Journal of Environmental Science and Engineering B 1 (2012) 82-91
Formerly part of Journal of Environmental Science and Engineering, ISSN 1934-8932

Analysis of Methodology for the Application of Stratified Random Sampling with Optimum Allocation: The Case Study of Forest Bioenergy

M.N. Tsatiris

Department of Forestry and Management of the Environment and Natural Resources, Democritus University of Thrace, Orestiada 68200, Greece

Received: April 6, 2011 / Accepted: May 13, 2011 / Published: January 20, 2012.

Abstract: In this paper, analysis of methodology was realized for the application of stratified random sampling with optimum allocation in the case of a subject of research which concerns the rural population and presents high differentiations among the three strata in which this population could be classified. The rural population of Evros Prefecture (Greece) with criterion the mean altitude of settlements was classified in three strata, the mountainous, semi-mountainous and flat population for the estimation of mean consumption of forest fuelwood for covering of heating and cooking needs in households of these three strata. The analysis of this methodology includes: (1) the determination of total size of sample for entire the rural population and its allocation to the various strata; (2) the investigation of effectiveness of stratification with the technique of analysis of variance (One-Way ANOVA); (3) the conduct of sampling research with the realization of face-to-face interviews in selected households and (4) the control of forms of the questionnaire and the analysis of data by using the statistical package for social sciences, SPSS for Windows. All data for the analysis of this methodology and its practical application were taken by the pilot sampling which was realized in each stratum. Relative paper was not found by the review of literature.

Key words: Analysis of methodology, stratified random sampling with optimum allocation, rural population, forest bioenergy.

1. Introduction

1.1 In General

The prefecture of Evros is the largest in extent and population prefecture of Thrace region (Greece). It has an extent of 4,242 square kilometers and a population of 143,752 residents [1], it northly borders on Bulgaria and Easternly on Turkey. The main branches of economy in the area are agriculture and cattle-breeding, while a significant growth of industry and manufacture is observed in the past few years.

30.5% of total extent of Prefecture is covered by forests [2] and the largest part of annual forest production is fuelwood (62.3%) emanating mainly from deciduous oaks, which is intended for the covering of needs of heating and cooking in households mainly of rural population of Prefecture as well as as a raw material for the production particleboards, fiberboards and paper [3]. The inhabitants of forest areas have the right of free collection of wood harvesting residues for covering of their individual needs for heating and cooking [2].

1.2 Formation of Strata

Stratified random sampling is generally used when the population is heterogeneous or dissimilar where certain homogeneous or similar subpopulations can be isolated (strata) [4-6].

The rural population of Evros Prefecture with criterion the mean altitude of settlements was classified in three strata, the mountainous, semi-mountainous and

Corresponding author: M.N. Tsatiris, assistant professor, Ph.D., main research fields: biomass and bioenergy. E-mail: tsatiris@fmenr.duth.gr.

flat population (according to the distinction of the National Statistical Service of Greece [1]) within the framework of an effort for a better approach of the allocated subpopulations and consequently of the achievement of more accurate estimations of the values of parameters of these subpopulations. This classification was realized with the idea that communities of higher altitudes would have a larger consumption of fuelwood in comparison with communities with lower altitudes and this mainly due to the lower family income in the communities of higher altitudes. Tsatiris [7] states that the Greek forests are situated mainly in mountainous areas, the residents of which have the lowest family income and because of this they consume for covering of their own individual heating needs larger quantities of the free collected fuelwood than the residents of communities with lower altitudes. Consequently, the quantitative variable used to form strata is in the case of this study the family income that is called a factor.

With this stratification we expect to have more homogeneity within the strata so that we need a smaller size of sample and the application of stratified random sampling to be more economic than the simple random sampling. With this stratification we expect to have as much as possible more homogeneous (similar) sampling units within each stratum, whilst on the contrary, among the strata to be as much as possible more non-homogeneous. Homogeneity and lack of homogeneity in these cases are estimated by using of variance. Thus, we expect that with this stratification we will have a low variance within each stratum and as much as possible higher variance among the strata.

