

Comparison of Sampling Methods for Annual Industry and Service Statistics Survey by TURKSTAT

Sibel ŞAHİN*¹, Süleyman ALPAYKUT²

¹Dokuz Eylül University, Institute of Natural and Applied Sciences, Statistics Department, 35160, İzmir

²Dokuz Eylül University, Faculty of Science, Statistics Department, 35160, İzmir

(Alınış / Received: 28.04.2017, Kabul / Accepted: 25.12.2017, Online Yayınlanma/ Published Online: 09.02.2018)

Keywords

Stratified sampling,
Poisson sampling,
Simple random sampling,
Systematic sampling,
Annual statistics estimations

Abstract: The Annual Industry and Service Statistics is one of the largest surveys, conducted by Turkish Statistical Institute, which aims to determine changes in economic structure in Turkey. Both full enumeration and sampling methods are used in this survey. Nevertheless, the percentage of full enumeration increases every year. Even though efforts have been made in order to be used administrative records in recent years, this could not satisfy all of the necessary information needed. Hence, it is believed that there is a requirement to decrease the size of the survey. In this study, it is aimed to propose a sampling method for part of the Annual Industry and Service Statistics Survey conducted with the enumeration and to compare the suggested methods. For that purpose, in the first phase, stratified sampling is used and then the comparison is made by using three different sampling methods within the strata, namely poisson, systematic and simple random sampling. The size of the survey is reduced by using sampling methods, but the economic activity classification together with the level of estimation to the regions increase. It is concluded that the best estimations and minimum variances are obtained when poisson and simple random sampling methods are applied together.

TÜİK Yıllık Sanayi ve Hizmet Araştırması için Örneklem Metotlarının Karşılaştırılması

Anahtar Kelimeler

Tabakalı örnekleme,
Poisson örnekleme,
Basit Rasgele örnekleme,
Sistematik örnekleme,
Yıllık istatistik tahminleri

Özet: Yıllık Sanayi ve Hizmet İstatistikleri, Türkiye'deki ekonomik yapıdaki değişiklikleri ortaya koymak için Türkiye İstatistik Kurumu tarafından yapılan en büyük araştırmalardan biridir. Bu araştırmada tamsayım ve örnekleme yöntemleri bir arada kullanılmaktadır. Ancak, her yıl tamsayım oranı artmaktadır. Son yıllarda bu araştırma için idari kayıtların kullanılması yönünde çalışmalar yapılsa da, bu ihtiyaç duyulan tüm bilgiyi sağlamayabilir. Dolayısıyla Yıllık Sanayi ve Hizmet İstatistikleri araştırmasının hacminin küçültülmesine ihtiyaç olduğu düşünülmektedir. Bu çalışma Yıllık Sanayi ve Hizmet İstatistikleri araştırmasının tamsayım ile yapılan kısmı için örnekleme metodu önermek ve önerilen metotları karşılaştırmak için yapılmıştır. İlk aşamada tabakalı örnekleme kullanılmıştır ve tabakalar içinde üç değişik örnekleme metodu, poisson, sistematik ve basit rastgele örnekleme kullanılarak karşılaştırması yapılmıştır. Örnekleme metotları kullanılarak araştırmanın hacmi azaltılmış ama ekonomik faaliyet sınıflaması ve bölgelere göre tahmin seviyesi artmıştır. En yakın tahminler ve en küçük varyanslar poisson ve basit rastgele örnekleme metotlarının birlikte kullanılmasından elde edilmiştir.

1. Introduction

The Annual Industry and Service Statistics Survey is one of the largest sample size of surveys conducted by the Turkish Statistical Institute (TURKSTAT). The main purpose of this survey is to determine changes

in the social and economic structure in the country. This survey is conducted in every European Union country. The countries, both members of the Union and the candidates to the Union, send their own results to the Statistical Office of the European Union (EUROSTAT) at the end of the survey. Each country

publishes own results, and EUROSTAT shares all the countries' results through its website. In order to compare results, the questionnaire contains common questions for all of the countries; however, it is also necessary for the local survey to contain additional questions. The frame of this study is based on the business registers of the TURKSTAT, and those registers have used some administrative records.

