n trees is:

$$Px = \frac{(k + x - 1)!}{x!\ \ (k - 1)} \cdot \frac{R^x}{q^k}$$

where $R = \frac{\mu}{k + \mu}$.

The probability for a given x is multiplied by N, the total number of plots counted, to obtain the expected frequency (Ø) of plots with x individuals. The frequency for the 0 count is computed as:

$$Ø_o = \frac{N}{q^k} .$$

With the entry for x = 1 as:

$$\emptyset_1 = k \cdot R \cdot \emptyset_o$$

and succeeding entries as:

$$\emptyset_{2 \cdots j} = \frac{(k + x - 1)}{x} \cdot R \cdot \emptyset_{x-1}$$

In the following illustrative example, the expected frequencies will be computed for the 16-inch class, District 1 (Table I).

TABLE I

DISTRIBUTION OF 16-INCH PINE, DISTRICT 1, OKLAHOMA
BLOCK DIERKS DIVISION OF WEYERHAUESER

Stem Count (x)	No. Sample Units (f)	Accumulated Frequency (Ax)	Total Count (fx)
0	102	69	0
1	37	32	37
2	20	12	40
3	7	5	21
4	3	2	12
5	1	1	5
6	1		6
	$\Sigma f = N =$ 171		$\Sigma =$ 121

The mean is estimated by:

$$\bar{x} = \frac{\Sigma fx}{N} = \frac{121}{171} = .7076$$

and the variance s^2 by:

$$\frac{\Sigma(fx^2) - \frac{(\Sigma fx)^2}{N}}{N} \doteq \frac{289 - \frac{(121)^2}{171}}{171 - 1} \doteq \frac{289 - 85.62}{170} = 1.1964.$$

In the search for the maximum likelihood estimate of k, the first estimate is made by:

$$k_1 = \frac{(\bar{x})^2}{s^2 - \bar{x}} = \frac{(.7076)^2}{1.1964 - .7076} = 1.0245$$

The first iteration in the search for $\hat{k}$, using $k_1 = 1.0245$ is shown in Table II.

TABLE II

FIRST ITERATION IN THE SEARCH FOR $\hat{k}$ BY THE MAXIMUM LIKELIHOOD METHOD

Stem Count	Ax	$\frac{1}{k_1 + x}$	$Ax \cdot \frac{1}{k_1 + x}$
0	69	.97609	67.3502
1	32	.49395	15.8064
2	12	.33063	3.9675
3	5	.24848	1.2424
4	2	.19902	.3980
5	1	.16599	.1659
			88.9306

In the search for $z_i \cong 0$

$$z_i = \Sigma \left(\frac{Ax}{k_1 + x}\right) - N \cdot \ln \left(1 + \frac{\bar{x}}{k_1}\right)$$

where k_1 is the value for the first iteration.

For ease of computation the second term of the equation is written as:

$$2.3025851 \ (N) \ \log_{10} \left(1 + \frac{\bar{x}}{k_1}\right),$$

then $z_1 = \Sigma(\frac{Ax}{k_1 + x}) - 2.3025851\ (N)\ \log_{10}\ (1 + \frac{\bar{x}}{k_1}) =$

$$(88.9306)\ -\ 2.3025851\ (171)\ \log_{10}\ (1 + \frac{.7076}{1.0245})$$

$$z_1 = 88.9306 - 89.79 = -.8668$$

The negative sign indicates that $k_2 < k_1$. The value k_2 should be chosen just large enough to give differing signs. This estimate is based on experience of estimating the parameter k̲. Therefore, $k_2 = .92$ is chosen and another iteration is performed.

$$\Sigma(\frac{Ax}{k_2 + x}) = 97.62718$$

$$2.3025851\ (171)\ \log_{10}\ (1 + \frac{.7076}{.92}) = 97.55347$$

$$z_2 = 97.62718 - 97.55347 = .07371\ .$$

By interpolation (Fisher, 1953), k_3 found by:

$$k_2 + \frac{(z_2)\ (k_1 - k_2)}{z_1\ -\ z_2}\quad \text{or}$$

$$.92 + \frac{(.07371\ \cdot\ .1045)}{.8668 + .07371} = .9281$$

Table III shows the additional iterations required in the search for $\underline{\hat{k}}$. The value k_4 was obtained by interpolation between k_2 and k_3.

Since $z_4 \doteq 0$ the value .9266 is chosen for $\underline{\hat{k}}$. The availability of modern computers makes the estimation of the maximum likelihood $\underline{\hat{k}}$ much easier.

The values q and R can now be calculated:

$$q = (1 + \frac{\bar{x}}{k}) = 1 + \frac{.7076}{.9266} = 1.7636$$

$$R = \frac{\bar{x}}{k + \bar{x}} = \frac{.7076}{.9266 + .7076} = .4330$$

TABLE III

ADDITIONAL ITERATIONS IN THE SEARCH FOR THE MAXIMUM LIKELIHOOD ESTIMATE OF k

	k_1=1.0245	k_2=.92	k_3-.9281	k_4=.9266
$\Sigma\ (Ax + \frac{1}{k_i + x})$	88.9306	97.62718	96.8877	97.02343
$2.3025851\ (N)\ \log_{10}\ (1 + \frac{\bar{x}}{k})$	89.79	97.55347	96.9034	97.02311
z_i	-0.8668	.073715	-.05704	.00032

Preliminary calculations for the computation of the expected frequencies are now complete.