2. Analysis of Methodology

2.1 In General

Data were collected by heads of households of mountainous, semi-mountainous and flat rural population of Evros Prefecture through face-to-face interviews and filling in of forms of a questionnaire. The questionnaire contained questions relative to subjects of forest bioenergy.

When the sampling units are human beings, the main methods of collecting information are: (1) face-to-face interviewing; (2) postal surveys; (3) telephone surveys and (4) internet [8, 9]. A questionnaire is a series of questions asked to individuals to obtain statistically useful information about a given topic. Questionnaires are frequently used in social researches. They are a valuable method of collecting a wide range of information from a large number of individuals, often referred to as respondents [10].

The problem of size of sample in the stratified random sampling has two legs. The first leg is reported in the determination of total size of sample (n) for entire the population, while the second leg is reported in the allocation of this size to the various strata (n_K). Both for these two legs exist four ways that can be applied for the determination of the above sizes [5, 11, 12]:

(1) Equal samples in each stratum;

(2) Proportional allocation;

(3) Optimum allocation with equal sampling costs in each stratum;

(4) Optimum allocation with varying sampling costs in each stratum.

In the case of this study, applying stratified random sampling, the size of sample of households was estimated for entire the rural population (n) (Eq. (1)) and afterwards for each stratum separately (n_K) (Eq. (2)) through the way of "Optimum Allocation with Equal Sampling Costs in Each Stratum".

Sampling frame refers to the set of all possible sample units [9, 13-16]. The lists of consumers of domestic electric current were used as a sampling frame in the case of this research. These lists were preferred, because they were more complete and more informed than those of other services, for example phone book, municipal rolls. The selected households (regular and surrogate) more accurately were located through these lists.

2.2 Determination of the Total Size of Sample for the Estimation of Numerical Mean

The Eq. (1) is applied for the determination of total

size of sample (n) for entire the rural population in the stratified random sampling when we want to estimate the numerical mean of population and in the case where the way of "Optimum Allocation with Equal Cost" has been selected [5, 11]:

$$n = \frac{t^2 \left(\sum_{K=1}^{L} N_K S_K \right)^2}{N^2 e^2 + t^2 \sum_{K=1}^{L} N_K S_K^2} \quad (1)$$

Nomenclature:

n: the total size of sample for entire the rural population (households);

N: the total number of households in entire the rural population;

N_K: the number of households in the stratum K (K = 1, 2,…, L);

N_1: the number of households of rural mountainous population;

N_2: the number of households of rural semi-mountainous population;

N_3: the number of households of rural flat population;

L: the number of strata;

The t value is related to the desired probability that the confidence interval will contain the mean of population;

S_K^2: the variance in the stratum K;

Variance is the sum of square of deviations of mean from the individual observations divided by a number by a unit smaller than the total number of observations [12]. The variance of individuals in a population is a measure of the dispersion of individual unit values about their mean. A large variance indicates wide dispersion, while a small variance indicates little dispersion [11, 17];

S_K: the standard deviation in the stratum K (K = 1, 2,…, L). The standard deviation is the square root of the variance;

Variance and standard deviation are indicators of variability among the units of a population. Standard error is a standard deviation but among the estimators (estimates) and not among the individual units of a population [11, 18, 19];

e: the accuracy of estimation.

2.3 Allocation of the Total Size of Sample to Various Strata

The total size of sample (n) should afterwards to be allocated to the three strata. Thus, the way of "Optimum Allocation with Equal Sampling Costs in Each Stratum" was used again. The total size of sample (n) is allocated to the three strata so that we have the lowest variance and moreover the cost of a sampling unit to be equal in all strata. The relative formula by which is estimated the size of sample in each stratum separately (n_K) is the following [5, 11]:

$$n_K = \frac{(N_K)(S_K)}{\sum_{K=1}^{L} ((N_K)(S_K))} \cdot n \quad (2)$$

Nomenclature:

n_K: the size of sample in the stratum K (households) (K = 1, 2,…, L).

The rest symbols have been explained above.