In recent years there have been some studies used administrative records for the Annual Industry and Service Statistics without any fieldwork, but, the information obtained solely from administrative records does not wholly satisfy the information needed for this survey. As Brick [1] says, "The purpose of the administrative records may not require the same level of quality as is needed for sampling purposes." Furthermore, the quality of data obtained from administrative records is also questionable. Thus, an additional mini-survey is needed to get information that could not be obtained from administrative records.

Full enumeration and sampling methods are both used for the Annual Industry and Service Statistics. Each year there exists some changes, but generally 60-65% of the frame consists of full enumeration. Some activity codes must be used with full enumeration because of the small size, but the others may be estimated by using statistical methods within a short time. This is one of the purposes of this study. As Brick [1] mentions, "The twentieth century saw a dramatic change in the way information was generated as probability sampling replaced full enumeration."

Another purpose of this study is to compare the suggested sampling methods. Mostly, stratified sampling is used. However, in some strata, full enumeration suggests due to the small population size. Except for these strata, simple random, systematic, and poisson sampling methods are used within the strata and then the results of these sampling methods are compared.

Currently, the results of this survey are given as the NACE Rev 2.2 classification (Nomenclature of Economic Activities) at the four-digit level for Turkey and NUTS2 (Nomenclature of territorial units for statistics) at the two-digit level for the regions. Giving NUTS2 estimations in four digits requires a much larger sample size in the current structure and it requires additional time, cost, and labor. Another purpose of this study is to give the results at four digits level NACE Rev 2.2 codes not only for Turkey but also for NUTS2 regions. This is important for determining regional policies and making decisions. This information is needed in the face of regional developments and it is an aid tool increasing regional policies. In the future, it is expected that it could be possible to discuss about the estimations given for NUTS3 (province) level.

The data used in this study is micro data and belongs to TURKSTAT. The allowance to be accessed and the usage of micro data depend on a protocol signed between the user and TURKSTAT. Use of direct results obtained from data by this protocol is restricted. So, while calculated statistics could be given in this study, unfortunately the value of the parameters could not be given due to the restriction mentioned.

2. Material and Method

2.1. Description of data

Approximately half of the total turnover is supplied from approximately 5-7% of enterprises as shown in Table 1. This information is calculated from data obtained from the database on the TURKSTAT website. The number of the enterprises is relatively small, and any change in the structure of these enterprises directly affects the economic structure. This importance and the relatively small number are the reasons why full enumeration is suggested for enterprises having more than 250 employees. This separation is also compatible with the European Union practices. In fact, Giovanninni [2] says one should regard "enterprises with fewer than 250 employees as small and medium-sized enterprises."

Table 1. Percentage of enterprises and turnover according to size class (number of people employed) year by year (2009-2013)

	Size Class	2009	2010	2011	2012	2013
Enterprise	20-249	93%	94%	95%	95%	94%
	>250	7%	6%	5%	5%	6%
	Total	100%	100%	100%	100%	100%
Turnover	20-249	45%	50%	49%	50%	51%
	>250	55%	50%	51%	50%	49%
	Total	100%	100%	100%	100%	100%

There are 514 NACE Rev 2.2 classes (4 digits) in the data. 387 of them have frequency smaller than 200 in country total. It is hard to divide 26 strata, so these activity codes are accepted as full enumeration data. For the remaining 127 activity codes, the sample size is calculated, but the results of six of them are almost equal to the population size, so it is decided to accept these codes as a full enumeration.

Before selecting samples, outlier values are determined and the data is removed. The values, far from the mean falling out of minus or plus 3σ , are accepted as outliers in this study. The reason of this elimination is due to the fact that the interval between -3σ and $+3\sigma$ covers 99.7% of all data and achieves a desired minimum data loss as much as

possible. The number of outliers to which are added to the full enumeration data is 1,520 over 110,420 units.