$$\emptyset_o = \frac{N}{q^k}\text{, or for ease of computation,}$$

$$\log \emptyset_o = \log N - k \log q$$

$$= 2.232996 - .9266\ (.2464)$$

$$\log \emptyset_o = 2.00468176$$

$$\emptyset_o = \qquad 101.10$$

$$\emptyset_1 = k \cdot R \cdot \emptyset_o \qquad = \quad 40.56$$

$$= .9266 \cdot .4330 \cdot 101.1$$

$$\emptyset_2 = \frac{.9266 + 2 - 1}{2}\ (.4330)\ (40.56) \quad = \quad 16.91$$

$$\emptyset_3 = \frac{.9622 + 3 - 1}{3}\ (.4330)\ (16.91) \quad = \quad 7.14$$

$$\emptyset_4 = \frac{.9266 + 4 - 1}{4}\ (.4330)\ (7.14) \quad = \quad 3.03$$

$$\emptyset_5 = \frac{.9266 + 5 - 1}{5}\ (.4330)\ (3.03) \quad = \quad 1.29$$

$$\emptyset_6 = \frac{.9266 + 6 - 1}{6}\ (.4330)\ (1.29) \quad = \quad .55$$

The most convincing test of the adequacy of the negative binomial is the chi-square test of the differences between the observed frequencies (f) and the expected frequencies (Ø). Table IV shows the chi-square test for the above fitted values of Ø.

TABLE IV

CHI-SQUARE TEST FOR THE GOODNESS OF FIT, 16-INCH TREE DISTRIBUTION DISTRICT 1, OKLAHOMA BLOCK, DIERKS DIVISION OF WEYERHAUESER

Stem Count (x)	Observed Values (f)	Expected Values (Ø)	Rounded Values	$\frac{(O - E)^2}{E}$
0	102	101.10	101	.00900
1	37	40.56	41	.39024
2	20	16.91	17	.52941
3	7	7.41	7	
4	3	3.03	3	
5	1	1.29	1	
6	1	.55	1	
	171		171	X^2 = .92955

The expected frequencies are pooled in an attempt to obtain a minimum expected value of at least 3. With N - 3 + 2 degrees of freedom, chi-square = .9266 has a probability of P = ~ .70, an acceptable fit. Diameter classes with fewer than four entries cannot be tested.

Errors Associated with the Selection of the Mean and *k*

The forecasting of distributions by the negative binomial is dependent on acceptable estimates for the two population parameters of the distribution, $\bar{x}$ and $\hat{k}$.. Errors in the estimation of these parameters will have varying effects on the predicted distribution.

Errors associated with the mean generally are a result of improper technique or insufficient sampling. Errors of this type can be reduced by adequate planning of the sampling procedures and computation of standard errors and confidence intervals to find the sampling intensity that will bring the estimate within acceptable limits of error. Walker (1970) reports that another source of error in computing means may occur in classes where rounded frequencies are used in distributions characterized by relatively large means coupled with low k values. Observed means will differ from the mean of the fitted distribution. This occurs in distributions with fairly long tails, (generally restricted to the 6-inch and 8-inch classes). Recommendations to alleviate this type of error were to adopt fixed radius plots smaller than one-fifth acre for the evaluation of these size classes.

Errors associated with k values may occur if values for the parameter are chosen on the basis of experience rather than being computed from observed data. The effect of inaccurate estimates of k is the overestimation or underestimation of the zero count and other count frequencies. While this type of error will not affect estimates of volume, it will give a distorted representation of the forest structure. The effects of the departure from the computed k is illustrated in Tables V and VI.

With the calculated k less than 1.0, the range of acceptable arbitrary values of k becomes critical. As shown in Table V (calculated $\hat{k}$ = .8959) arbitrary values of k = .8 and 1.0 generate acceptable distributions while all others yield unacceptable distributions. In Table VI (calculated $\hat{k}$ = 1.8225), k = 1.0 and 2.0 generate acceptable distributions, indicating a wider range of arbitrary values for k is available. Those values of k greater than 2 generate comparable frequencies indicating that k values can be chosen with confidence if the actual value of k for the distribution is greater than 2.

TABLE V

EFFECT OF THE DEPARTURE FROM THE CALCULATED $\hat{k}$ ON THE DISTRIBUTION OF 12-INCH TREES, DISTRICT 2

N = 188, $\bar{x}$ = 1.0943, $\hat{k}$ = .8959

Stem Count	k=0.5	k=0.6	k=0.7	k=0.8	Calculated $\hat{k}$=.8958	k=1.0	k=10	k=20
0	86	80	75	71	68	65	33	31
1	34	36	38	40	41	43	53	53
2	20	22	24	25	27	28	46	48
3	13	15	16	17	18	18	30	31
4	9	10	11	11	12	12	15	15
5	7	7	7	8	8	8	7	7
6	5	5	5	5	5	5	3	2
7	4	4	4	4	3	3	1	1
8	3	3	3	2	2	2		
9	2	2	2	2	2	1		
10	1	1	1	1	1	1		
11	1	1	1	1	1	1		
12	1	1	1	1				
13	1	1						
14	1							

TABLE VI

EFFECT OF THE DEPARTURE FROM THE CALCULATED k ON THE DISTRIBUTION OF 14-INCH TREES, DISTRICT 1

N = 171, $\bar{x}$ = 1.3743, $\hat{k}$ = 1.8225

Stem Count	k=0.5	k=0.8	k=1.0	Calculated $\hat{k}$=1.8225	k=2	k=3	k=5	k=10
0	88	77	72	61	60	55	51	47
1	32	39	42	48	49	52	55	57
2	18	22	24	29	30	33	35	38
3	11	13	14	16	16	17	18	18
4	7	8	8	8	8	8	8	8
5	5	5	4	4	4	4	3	3
6	3	3	3	2	2	1	1	1
7	2	2	2	1	1			
8	1	1	1	1				
9	1	1	1	1				
10	1							
11	1							
12	1							

CHAPTER V

RESULTS AND DISCUSSION

Forecasting Distributions with a Reduced Sample Size

The development of stem distribution information provides the forest manager with a number of management tools. One such piece of information is the basis for this study, that is, the prediction of forest structures with a much reduced sample size.