2.4 Effectiveness of Stratification

After the formation of strata and the gathering of data, the effectiveness of stratification should be investigated with the technique of Analysis of Variance. The technique of Analysis of Variance in order to be applied, the following hypotheses (assumptions) should be in force [20-22]:

(1) The K samples should be random and independent. The observations of each sample should be taken so as they are not influenced by those that were taken up to now;

(2) The distribution of populations by which the K samples were taken should be approximate normal;

(3) All populations by which the K samples were taken should have equal variance, σ^2;

(4) The last hypothesis is the null hypothesis that we want to test.

The null hypothesis (Ho) declares that the K populations by which the samples were taken have equal numerical mean:

$$Ho: \mu_1 = \mu_2 = \mu_3$$

The alternate hypothesis H_1 declares that all numerical means are not equal.

Test of homogeneity of variances for the three (3) samples was realized by the test of Levene for the dependent variable "consumption of fuelwood". Norusis [23] states that the test of homogeneity of variances calculates the Levene statistic, testing whether the variance of the dependent variable (e.g. consumption of fuelwood) is equal for all the groups.

The total variance of rural population was separated in two components. The one component is that that exists among the strata while the other one within the strata.

Analysis of variance by a factor (independent variable) was held by the subprogram One-Way ANOVA of SPSS for Windows [23-26]. In this paper, we'll use the One-Way ANOVA procedure. According to Norusis [25] it's called One-Way analysis of variance because cases are assigned to different groups based on their values for one variable. In this paper, we form the groups (strata) based on the values of the variable "family income". The variable used to form groups is the "family income" that is called a factor. If the mean annual family income of strata (factor or independent variable) influences the mean consumption of fuelwood (dependent variable) is tested [27]. All data for the realization of One-Way analysis of variance were taken by a pilot sampling [12, 28-30].

2.5 Choice of Households and Conduct of the Sampling Research

The number of households for example of rural mountainous population results as a sum of the households of all rural mountainous communities. The number of households of the other categories of population results at a proportional way. The households which are included in the lists of consumers of domestic electric current of each one category of population (stratum) were numbered at a serial order. A simple random sample of households was taken by each one stratum. Tables of random numbers and the method of simple random sampling without repetition were used for the choice of simple random sample [12, 31]. Thus, all selected households (regular households) in each sample are different at each other, namely each selected household is taken in the sample only once. Each selected regular household was determined in these lists by the corresponding full name and address.

Surrogate households besides were taken because "deniers'' would result at the conduct of sampling research. The quantity of surrogate households in each stratum was taken equal to 30% of the sample size of each stratum [13, 14, 32]. Tables of random numbers were used for the choice of surrogate households from each stratum [12, 31]. Each selected surrogate household was determined in these lists by the corresponding full name and address.

In all selected households (regular and surrogate), registered letters were sent by which the aim of this research is explained and the date and time at which the interviewer will meet the interviewee in the place of residence of the last one are predetermined [9, 14, 33]. In other words, the conduct of sampling research with the realization of face-to-face interviewing between the researcher and the heads of selected households and filling in of the forms of a questionnaire was decided.

2.6 Control of the Forms of the Questionnaire and Analysis of Data

After the gathering of forms of the questionnaire they were controlled for the detection of by any chance serious lacks. This control aimed at the ascertainment of completeness, consequence and plausibility of the answers in the forms of the questionnaire [9, 10, 14]. The control of completeness aims at the ascertainment of by any chance unanswered questions. The control of consequence becomes through the questions of indirect control by which the by any chance inconsequence of the respondent in questions that measure the same characteristic is ascertained. The control of plausibility aims at the revelation of logical consequence at the answers that are given and it becomes with the

comparison of answers in proportional relative questions of the questionnaire. Moreover, the reliability of the pilot sample of households from each stratum (category of subpopulation) was checked by the errors of pilot sampling (standard errors) [12, 28]. The analysis of data was realized by using the statistical package for social sciences, SPSS for Windows. Specifically, the results (output) which were extracted concern mainly the Descriptive Statistics, Test of Homogeneity of Variances (Levene's test) and One-Way ANOVA.

3. Results and Discussion

3.1 Results of Pilot Sampling

3.1.1 Size of Sample

The following data were taken for the determination of total size of sample of households (n) for entire the rural population:

N: 19,142 households [1];

N_1: 1,798 households [1];

N_2: 3,106 households [1];

N_3: 14,238 households [1];

L: 3 strata;

t: 2 for probability (1-α) % = 95%.