Each of the NACE Rev 2.2 (4-digits) class is accepted as a discrete population. There are 121 discrete populations in this study. Besides, there are 26 NUTS2 regions in Turkey. Each of the NUTS2 regions have similar economic and geographic characteristics within itself, but at the same time they are significantly different from the other regions. This is one of the reasons of the fact that stratified sampling is used. Each population has 26 strata consisting of 26 regions. Scheaffer, Mendenhall, & Ott [3] say that "generally, more than five or six strata are not chosen when using this method", but in the Annual Industry and Service Statistic Survey, it is asserted that estimations must be given for all regions and all NACE Rev 2.2 (2-digits) divisions. The reasons of having 26 strata and not including cluster sampling method in the study can be summarized as follows: One of the aims of this study is to give the estimations for all regions and all NACE Rev 2.2 (4 digits) classes instead of just NACE Rev 2.2 (2-digits). Scheaffer, Mendenhall, & Ott [3] point out that by using the stratified sampling method, "separate estimates can be obtained for individual strata without selecting another sample."

Samples are selected in different ways: simple random, poisson, or systematic sampling. The activity records within the strata, having a coefficient of variation more than 2.5, are sampled through those mentioned ways and then the results are compared. Simple random sampling in the strata is used for all other activity records.

There are many variables in the Annual Industry and Service Statistics. Approximately 200 cells are filled. Five of the most important and published variables are chosen for estimations:

D12110: Turnover
 D12120: Production Value
 D12150: Value-added at factor cost
 D13110: Total purchases of goods and services
 D16110: Number of people employed

The variable D12110 turnover is used for calculation. Variables codes like D12110 are defined from EUROSTAT and taken from the EUROSTAT website. It is a standard code for all the countries in the European Union; that is, D12110 means turnover in Turkey, and it also means turnover in France, vice versa.

2.2. Sample size

Chambers [4] indicates, "In practice, surveys are concerned with many population variables. However, most of the theory for sample surveys is developed

for a small number of variables, typically one or two". As he says, in this study, only one variable should be chosen to determine sample size.

Table 2. Sample sizes calculated from variables (for year 2013)

	D12110 with B=200,000	D12120 with B=100,000	D12150 with B=50,000	D13110 with B=200,000	D16110 with B=2
Ni	ni	ni	ni	ni	ni
84,555	52,663	36,071	12,763	36,220	2,595

Firstly, the sample size is calculated for all five variables that would be estimated. In Table 2 the sample sizes calculated for D12110, D12120, D12150, D13110, and D16110 variables is given. The largest sample size is obtained for the D12110 variable, and also turnover is one of the most important economic indicators. So, the D12110 variable is chosen for calculations, but the results are given for all five variables. Calculations are made by formulas of stratified sampling and Neyman allocation. Supposed cost is not important, but variances are considered as the opposite. Therefore, it is decided that the Neyman allocation is the most appropriate allocation. Yamane [5] shows that the efficiency of the Neyman allocation is more than the optimum allocation if sizes and variances of strata have large differences. There are also some studies about how to choose allocation methods. Mathew, Sola, Oladiran, & Amos [6] and Barnabas & Sunday [7] study the efficiency of allocation methods; both studies conclude that the Neyman allocation is the most efficient allocation. Winkler [8] says, "The Neyman allocation is known to be theoretically optimal in comparison with proportional allocation".

The sample sizes are calculated by formulas of the Neyman allocation in the strata using poisson, simple random, and systematic sampling because of the randomness of the sample size in poisson sampling and the need for a fixed sample size.