If it can be shown that the univariate negative binomial distribution fits well the observed size class distributions, then it becomes possible to make forecasts of the distribution on smaller areas within, and areas contiguous to the sample area, by estimating the population parameters based on low intensity samples. Forecasts can possibly be made in two ways; (1) by estimating from reduced samples, the mean and $\hat{\underline{k}}$ of the negative binomial series and projecting the distribution and (2) using the computed population parameter $\underline{k}$ and projecting the distribution by estimating the mean from the reduced sample.

The first step in the procedure is to determine the goodness of fit of the observed size class distributions with the univariate negative binomial distribution. Data from 1139 CFI plots were analyzed. Frequency tables were constructed by district. The mean and the exponent $\hat{\underline{k}}$ were computed for each diameter class. The expected frequencies were then computed and chi-square tests for goodness of fit calculated and recorded. The results of the analysis is shown in Table VII. The table shows the mean, $\hat{\underline{k}}$, and the probability obtained in the chi-square tests with the observed size class distribution. Only District 1 contained sufficient entries in the 20-inch class for testing with chi-square. One distribution, (6-inch class, district 2), failed to fit well with the observed distribution. Values for the maximum likelihood estimate of $\underline{k}$ varied from .1835 (20-inch class, district 1) to 1.8225 (14-inch class, district 1). Most calculated values of $\underline{k}$ were less than 1.0. The results of the chi-square tests indicate that the observed size class distributions, obtained from a large-scale sampling project fits well with the negative binomial distribution and are in accordance with the findings of Walker (1970).

The Sampling Intensity

The question of how many units should be taken depends on a number of factors, including such economic considerations as the degree of accuracy desired and the amount of money available for the survey. Also involved are such characteristics as the size of the tract to be sampled, the natural variation of the forest conditions within the sampling units and the size and value of the timber to be sampled. According to statistical theory for large forest areas, only a very small proportion of the area must be sampled to obtain an accurate estimate. For tracts of a few hundred acres or less, fairly heavy sampling is needed (5-20 percent intensity). For proportionately smaller areas a larger percentage of the area must be sampled to obtain accurate estimates (Spurr, 1952).

Assume at this point that the 1139 quarter-acre sample units represents 284.75 acres, that the data represents a 100 percent cruise of the

TABLE VII

RESULTS OF DETERMINING THE GOODNESS OF FIT OF THE OBSERVED SIZE CLASS DISTRIBUTIONS WITH THE EXPECTED DISTRIBUTIONS OF THE NEGATIVE BINOMIAL SERIES

		Diameter Class								
		4-inch	6-inch	8-inch	10-inch	12-inch	14-inch	16-inch	18-inch	20-inch
District 1	x̄	3.1053	2.5439	1.9006	3.9591	2.5439	1.3743	.7076	.3041	.1930
N=171	k	.7523	1.0802	.7812	1.2513	1.4105	1.8225	.9266	.2533	.1835
	≈ P	.35	.45	.60	.95	.30	.70	.70	.70	.60
District 2	x̄	3.9202	2.877	1.7660	3.0479	1.9043	1.2553	.5372	.2181	
N=188	k	.6377	.7026	.6761	.5764	.8959	.9478	.8465	.6649	
	~ P	.73	.01	.35	.50	.20	.30	.45	.30	
District 3	x̄	2.6809	2.1986	1.6170	2.4823	1.8794	1.2553	.4539	.1773	
N=141	k	.3861	.4787	.7067	.8609	.9552	1.1106	.5654	.9812	
	~ P	.65	.60	.55	.40	.75	.35	.15	.60	
District 4	x̄	4.4639	3.5309	2.1186	3.6237	2.2371	1.3814	.6031	.1598	
N=194	k	.7563	.9766	1.0198	.9829	1.2401	.9297	1.0694	.6928	
	~ P	.45	.85	.20	.30	.30	.15	.40	.60	
District 5	x̄	5.5204	4.2217	2.4480	3.2851	1.8869	1.0181	.4932	.1991	
N=221	k	.7410	1.0843	1.1679	.7338	1.3235	1.5373	1.3012	.4785	
	~ P	.45	.85	.20	.30	.30	.15	.40	.60	
District 6	x̄	5.4911	3.2411	1.3348	1.9643	1.5357	1.1295	.6964	.2143	
N=224	k	.4433	.6892	.7161	.5233	.8099	1.2284	1.1721	1.0985	
	~ P	.83	.83	.20	.90	.80	.92	.60	.15	
Block	x̄	4.3565	3.1870	1.8753	3.045	1.9781	1.2239	.5865	.2116	.0658
	k	.5741	.7887	.8144	.7350	1.0507	1.1678	.9483	.4893	.1283
	~ P	.40	.80	.10	.45	.97	.35	.15	.15	.35

intensity in order to make forecasts of the distribution by the two methods previously mentioned.