The unknown parameters e, S_K, S_K^2 were determined by data of a pilot sampling which was realized within each stratum [12, 28]. Specifically, they randomly were taken by the [12, 31]:

- Rural mountainous population 43 households (stratum 1);
- Rural semi-mountainous population 37 households (stratum 2);
- Rural flat population 42 households (stratum 3).

The characteristic sizes (output) of quantitative variables: (1) "annual consumption of fuelwood per household"; (2) "annual consumption of charcoal per household"; and (3) "mean annual family income" for each stratum emanating from the pilot sampling are presented in Tables 1, 2 and 3 respectively (Descriptive Statistics). For each stratum, the mean, standard deviation, standard error of the mean, minimum, maximum and a 95% confidence interval for the mean are calculated by the Descriptive Statistics.

More specifically, the quantitative variable "annual consumption of fuelwood per household" was selected, which comparatively with the quantitative variable "annual consumption of charcoal per household" presented higher standard deviation in relation to the mean and afterwards the standard deviation of this variable as it was estimated in each stratum, it was used in the formula (1) as a standard deviation (S_K). The quantitative variable "mean annual family income" was excluded, because it presented excessively high

Table 1 Characteristic sizes of the quantitative variable "annual consumption of fuelwood per household" for each stratum (in tons of fuelwood/year, household) emanating from the pilot sampling.

	N	Mean	Std. deviation	Std. error	95% confidence interval for mean		Minimum	Maximum
					Lower bound	Upper bound		
Rural mountainous	43	3.521	1.1422	0.1741	3.173	3.869	1.60	8.00
Rural semi-mountainous	37	2.353	0.7341	0.1206	2.112	2.594	1.00	4.80
Rural flat	42	0.842	0.3215	0.0496	0.743	0.941	0.00	3.90
Total	122	2.244	0.7358	0.0667	2.111	2.378	0.00	8.00

Table 2 Characteristic sizes of the quantitative variable "annual consumption of charcoal per household" for each stratum (in kilos of charcoal/year, household) emanating from the pilot sampling.

	N	Mean	Std. deviation	Std. error	95% confidence interval for mean		Minimum	Maximum
					Lower bound	Upper bound		
Rural mountainous	43	1.724	0.7341	0.1119	1.500	1.948	0.00	25.00
Rural semi-mountainous	37	1.655	0.4213	0.0692	1.516	1.793	0.00	14.00
Rural flat	42	1.411	0.1124	0.0173	1.376	1.446	0.00	11.00
Total	122	1.595	0.4252	0.0384	1.518	1.672	0.00	25.00

Table 3 Characteristic sizes of the quantitative variable "mean annual family income" for each stratum (in euros of the year 2009) emanating from the pilot sampling.

	N	Mean	Std. deviation	Std. error	95% confidence interval for mean		Minimum	Maximum
					Lower bound	Upper bound		
Rural mountainous	43	4704	1046	159.514	4384	5023	2688	8064
Rural semi-mountainous	37	5371	1315	216.186	4938	5803	2688	9132
Rural flat	42	7705	1713	264.323	7176	8233	2688	13100
Total	122	5939	1357	122.857	5693	6184	2688	13100

values of standard deviation. The standard deviation for which reference becomes in this point is that among the units of each stratum.

The accuracy of estimation (e) was taken equal to 2% of the numerical mean of the entire rural population [31, 34], namely e=0.0448. Because the sampling fractions $f = \frac{n}{N}$ were smaller than 5%, the second part of the denominator in the Eq. (1) was ignored [11, 12].

Nomenclature:

n: the number of observations;

N: the size of "population"-number of households in all strata.