Cochran [9] point out that "The specification of the degree of precision wanted in the results is an important step". Since the variance in the data is sometimes very large, this causes large bounds, so to maintain the sensitivity, after many trials it is decided to use another calculation method instead of variance. The bound on the error of estimation is changed in each of the activity codes, but a standard is needed for every activity code using the previous year data (in this study, data from the year 2012). 5% of the turnover mean is calculated. If this value is greater than or equal to 150,000 TL, the bound on the error of estimation is accepted as 200,000. If this value is smaller than 150,000 TL, the calculated value is rounded down. For example, if the calculated value is 125,000, the bound on the error of estimation is accepted as 120,000.

The approximate sample size is

$$n = \frac{\sum_{i=1}^L N_i^2 \sigma_i^2 / w_i}{N^2 D + \sum_{i=1}^L N_i \sigma_i^2} \quad (1)$$

where w_i is the fraction of observations allocated to stratum i , and σ_i^2 is the population variance for stratum i . Since the costs per observation are ignored, w_i is

$$w_i = \frac{N_i \sigma_i}{\sum_{i=1}^L N_i \sigma_i} \quad (2)$$

and

$$D = \frac{B^2}{4} \quad (3)$$

when estimating μ

$$D = \frac{B^2}{4N^2} \quad (4)$$

when estimating τ , where B is the bound on the error of estimation.

The stratum size is

$$n_i = n w_i \quad i = 1, 2, \dots, L. \quad (5)$$

2.3. Poisson sampling applications

Inside strata, simple random, systematic and poisson sampling are used. Except for poisson sampling, other methods are well-known and renowned methods. As Lohr [10] says, a simple random sample "provides the theoretical basis for the more complicated forms". For this reason, only the poisson sampling applications and their formulas are given in this article.

Poisson sampling was introduced into the literature by Hájek [11,12]. Williams, Schreuder, & Terraza [13] define poisson sampling "as a sampling design in which the sample units have unequal probabilities of selection, π_i . In addition, the units in the population are independent and the sample size, n , is a random variable" using Hájek's work. Aires' [14] definition is as follows: "A poisson sample may be realized by using N independent Bernoulli trials to determine whether the individual under consideration is to be included in the samples or not."

When using poisson sampling, deciding on the sample size presented some difficulties due to the randomness of the sample size. n may take values between 0 and N , according to inclusion probabilities and random numbers which are used for calculations. For that reason, it is decided to use Conditional Poisson Sampling, since it has a fixed n . Grafström [15] says that "if a fixed sample size n is desired, it is possible to generate poisson samples and to accept

the sample only if the sample size is n . The resulting design is called conditional poisson sampling". Also, Grafström [16] define conditional poisson sampling as "a modification of Poisson sampling. Each unit i in the population is included with a given probability π_i but only samples of size n are accepted." In conditional poisson sampling, n is fixed, and it is tried to find this fixed n . In this case the number of trials is uncertain. For example, finding 100 samples with size two and population size 14, 342 samples must be drawn. 100 of 342 has the sample size two, and the others' size changes between zero and 14.

A characteristic which is easily observed or previously known and existing in each unit of the population (x_i) is selected. This value and the previously decided n value are used to calculate inclusion probabilities. Ghosh & Vogt [17] give the definition of inclusion probability as "the probability that an individual unit will be in the final sample." If a unit has to be in the sample, its inclusion probability is determined as 1 without making any calculations. Saavedra [18] says, "There is no known analytic formula that permits us to calculate probabilities of selection". The most important point before selecting a sample is to decide the x_i values. Williams, Ebel, & Wells [19] say, "In the development of poisson sampling, it was mentioned that the characteristic x_i is chosen so that it is positively correlated with y_i ", and Brewer, Early, and Hanif [20] also say, "If the y_i are roughly proportional to the π_i , it is more efficient for samples of any size." Lundquist [21] define auxiliary variables as "variables which are not our primary interest, but it is reasonable to assume they are connected to our study variable in some way." In this study, x_i value is a kind of transformation of the previous year's mean and variance of the turnover data. Then, inclusion probabilities π_i 's are calculated using the formula $\pi_i = \frac{n x_i}{X} \quad i = 1, \dots, N$ and $X = \sum_{i=1}^n x_i$. For determining the inclusion probabilities, a dummy variable (x_i) that is calculated by using the mean and variance of D12110 (turnover) variable is used. Here, y_i is the turnover value, and x_i is the dummy variable that is produced via turnover value's mean and variance for the previous year.