Forecasting Distributions by Estimating the Mean and *k* from a Reduced Sample

The first method of making forecasts of the distributions to estimate the mean and $\hat{k}$ from a random sample. The initial sample consisted of drawing 20 percent of the sample plots in each district at random. The mean and $\hat{k}$ were estimated for each distribution. The expected frequencies were computed by substituting the estimates of the mean and $\hat{k}$ in the formula to obtain the expected negative binomial distribution. The N value used in the computations were the N values for the district in order to obtain comparable distributions that could be tested with chi-square. The observed size class distributions were then compared with the expected distributions and tested with chi-square. Visual examinations of the estimated values for the mean and $\hat{k}$ revealed that reasonable estimates of the mean were made but the estimates of $\hat{k}$ varied widely from the actual *k* values. The resulting expected distributions when compared to the observed distributions proved to be unsatisfactory. Distributions from the sample using only the estimated mean and the actual $\hat{k}$ values generated a larger number of acceptable distributions although there were still many failures. Further testing using the estimated values of the mean and $\hat{k}$ revealed equally unsatisfactory results. The negative results most likely originate in the parameter *k*. Since most values of $\hat{k}$ are less than 1 and all values are less than 2, they lie in the region where the estimates become very critical. Very little deviation from the actual *k* generates unacceptable distributions as demonstrated previously in this paper. In light of the poor results, further attempts at forecasting by this method were abandoned.

Forecasting Distributions by Estimating the Mean from a Reduced Sample

Attempts to predict distributions by the estimation of the mean from reduced samples and using the actual $\underline{\hat{k}}$ values for the population gave satisfactory results. The results of the sampling are shown in Tables XIII through XVIII in the Appendix. Each table depicts the results obtained in the sampling for each district. The first two rows show the calculated mean and the probability obtained in testing the goodness of fit with the univariate negative binomial distributions computed for each size class. The other rows show examples of the results obtained at the three sample intensities tested. The intensities of sampling were 20, 30, and 35 percent. The tables show four examples at the 35 percent intensity, the level at which consistently good results were obtained. In general, the results improved as the sampling intensity increased, but increasing the intensity beyond the 35 percent level did not yield improvements in the estimates.

From each sample the mean was estimated for each size class. The estimated mean and the population value for $\underline{k}$ were used in computing the expected frequencies. The distributions were then tested with chi-square and the probabilities recorded.

The initial sampling proceeded by drawing 20 percent of the sample units from each district. This level of intensity was unsatisfactory for two reasons; (1) of the forty-six distributions tested, fifteen failed to fit satisfactorily and (2) there was tendency for the 18-inch class to be skipped in the sampling where there were only a small number of stems available for sampling.

At the 30 percent sampling intensity the number of acceptable fits increased significantly. Only four of the forty-nine distributions failed to fit satisfactorily. With the 30 percent intensity samples, the 18-inch class was represented in all districts.

The best success rate was attained at the 35 percent sampling intensity. In one sample only one failure occurred. Two samples had two failures each and the fourth had four failures of the forecast distributions. By district, District 5 had no failures; District 2 and 6 had one failure each; Districts 1 and 3 had two failures and in District 4, three failures occurred. In general, the rate of success increased as N (the number of plots per district) increased.

CHAPTER VI

CONCLUSIONS

The poor results obtained in the 20 percent sample are most likely due simply to insufficient sampling.

Results improved significantly in the 30 percent sample but the 35 percent sample proved to be best for forecasting the distributions. It was expected that where a weak fit existed bewteen the observed distribution and the expected negative binomial distribution, a weak fit would also occur in any sample drawn from that distribution. In District 2 the only failure occurred in the 6-inch class where a very weak fit was found between the observed and expected univariate distributions. The same situation exists in Districts 4 and 6. Failures occurred only where the chi-square probability in the test of the observed with the fitted distributions was .15 or less. With high chi-square probabilities between the observed and expected distributions of the population, failures of the sample to adequately forecast the distribution occurred randomly. This occurred only four times.

In many cases where a weak fit existed between the observed distribution and the expected negative binomial distribution, the sample-derived distributions were not significantly different from the expected distributions, even though the chi-square test indicated a poor fit. An example of such a case is presented in Table VIII.

The probability in testing the population distribution is .15. Two

TABLE VIII

COMPARISON OF THREE DISTRIBUTIONS FORECASTED FOR THE 4-INCH SIZE CLASS, DISTRICT 4

X	Expected (f)	Forecasted (35-1) (f)	Forecasted (35-3) (f)
0	45	48	49
1	29	30	31
2	22	23	23
3	17	17	18
4	14	14	14
5	11	11	11
6	9	9	9
7	8	7	7
8	6	6	6
9	5	5	5
10	4	4	4
11	4	3	3
12	3	3	3
13	3	2	2
14	2	2	2
15	2	2	1
16	2	1	1
17	1	1	1
18	1	1	1
19	1	1	1
20	1	1	1
21	1	1	1
22	1	1	
23	1	1	
24	1		

of the 35 percent intensity samples had chi-square probabilities of .025 and .01. A comparison of the three distributions reveals that they are almost identical, and to say that they fail to forecast the same distribution is questionable.

The results indicate that where the observed size class distributions fit the negative binomial distribution well, a forecast of the distribution can be made by estimating the mean from small intensity samples. The limits of error or confidence intervals are still in question, however. The point being stressed here is that the data analyzed for this project was a forecast for an ownership of 890,000 acres of forest land. If the reduced samples drawn here adequately forecast the data, then it is reasonable to assume that the samples also forecast the distribution for the entire management block.

CHAPTER VII

RECOMMENDATIONS FOR FUTURE RESEARCH

The purposes of this project were served by the testing of the observed distributions with the univariate negative binomial and the estimating of these distributions with reduced sample sizes. The results indicate that these objectives have been met. Future work, using this type of analysis, should concentrate on (1) the establishment of confidence limits on the estimated parameters and (2) investigating alternative tests of goodness of fit for the reduced samples with the observed distributions.