Thus, the Eq. (1) is simplified in:

$$n = \frac{t^2 \left(\sum_{K=1}^{L} N_K S_K \right)^2}{N^2 \cdot e^2} \quad (3)$$

The Eq. (3) in the case of this research analytically is presented as follows:

$$n = \frac{t^2 (N_1 S_1 + N_2 S_2 + N_3 S_3)^2}{N^2 \cdot e^2}$$

Replacing the data in this analytic formula we have:

$$n = \frac{2^2 \times (1{,}798 \times 1.1422 + 3{,}106 \times 0.7341 + 14{,}238 \times 0.3215)^2}{(19{,}142)^2 \times (0.0448)^2} = 432$$

Thus, the total size of sample for entire the rural population amounted in n = 432 households. Afterwards the total size of sample (n) was allocated to the three (3) strata so that we have the lowest possible standard error and moreover the cost of a sampling unit to be equal in all strata. Thus, the size of sample in each stratum (nK) was estimated, which has as follows:

Stratum 1, Rural Mountainous Population, n_1 = 100 households;

Stratum 2, Rural Semi-mountainous Population, n_2 = 110 households;

Stratum 3, Rural Flat Population, n_3 = 222 households;

Total (Rural Population), n = 432 households.

3.1.2 Effectiveness of Stratification

The effectiveness of stratification was investigated with the technique of Analysis of Variance. The test of homogeneity of variances for the three (3) samples was realized by the test of Levene for the consumption of fuelwood (Table 4).

From Table 4, it appears that the statistical value of Levene test is 0.362 for the consumption of fuelwood with 2 and 119 degrees of freedom and a probability (significance) of 0.658. According to Einspruch [35], because this probability is higher than the usually used of 0.05, it is concluded that the three samples have equal variances, that is to say homogeneity of variances exists, fact that allows us to apply the technique of analysis of variance.

Thus, the table of One-Way analysis of variance for the dependent variable "consumption of fuelwood" is presented below (Table 5).

In order to obtain the output of One-Way ANOVA table, we selected the "consumption of fuelwood" as a dependent variable and the "family income" as a factor (independent variable) in the One-Way ANOVA dialog box.

We test the null hypothesis (H_0). The null hypothesis (H_0) declares that the three (3) subpopulations by which the samples were taken have equal numerical mean:

$$H_0: \mu_1 = \mu_2 = \mu_3$$

The alternate hypothesis (H_1) declares that all numerical means are not equal.

Table 4 Test of homogeneity of variances for the consumption of fuelwood by using of Levene test (for this test the data concern the quantitative variable "annual consumption of fuelwood per household" for each stratum, in tons/year, household, emanating from the pilot sampling).

	Levene statistic	df1	df2	Sig.
Consumption of fuelwood	0.362	2	119	0.658

Table 5 One-Way analysis of variance for the consumption of fuelwood.

		Sum of Squares	df	Mean Square	F	Sig.
Consumption of fuelwood	Between Groups	426.237	2	213.118	328.379	0.000
	Within Groups	77.344	119	0.649		
	Total	503.581	121			

The F ratio which is equivalent to the variance among the strata divided by the variance within the strata has in this case the value F = 328.379. The critical value of $F_{.05, 2,119}$ the given from the tables is 3.07, for a level of significance of α=0.05 and 2 (K-1 = 3-1) and 119 (n-K = 122-3) degrees of freedom, that is to say F = 328.379 > $F_{.05, 2,119}$ = 3.07 and in consequence the null hypothesis is rejected and we accept the alternate hypothesis (H_1), matter that means that all numerical means are not equal [27]. The acceptance of alternate hypothesis (H_1) means that the mean annual family income of the strata (factor or independent variable) influences the mean consumption of fuelwood (dependent variable). This confirms that the choice of One-Way ANOVA was successful.

Tsatiris [7] states that between the mean consumption of fuelwood and the mean annual family income in seven categories of population of Thessaly region exists a very powerful negative linear correlation, fact that means that while the mean annual family income of the strata increases, the mean consumption of fuelwood per year and household is decreased and this effect is very powerful. Thus, while the mean annual family income of the strata increases, the households of these strata tend to consume a larger quantity of alternate fuels, as petroleum, liquid gas and electric energy with reference to the satisfaction of their heating needs. Fuelwood could be classified as an inferior virtuous and as a cheap fuel which is mainly consumed by categories of population or households of low income.