Then, random numbers which come from uniform distribution are generated, and every unit is assigned to a random number. If the random number is smaller than the inclusion probability of case i , case i is selected in the sample. If the number of selected units is smaller or larger, second order inclusion probabilities are calculated by using the new n value which is the number of selected cases in the first iteration. This procedure is repeated until a constant n value is obtained. That means that at this point, whatever the iteration number is, n cannot be changed. If this n value equals to the previously decided n value, the selected units consist of the sample, and it can be calculated with statistics from this sample. But, if this n value does not equal to the

previously decided n value, this sample was rejected. In this case, new random numbers are generated, and all the procedure are done again.

Särndal, C.E., Swenson, B. & Wretman, J. [22] gives the formulas of estimators of poisson sampling as follows;

The estimator of population total $\tau = \sum_U y_i$ is

$$\hat{\tau} = \sum_s \check{y}_i = \sum_s \frac{y_i}{\pi_i} \quad (6)$$

Hajek [12] say Horwitz and Thompson show that $\sum_{ies} (y_i/\pi_i)$ is an unbiased estimator of population total if $\pi_i > 0, i=1, \dots, N$, for any sampling design.

$$V(\hat{\tau}) = \sum_U \left(\frac{1}{\pi_i} - 1 \right) y_i^2 \quad (7)$$

An unbiased variance estimator is

$$\hat{V}(\hat{\tau}) = \sum_s (1 - \pi_i) \check{y}_i^2 \quad (8)$$

Ardilly & Tille [23] give the proof of formulas and some examples of poisson sampling and the other methods above.

An activity code estimated total is

$$\hat{\tau} = \sum_{i=1}^{26} \hat{\tau}_i \quad (9)$$

An activity code estimated mean is

$$\bar{y} = \frac{\hat{\tau}}{N} \quad (10)$$

An activity code estimated variance of \bar{y} is

$$\hat{V}(\bar{y}) = \frac{1}{N^2} \sum_{i=1}^{26} N_i^2 \left(\frac{N_i - n_i}{N_i} \right) \left(\frac{s_i^2}{n_i} \right) \quad (11)$$

where $N = \sum_{i=1}^{26} N_i$, $i = TR10, TR21, \dots, TRC3$ (26 NUTS2 regions).

If there exists a full enumeration strata n accepted as $N-1$ in formulas so as not to lose the variation, the calculation variance of an activity code comes from the full enumeration strata, and the effects of variances of full enumeration strata are visible.

3. Results

The estimated values of all variables are calculated for all NUTS2 regions and all activity codes. Since the number of activity codes (121) and the number of NUTS2 regions (26) is large, for Turkey only the total estimation is given. For estimations of Turkey

stratified sampling formulas are used with all activity codes used as strata.

The estimated values of the Turkey totals for all variables are given in Table 3. These values from the poisson sampling and the simple random sampling together are closer to population values except for a few points and small variances. All variables are estimated with approximately 0.1% or more smaller difference with real values.