The establishment of confidence limits needs little explanation. To this point it has been stated only that the distributions could be predicted with reduced sample sizes but no investigation of the errors involved has been made. The establishment of confidence limits for the parameters $\underline{k}$ and the mean would lend reliability to the system.

The other area needing research is the testing of goodness of fit between the sample-derived distributions and the actual distributions. For the purposes of this project, the mean was estimated and combined with the computed value of $\underline{\hat{k}}$. The distributions was then expanded with the N value for the population to obtain distributions that could be tested with chi-square. In each case the sample-derived distribution was tested for agreement with the actual distribution. An alternative method might be the testing of the sample-derived distribution with the theoretical univariate negative binomial that best fits the observed distribution. One such procedure is the Kolmogorov-Smirnov one-sample test. The Kolmogorov-Smirnov one-sample test is concerned with the agreement between the distribution of a set of sample values and some specified theoretical distribution (the negative binomial distribution in this case). One advantage of the Kolmogorov-Smirnov test is that because the relative frequencies are cast into cumulative frequencies, no information is lost through the combining of categories, such as is done with the testing for goodness of fit with chi-square that requires the combining of categories to obtain a minimum expectation of three. Another advantage of the Kolmogorov-Smirnov is that the test is applicable to very small distributions with only two or three observations and the chi-square test is not. Other appropriate tests are most likely available and need to be investigated.

Another area worthy of investigation is the prediction of joint frequency (association) between adjacent size classes. The analysis of association between diameter classes provides the forest manager information on the occurrence of product combinations such as posts, pulpwood, or poles, which come from restricted size classes. The identification of the size class combinations can be made with joint frequency tables. The identification of an overabundance of these products can be made by examining the association present between the respective size classes.

Association between size classes can be measured in a number of ways. Contingency tables with chi-square tests is the most common method of measuring association, however, the 2 X 2 contingency table fails to give information on the joint frequency and does not yield a coefficient of association useful in predicting association. Walker

(1970), to overcome these difficulties, broke the joint frequency tables down into seven categories as follows:

1. Sample units having less than a mean number of trees in both size classes ($\bar{x}$, $\bar{x}$).
2. Units having more than a mean number of trees in both diameter classes ($\bar{x}$, $\bar{x}$).
3. Units having less than a mean number of trees in Y diameter class ($\bar{x}$, $\bar{x}$).
4. Units having more than a mean number of trees in X diameter class and less than a mean number Y-inch trees ($\bar{x}$, $\bar{x}$).
5. Units having a number equal to the integer nearest the mean in each diameter class ($\bar{x}$, $\bar{x}$).
6. Units having a number equal to the integer nearest the mean of X-inch trees and a non-mean of Y-inch trees ($\bar{x}$, non-$\bar{x}$).
7. Units having a non-mean number of X-inch trees and a number equal to the integer nearest the mean of Y-inch trees (non-$\bar{x}$, $\bar{x}$).

Association can also be measured using the bivariate negative binomial series developed by Bates and Neyman (1952). High probabilities obtained in testing the actual values with the expected values are an indication of positive association.

The following is an illustrative example of measuring association by the methods previously mentioned. Table IX is the stem frequency distribution of the 12-inch and 14-inch trees in District 2. Table X is the joint distribution of the 12-inch and 14-inch trees in District 2.

TABLE IX

STEM FREQUENCY OF 12- AND 14-INCH TREES IN DISTRICT 2

Stem Count	DBH Classes	
	12	14
0	66	84
1	45	46
2	30	28
3	13	12
4	10	9
5	7	2
6	3	1
7	2	4
8	7	0
9	0	2
10	4	0
11	1	0

TABLE X

JOINT FREQUENCY TABLE FOR 12- AND 14-INCH TREES IN DISTRICT 2

Stem Count of 12-Inch Trees	Stem Count of 14-Inch Trees									
	0	1	2	3	4	5	6	7	8	9
0	44.	16.	5.	1.	0.	0.	0.	0.	0.	0.
1	22.	12.	9.	1.	1.	0.	0.	0.	0.	0.
2	8.	8.	6.	4.	3.	0.	0.	1.	0.	0.
3	2.	5.	3.	1.	1.	0.	0.	0.	0.	1.
4	3.	4.	2.	1.	0.	0.	0.	0.	0.	0.
5	2.	0.	1.	1.	2.	0.	0.	1.	0.	0.
6	1.	0.	0.	1.	0.	0.	0.	0.	0.	0.
7	0.	0.	0.	0.	0.	0.	1.	1.	0.	0.
8	2.	1.	1.	0.	1.	1.	0.	0.	0.	1.
9	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
10	0.	0.	1.	1.	1.	0.	0.	1.	0.	0.
11	0.	0.	0.	1.	0.	0.	0.	0.	0.	0.

The mean of the 12-inch class is 1.90425 trees per plot while the mean number of 14-inch trees is 1.2553 trees per plot. A 2-count for the 12-inch trees and a 1-count for the 14-inch trees will be used as approximate means. The expected number of plots in each of the seven categories is obtained from the stem frequency distribution tables and the actual number is obtained from the joint frequency table. The seven categories were found to be:

Category	12-inch units	14-inch units	12-inch P	14-inch P	Expectation		Actual units	$\frac{(O-E)^2}{E}$
					P	Units		
1	111	84	.5904	.4468			66	5.897
2	47	58	.2500	.3085	.0771	14	27	12.071
3	111	58	.5904	.3085	.1822	34	14	8.500
4	47	84	.2500	.4468	.1117	21	10	5.761
5	30	46	.1596	.2447	.0391	8	8	0.0
6	30	142	.1596	.7553	.1205	23	22	0.043
7	158	46	.8404	.2447	.2056	39	38	
					1.0000	188	188	

With 4 degrees of freedom, $\chi^2 = 32.279$, P< 0.001, $\chi^2 = 32.279$

Positive association (an overabundance on some plots of the size classes in question)is indicated when the actual number of sample units is greater than expectation in categories 1 and 2, and less than expectation in categories 3 and 4. Negative association is indicated when the reverse occurrs (Walker, 1970).