From the Table 5 of One-Way analysis of variance for the consumption of fuelwood, it results that the variance among the strata found to constitute 84.6% of the total variance (very high), whilst the variance within the strata amounted in 15.4% of the total variance (very low). That means that 84.6% of the total variance due to the differentiation among the three strata and only 15.4% of the total variance due to the variance within the strata. The variance of the stratified sample is measured by the variance within the strata, which significantly has been decreased by the division of rural population in three strata. From these, it results that the stratification was successful [12, 36].

3.1.3 Choice of Households and Conduct of the Sampling Research

The number of households that were selected through the lists of consumers of domestic electric current is presented in Table 6.

Afterwards the realization of the first phase of

Table 6 Number of selected regular and surrogate households.

Stratum	Category of subpopulation	Size of sample, n_K (regular households)	Surrogate households
1	Rural mountainous population	100	30
2	Rural semi-mountainous population	110	33
3	Rural flat population	222	67
Total		432	130

face-to-face interviews, the total "refusals" amounted in the various strata from 4.1%-4.9% of the initially selected households. These "refusals" due mainly to direct refusals (refusal for interview) and to absentees. In second phase, the void that resulted from the "deniers" was covered with the realization of face-to-face interviews in surrogate households which had been selected for this purpose by the corresponding strata.

The relatively large quantity of the required for gathering information and the possibility of direct clarification of questions contributed in the appointment of face-to-face interviewing as the most reliable and suitable technique of information collection in this case, although that the face-to-face interviews were expensive in money and time, where it concerned in the time of locomotion, approach of the area of research and interview of the respondents. With the method of face-to-face interviewing, a relatively high percentage of response from the side of the respondents was achieved, fact which would not happen through the postal returned questionnaire survey, if we furthermore count in the educational level particularly in the case study of rural population [9, 13, 14].

Thus, 100 forms of the questionnaire were filled in by the first stratum, 110 forms of the questionnaire by the second stratum and 222 forms of the questionnaire by the third stratum. 432 forms of the questionnaire totally were filled in by all the strata. After the gathering of 432 forms of the questionnaire, control of these forms took place. A small number of individual questions were unanswered and were considered as invalid.

The errors of pilot sampling or standard errors of the quantitative variable "annual consumption of fuelwood per household" emanating from the data of the pilot sampling and for each pilot sample of households were ranged from 0.04 to 0.17 which are considered as slight ones [12, 15, 28-30] (see Table 1).

4. Conclusions

In this paper, analysis of methodology was realized for the application of stratified random sampling with optimum allocation in the case of a subject of research which concerns the rural population and presents high differentiations among the three strata in which this population could be classified. The rural population of Evros Prefecture (Greece) with criterion the mean altitude of settlements was classified in three strata, the mountainous, semi-mountainous and flat population within the framework of an effort for a better approach of the allocated subpopulations and consequently of the achievement of more accurate estimations of the values of parameters of these subpopulations. The quantitative variable used to form strata is in the case of this study the family income that is called a factor. With this stratification we expect to have more homogeneity within the strata so that we need a smaller size of sample and the application of stratified random sampling to be more economic than the simple random sampling. Specifically, the methodology of application of stratified random sampling with optimum allocation is analyzed for the estimation of mean consumption of forest fuelwood for covering of heating and cooking needs in households of mountainous, semi-mountainous and flat rural population of Evros Prefecture.

The analysis of this methodology includes: (1) the determination of total size of sample for entire the rural population and its allocation to the various strata; (2) the investigation of effectiveness of stratification with the technique of analysis of variance (One-Way ANOVA); (3) the conduct of sampling research with the realization of face-to-face interviews in selected households; and (4) the control of forms of the questionnaire and the analysis of data by using the statistical package for social sciences, SPSS for Windows. Specifically, the results (output) which were extracted concern mainly the Descriptive Statistics, Test of Homogeneity of Variances (Levene's test) and One-Way ANOVA.

All data for the analysis of this methodology and its practical application were taken by the pilot sampling

which was realized in each stratum. Relative paper was not found by the review of literature.