A bivariate negative binomial fitting is given in Table XI. The upper figure in each cell is the observed value while the lower figure is the expectation derived from the bivariate negative binomial. The computer calculated chi-square value as 6.95, with 9 degrees of freedom. The probability in this case is = .60, an indication of strong positive association. Many of the distributions tested in this project gave equally good results.

TABLE XI

BIVARIATE NEGATIVE BINOMIAL FITTING FOR 12- AND 14-INCH TREES IN DISTRICT 2

Stem Count of 12-Inch Trees (Y)	Stem Count of 14-Inch Trees (X)									
	0	1	2	3	4	5	6	7	8	9
0	44. 46	16. 14	5. 4	1. 1	0. 1	0.	0.	0.	0.	0.
1	22. 21	12. 12	9. 6	1. 2	1. 1	0. 1	0.	0.	0.	0.
2	8. 9	8. 8	6. 5	4. 3	3. 1	0. 1	0. 1	1.	0.	0.
3	2. 4	5. 5	3. 4	1. 2	1. 1	0. 1	0.	0.	0.	1.
4	3. 2	4. 3	2. 2	1. 2	0. 1	0. 1	0.	0.	0.	0.
5	2. 1	0. 2	1. 2	1. 1	2. 1	0. 1	0.	1.	0.	0.
6	1. 1	0. 1	0. 1	1. 1	0. 1	1. 1	0.	0.	0.	0.
7	0.	0. 1	0. 1	0. 1	0. 1	0.	1.	1.	0.	0.
8	2.	1.	1. 1	0. 1	1. 1	1.	0.	0.	0.	1.
9	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
10	0.	0.	1.	1.	1.	0.	0.	1.	0.	0.
11	0.	0.	0.	1.	0.	0.	0.	0.	0.	0.

D.F = 9, $\chi^2 = 6.95$, P .60

A chi-square test for the goodness of fit based on the seven categories confirms the results of the bivariate negative binomial fitting. In most cases it gives a P value close to the value obtained in the bivariate fitting.

Category	Actual units	Expected units	$\frac{(A-E)^2}{E}$
1	66	67	0.0149
2	27	30	0.3000
3	17	16	0.0625
4	10	8	0.5000
5	8	9	0.1111
6	22	20	0.2000
7	38	38	0.0000
		χ^2 =	1.1884

With 4 degrees of freedom, P ~ .85

More experimenting needs to be done with the bivariate negative binomial before its usefulness can be determined. The mathematical equations of the bivariate negative binomial forces positive association for any distribution and, therefore, a poor fit may be due to negative association or a lack of association, or, if positive association exists, it may simply reflect a poor fit of the data. The possibility, however, of predicting both the univariate and bivariate distributions from small sample intensities would provide the forest manager a powerful tool.

CHAPTER VIII

SUMMARY

The purpose of this research were twofold: (1) to determine if the observed size class distributions of the data obtained for this project were in agreement with the univariate negative binomial series and (2) if the data fit well with the negative binomial series to make forecasts of the distribution with reduced sample sizes.

Stem frequency tables were constructed for each 2-inch diameter class in each district of the forest ownership. The population parameters ($\bar{x}$, $\hat{k}$) were computed for each diameter class, as were the negative binomial distributions most adequately describing each distribution. The theoretical distributions were compared with the actual distributions and tested with chi-square. It was found that, with one exception, the actual distributions were in agreement with the negative binomial series.

Forecasts of the distributions using reduced sample sizes were made in two ways. The first involved estimating the mean and $\underline{\hat{k}}$ from the sample and computing a negative binomial distribution based on these estimates. The results obtained were unsatisfactory and the method was abandoned. The second method involved estimating the mean from the reduced samples and combining this with the computed parameter $\underline{\hat{k}}$ to compute the expected negative binomial distribution. The expected distributions were then compared to the actual distributions and tested with chi-square. Samples of 20 percent, 30 percent, and 35 percent were drawn.

The 20 percent sample yielded unsatisfactory results. The 30 percent sample proved much better, but was still considered unsatisfactory. The 35 percent samples, however, consistently gave good results, thus, it was determined that adequate forecasts could be made at this intensity level.

Suggestions were made that any future work concentrate on alternative tests of agreement and the establishment of confidence limits to add reliability to the system. The possibility of forecasting joint distributions (association) from reduced samples, by means of the bivariate negative binomial, was also suggested.

LIST OF SELECTED REFERENCES

Archibald, E. E. A. 1948. Plant Populations, I. A New Application of Neymans Contagious Distribution, Ann. Bot. (NS) 12: 221-235.

Bates, Grace and Jerzy Neyman. 1952 a,b. Contributions to the Theory of Accident Proness: I. An Optimistic Model of the Correlation Between Light and Severe Accidents. II. True or False Contagion. University of California Publications in Statistics. 1, (9) (10) 215-254 and 255-275.

Blackman, G. E. 1935. A Study of the Distribution of Species in Grassland Associations. Ann. Bot. xlix. 749.

Bliss, C. I. 1953. Fitting the Negative Binomial Distribution to Biological Data. Biometrics 9: 176-196.

Clapham, A. R. 1936. Over-Dispersion in Grassland Communities and the Use of Statistical Methods in Plant Ecology. Jour. of Ecology 24: 232-251.

Cole, L. C. 1946. A Study of the Cryptozoa of an Illinois Woodland. Ecological Monographs. 16: 49-86.

Feller, W. 1943. On a General Class of Contagious Distributions. Ann. Math. Statistics 14: 389-400.

Fisher, R. A. 1953. Note on the Efficient Fitting of the Negative Binomial. Biometrics 9: 197-200.

Greig-Smith, P. 1957. Quantitative Plant Ecology. Second Edition. Butterworths Printing Company, Washington, D. C.

Grosenbaugh, L. R. 1952. Plotless Timber Estimates, New, Fast, and Easy. Jour. of Forestry 50: 32-37.

Pielou, E. C. 1960. A Single Mechanism to Account for Regular, Random, and Aggregated Populations. Jour. of Ecology 48: 575-584.

Robinson, P. 1954. The Distribution of Plant Populations. Ann. Bot. (NS) 18: 35-45.

Spurr, S. H. 1952. Forest Inventory. Ronald Press Company, New York.

Walker, N. 1970. Tree Stem Frequency Distribution and Size Class Association in Natural Forest Types. Ph.D. Thesis. N.C. State Univ.

APPENDIX

TABLE XII

ESTIMATED MEANS AND PROBABILITIES OBTAINED IN THE CHI-SQUARE TESTS FOR A NUMBER OF RANDOM SAMPLES DRAWN FROM DISTRICT 1

		Diameter Class								
		4-inch	6-inch	8-inch	10-inch	12-inch	14-inch	16-inch	18-inch	20-inch
Observed	$\bar{x}$	3.1053	2.5439	1.9006	3.9591	2.5439	1.3743	.7076	.3041	.1930
	~P	.35	.45	.60	.95	.3	.70	.70	.70	.60
Sample size and number										
20%-1	$\bar{x}$	3.8529	3.2941	2.0294	3.8824	2.50	1.4706	.50	.3529	.3353
	~P	.40	.25	.80	.80	.20	.80	.005	.40	.40
30%-1	$\bar{x}$	3.2745	2.9412	2.2157	4.3725	2.7451	1.5294	.8039	.2157	.2157
	~P	.40	.30	.60	.90	.30	.80	.50	.50	.60
35%-1	$\bar{x}$	2.850	2.70	2.2167	4.1167	2.6167	1.5167	.8167	.2333	.1833
	~P	.35	.40	.80	.90	.30	.80	.55	.07	.30
35%-2	$\bar{x}$	2.8333	2.2333	1.6167	3.55	2.4167	1.5333	.70	.1883	.1167
	~P	.40	.20	.20	.91	.15	.72	.80	.005	.06
35%-3	$\bar{x}$	2.8333	2.253	1.6233	3.5427	2.60	1.5333	.70	.2057	.1167
	~P	.40	.20	.20	.90	.15	.78	.80	.50	.06
35%-4	$\bar{x}$	2.733	2.30	1.90	3.60	2.65	1.4333	.6333	.2833	.10
	~P	.20	.15	.65	.92	.28	.70	.20	.60	.005

TABLE XIII

ESTIMATED MEANS AND PROBABILITIES OBTAINED IN THE CHI-SQUARE TESTS FOR A NUMBER OF RANDOM SAMPLES DRAWN FROM DISTRICT 2

		Diameter Class							
		4-inch	6-inch	8-inch	10-inch	12-inch	14-inch	16-inch	18-inch
Observed	$\bar{x}$	3.9202	2.877	1.7660	3.0479	1.9043	1.2553	.5372	.2181
	~ P	.73	.01	.35	.5	.2	.3	.45	.30
Sample size and number									
20%-1	$\bar{x}$	4.3784	2.4324	1.5405	2.7027	1.4865	1.1081	.3783	.2432
	~ P	.80	.005	.20	.30	.005	.40	.009	.09
30%-1	$\bar{x}$	3.5893	2.3214	1.4286	2.6607	2.2301	1.2143	.6250	.1250
	~ P	.60	.025	.005	.52	.06	.60	.20	.005
35%-1	$\bar{x}$	3.7121	2.50	1.5909	2.7424	2.1818	1.1818	.6515	.1818
	~ P	.30	.03	.20	.30	.30	.55	.30	.15
35%-2	$\bar{x}$	4.3636	2.7727	1.8333	3.5909	1.5909	1.1970	.4242	.1970
	~ P	.93	.09	.20	.40	.09	.40	.08	.15
35%-3	$\bar{x}$	4.2206	2.7794	1.7941	3.7206	1.6029	1.1324	.4706	.1618
	~ P	.80	.09	.30	.28	.15	.48	.45	.07
35%-4	$\bar{x}$	4.2121	3.4242	2.1364	3.7273	2.0758	1.4091	.5152	.1970
	~ P	.80	.05	.20	.27	.20	.55	.30	.15