Applying stratified random sampling, the total size of sample of households was estimated for entire the rural population and afterwards for each stratum separately through the way of "Optimum Allocation with Equal Sampling Costs in Each Stratum". The total size of sample (n) was allocated to the three (3) strata so that we have the lowest possible standard error and moreover the cost of a sampling unit to be equal in all strata.

The effectiveness of stratification was investigated with the technique of analysis of variance. The Test of Homogeneity of Variances for the three (3) samples was realized by the test of Levene for the consumption of fuelwood. The statistical value of Levene test is 0.362 for the consumption of fuelwood with 2 and 119 degrees of freedom and a probability of 0.658. Because this probability is higher than the usually used of 0.05, it is concluded that the three samples have equal variances, that is to say homogeneity of variances exists, fact that allows us to apply the technique of analysis of variance. In this paper, we used the One-Way ANOVA procedure. Analysis of variance by a factor (independent variable) was held by the subprogram One-Way ANOVA of SPSS for Windows. We formed the groups (strata) based on the values of the quantitative variable "family income" that is called a factor. In order to obtain the output of One-Way ANOVA table, we selected the "consumption of fuelwood" as a dependent variable and the "family income" as a factor (independent variable) in the One-Way ANOVA dialog box. In other words, if the mean annual family income of strata (factor or independent variable) influences the mean consumption of fuelwood (dependent variable) was tested. We tested the null hypothesis and we proved that the null hypothesis is rejected while the alternate hypothesis be accepted. The acceptance of alternate hypothesis means that the mean annual family income of the strata influences the mean consumption of fuelwood. This confirms that the choice of One-Way ANOVA was successful.

From the table of One-Way analysis of variance for the consumption of fuelwood, it results that the variance among the strata found to constitute 84.6% of the total variance (very high), whilst the variance within the strata amounted in 15.4% of the total variance (very low). That means that 84.6% of the total variance due to the differentiation among the three strata and only 15.4% of the total variance due to the variance within the strata. The variance of the stratified sample is measured by the variance within the strata, which significantly has been decreased by the division of rural population in three strata. From these, it results that the stratification was successful.

The lists of consumers of domestic electric current were used as a sampling frame in the case of this research. The selected households (regular and surrogate) more accurately were located through these lists. Tables of random numbers were used for the choice of regular and surrogate households from each stratum. Each selected regular and surrogate household was determined in these lists by the corresponding full name and address. In other words, the conduct of sampling research with the realization of face-to-face interviewing between the researcher and the heads of selected households and filling in of the forms of the questionnaire was decided. Thus, data were collected by heads of selected households of mountainous, semi-mountainous and flat rural population of Evros Prefecture through face-to-face interviews and filling in of forms of the questionnaire.

After the gathering of forms of the questionnaire they were controlled for the detection of by any chance serious lacks. This control aimed at the ascertainment of completeness, consequence and plausibility of the answers in the forms of the questionnaire. Moreover, the reliability of the pilot sample of households from each stratum (category of subpopulation) was checked by the errors of pilot sampling (standard errors). This analysis could constitute a useful guide in the case of application of stratified random sampling with

optimum allocation in matters of social researches that require the realization of polls in a population which lives in a relatively wide geographic region where this population could be classified in more than one different categories of subpopulations (strata) on the basis of some feature (criterion).

References

[1] National Statistical Service of Greece, Households of rural flat, semi-mountainous and mountainous areas, National Census for Greece on March 18th 2001, Prefecture of Evros, Athens 5 (2005) 125-137.

[2] Results of the First National Inventory of Greek Forests, Ministry of Agriculture of Greece, Athens, 1989, p. 486.

[3] Forest Commission of Evros Prefecture, Forest Production of Evros Prefecture, Alexandroupolis, 2008, p. 128.

[4] L. Kish, Survey Sampling, John Wiley, New York, 1965, p. 643.

[5] W. Cochran, Sampling Techniques, 3rd ed., John Wiley and Sons, New York, 1977, p. 428.

[6] H. Schreuder, T. Gregoire, G. Wood, Sampling Methods for Multiresource Forest Inventory, John Wiley and Sons, New York, 1993, p. 464.