TABLE XIV

ESTIMATED MEANS AND PROBABILITIES OBTAINED IN THE CHI-SQUARE TESTS FOR A NUMBER OF RANDOM SAMPLES DRAWN FROM DISTRICT 3

		Diameter Class							
		4-inch	6-inch	8-inch	10-inch	12-inch	14-inch	16-inch	18-inch
Observed	$\bar{x}$	2.6809	2.1986	1.6170	2.4823	1.8794	1.2553	.4539	.1773
	~P	.65	.6	.55	.40	.74	.35	.15	.60
Sample size and number									
20%-1	$\bar{x}$	1.8214	1.5357	1.0357	2.2143	2.00	.7143	.7143	
	~P	.09	.05	.005	.40	.93	.005	.025	
30%-1	$\bar{x}$	2.7143	2.5476	1.8333	2.8095	2.2381	1.3571	.6190	.2143
	~P	.48	.40	.52	.60	.77	.20	.06	.30
35%-1	$\bar{x}$	2.7551	2.5918	1.8163	2.8980	2.3061	1.3469	.5306	.2245
	~P	.45	.45	.53	.55	.55	.09	.15	.30
35%-2	$\bar{x}$	3.1020	2.6327	1.4898	2.5510	1.9184	1.0816	.5510	.2041
	~P	.60	.60	.55	.30	.93	.09	.10	.13
35%-3	$\bar{x}$	3.0816	2.6122	1.5102	2.4490	1.8163	1.0612	.5102	.2041
	~P	.60	.40	.70	.70	.80	.08	.20	.40
35%-4	$\bar{x}$	2.8980	2.2041	1.4082	1.8980	1.2857	1.0204	.4490	.1429
	~P	.70	.70	.40	.02	.005	.09	.15	.35

TABLE XV

ESTIMATED MEANS AND PROBABILITIES OBTAINED IN THE CHI-SQUARE TESTS FOR A NUMBER OF RANDOM SAMPLES DRAWN FROM DISTRICT 4

		Diameter Class							
		4-inch	6-inch	8-inch	10-inch	12-inch	14-inch	16-inch	18-inch
Observed	$\bar{x}$	4.4639	3.5309	2.1186	3.6237	2.2371	1.3814	.6031	.1598
	~P	.15	.43	.15	.80	.45	.30	.55	.30
Sample size and number									
20%-1	$\bar{x}$	3.9655	3.5789	1.8158	2.9737	2.4737	1.6316	.5263	
	~P	.005	.20	.02	.15	.45	.40	.40	
30%-1	$\bar{x}$	3.9655	3.1552	2.2241	3.8276	2.2759	1.4483	.5862	.1207
	~P	.01	.05	.10	.65	.30	.30	.40	.22
35%-1	$\bar{x}$	4.0735	3.3676	2.3235	3.9853	2.3382	1.4559	.6324	.1176
	~P	.025	.28	.07	.65	.20	.30	.50	.20
35%-2	$\bar{x}$	4.1324	4.1324	2.6912	4.1912	2.5147	1.3529	.6765	.1765
	~P	.05	.15	.03	.48	.16	.30	.30	.60
35%-3	$\bar{x}$	3.9559	3.9118	2.4412	3.8382	2.3971	1.4118	.7206	.1765
	~P	.01	.15	.09	.70	.15	.20	.10	.30
35%-4	$\bar{x}$	4.9412	3.2059	1.9412	3.2794	2.0441	1.5147	.6176	.1912
	~P	.05	.40	.09	.45	.08	.40	.40	.35

TABLE XVI

ESTIMATED MEANS AND PROBABILITIES OBTAINED IN THE CHI-SQUARE TESTS FOR A NUMBER OF RANDOM SAMPLES DRAWN FROM DISTRICT 5

		Diameter Class							
		4-inch	6-inch	8-inch	10-inch	12-inch	14-inch	16-inch	18-inch
Observed	$\bar{x}$	5.5204	4.2217	2.4480	3.2851	1.8869	1.0181	.4932	.1991
	~P	.45	.85	.20	.30	.30	.15	.40	.60
Sampel size and number									
20%-1	$\bar{x}$	5.4318	3.4091	1.3636	2.2273	1.3409	1.1591	.4318	
	~P	.3	.025	.005	.005	.005	.15	.10	
30%-1	$\bar{x}$	5.6212	4.50	2.7424	3.5303	1.9242	1.1515	.5909	.2121
	~P	.40	.75	.40	.40	.15	.15	.15	.35
35%-1	$\bar{x}$	5.9221	4.5844	2.7403	3.4935	2.00	1.1169	.5714	.2208
	~P	.60	.77	.13	.30	.40	.10	.20	.13
35%-2	$\bar{x}$	5.2338	4.3247	2.7403	3.0390	2.2857	1.558	.5584	.1688
	~P	.30	.80	.15	.30	.20	.13	.15	.20
35%-3	$\bar{x}$	5.0130	4.1169	2.7663	3.1039	2.2597	1.1039	.5455	.1688
	~P	.60	.80	.09	.30	.20	.20	.20	.20
35%-4	$\bar{x}$	5.0390	3.8052	2.2078	2.7662	1.7662	1.558	.5159	.2078
	~P	.55	.60	.05	.06	.30	.15	.20	.60

VITA

Michael John Dahlem

Candidate for the Degree of

Master of Science

Thesis: LOW INTENSITY FOREST SAMPLING THROUGH USE OF STEM FREQUENCY DISTRIBUTION AND POPULATION PARAMETERS

Major Field: Forest Resources

Biographical:

Personal Data: Born at Clarksville, Arkansas, May 20, 1942, the son of Mr. and Mrs. Alfred R. Dahlem.

Education: Graduated from St. Annes Academy, Fort Smith, Arkansas in 1960; received Bachelor of Science degree, with a major in Forestry, at Oklahoma State University, in May, 1970; completed the requirements for a Master of Science degree in July, 1972.

Professional Experience: Worked as a graduate/research assistant for the Oklahoma State University Forestry Department in the years 1970 through 1972.