[7] M. Tsatiris, Forecasting of supply and demand of forest biomass for energy production: The case study of Thessaly region, Doctoral Dissertation, Department of Forestry and Natural Environment, Aristotelian University, Thessaloniki, 2001, p. 272.

[8] J. Bell, Methodological Planning of Social Research, The Gutenberg Press Co., Athens, 1997, p. 326.

[9] G. Siardos, Methodology of Agricultural Sociological Research, The Ziti Press Co., Thessaloniki, 1997, p. 367.

[10] C. Javeau, The Research by Using of a Questionnaire, The Typothito Press Co., Athens, 1996, p. 412.

[11] F. Freese, Elementary forest sampling, in: Agricultural Handbook No. 232 of USDA Forest Service, USA, 1962, p. 91.

[12] K. Matis, Forest Sampling, Aristotle University Publications, Thessaloniki, 1988, p. 200.

[13] V. Filias, An Introduction to the Methodology and Techniques of Social Research, The Gutenberg Press Co., Athens, 1977, p. 689.

[14] G. Daoutopoulos, Methodology of Social Researches in the Rural Area, Aristotle University Publications, Thessaloniki, 1994, p. 287.

[15] E. Johnson, Forest Sampling Desk Reference, CRC Press, Florida, 2000, p. 1008.

[16] D. Mandallaz, Sampling Techniques for Forest Inventories, Chapman and Hall & CRC, Boca Raton, 2008, p. 421.

[17] G. Snedecor, W. Cochran, Statistical Methods, 7th ed., The Iowa State University Press, Iowa, USA, 1976, p. 586.

[18] D. Psoinos, Statistics, The Ziti Press Co., Thessaloniki, 1999, p. 396.

[19] G. Halkos, Statistics, The Typothito Press Co., Athens, 2000, p. 309.

[20] A. Katos, Statistics, The Observer Press Co., Thessaloniki, 1984, p. 708.

[21] K. Matis, Forest Biometrics I. Statistics, Aristotle University Publications, Thessaloniki, 1989, p. 358.

[22] E. Bora-Senta, H. Moisiadis, Applied Statistics, The Ziti Press Co., Thessaloniki, 1997, p. 271.

[23] M. Norusis, SPSS 16.0, Guide to Data Analysis, Prentice Hall, New Jersey, 2008, p. 637.

[24] SPSSInc., SPSS 16.0 Brief Guide, Publisher: SPSSInc., Chicago, 2007, p. 171.

[25] M. Norusis, SPSS Statistics 17.0. Guide to Data Analysis, Prentice Hall, New Jersey, 2008, p. 467.

[26] J. Kulas, SPSS Essentials, Managing and Analyzing Social Sciences Data, JOSSEY-BASS, A Wiley Imprint, San Francisco, 2009, p. 432.

[27] K. Bagiatis, N. Tsantas, The Statistics by Using SPSS, Aristotle University Publications, Thessaloniki, 1993, p. 234.

[28] A. Gofas, The Sampling in Forestry and Agronomy, Aristotle University Publications, Thessaloniki, 1970, p. 376.

[29] K. Dalianis, Designing and Analysis of Experiments, Agricultural University Publications, Athens, 1984, p. 586.

[30] H. Damianou, Sampling Methodology, Techniques and Applications, The Ethra Press Co., Athens, 1995, p. 365.

[31] K. Matis, Personal Communication, Thessaloniki, 2007.

[32] G. Daoutopoulos, Personal Communication, Thessaloniki, 2007.

[33] M. Tsatiris, An. Papastavrou, Ath. Karameris, D. Pallis, Organization and self-financing of the city parks of Thessaloniki for recreational purposes, Medit. Journal of Economics, Agriculture and Environment, Edagricole, Bologna 4 (1997) 30-33.

[34] G. Brister, R. Schultz, Probability level, accuracy and sample size in forest sampling, Forest Resource Paper 89, Georgia Forest Resource Council, Macon, Georgia, 1977, p. 96.

[35] E. Einspruch, An Introductory Guide to SPSS for Windows, Sage Publications, London, 1998, p. 147.

[36] M. Prodan, Forest Biometrics, Pergamon Press, Oxford, 1968, p. 447.