Table 3. Estimated total of Turkey (for year 2013)

	$\hat{\tau}$ (poisson and simple random sampling used together)	$\hat{\tau}$ (systematic and simple random sampling used together)	$\hat{\tau}$ (only simple random sampling used)
D12110	745,455,721,619	743,641,098,497	740,822,016,427
D12120	426,406,194,396	425,913,245,604	428,212,238,425
D12150	107,758,066,110	107,465,821,686	107,235,044,723
D13110	663,007,049,935	660,424,519,776	662,195,321,458
D16110	3,434,500	3,442,350	3,437,053

Variances of $\hat{\tau}$ for all Turkey are calculated using with poisson and simple random sampling together, systematic and simple random sampling together, and only simple random sampling in strata are given in Table 4. For variables D12110, D12120, and D13110, the total estimations obtained from poisson and simple random sampling are the closest estimations and also have the minimum sample variances. For D12150, the simple sampling total estimation is the closest, but the poisson and simple random sampling variance are still minimum. For D16110, the systematic and random estimation is the closest, and simple random sampling variance is the minimum. It should also be considered that all calculations are made for estimating variable D12110.

Table 4. Variances of $\hat{\tau}$ for Turkey (for 2013)

	$\hat{V}(\hat{\tau})$ (poisson and simple random sampling used together)	$\hat{V}(\hat{\tau})$ (systematic and simple random sampling used together)	$\hat{V}(\hat{\tau})$ (only simple random sampling used)
D12110	6.52E+18	9.87E+18	1.37E+19
D12120	2.35E+18	4.61E+18	1.04E+19
D12150	2.11E+17	4.02E+17	4.85E+17
D13110	5.97E+18	9.33E+18	1.74E+19
D16110	1.79E+08	1.57E+08	1.50E+08

Turkey total estimations formulas are:

Turkey's estimated total,

$$\hat{\tau}_{turkey} = \sum_{i=1}^{121} \hat{\tau}_i \quad (12)$$

Turkey's estimated mean,

$$\bar{y}_{turkey} = \frac{\hat{\tau}_{turkey}}{N} \quad (13)$$

Turkey's estimated variance of \bar{y} ,

$$\hat{V}_{turkey}(\bar{y}_{turkey}) = \frac{1}{N^2} \sum_{i=1}^{121} N_i^2 \left(\frac{N_i - n_i}{N_i} \right) \left(\frac{s_i^2}{n_i} \right) \quad (14)$$

Turkey's estimated variance of $\hat{\tau}$,

$$\hat{V}_{turkey}(\hat{\tau}_{turkey}) = \hat{V}_{turkey}(N\bar{y}_{turkey}) \quad (15)$$

where $N = \sum_{i=1}^{121} N_i$, $i = 0520, 0811, \dots, 9522$ (121 sampled activity codes).

4. Discussion and Conclusion

In this study, estimations of the full enumeration part of the Annual Industry and Service Statistics in different sampling methods and their comparisons are presented. The sample size in this study is 62% of the full enumeration data. So, the size of survey decreases, obtained results are almost same with using poisson and simple random sampling together. Another important result of this study is that the size of the survey is reduced, but the level of estimates for the activity codes increased. This is important and desirable for regional decisions. Except for a few points, poisson sampling estimations have a smaller difference from population values and also smaller variances than other sampling methods. For all variables, there is 95% confidence that the intervals for μ contain the real value of μ for all different sampling methods.

References

- [1] Brick, J.M. 2011. The future of survey sampling. *Public Opinion Quarterly* 75(5), 872-888.
- [2] Giovannini, E. 2008. Understanding economic statistics: an OECD perspective (p. 195). Bologna: Societa Editrice il Mulino.
- [3] Scheaffer, R.L., Mendenhall, W., Ott, L. 1990. Elementary survey sampling, 5th ed. (p. 501). Belmont: Duxbury Press.
- [4] Chambers, R. 2003. An introduction to model-based survey sampling (p. 90). Southampton: EUROSTAT.
- [5] Yamane, T. 2010. Temel örnekleme yöntemleri. (A. Esin, C. Aydın, M. A. Bakır, E. Gürbüzsel, Trans.) (p. 509). İstanbul: Literatür Yayıncılık.
- [6] Mathew, O.O., Sola, A.F., Oladiran, B.H. Amos, A.A. 2013. Efficiency of Neyman allocation procedure over other allocation procedures in stratified random sampling. *American Journal of Theoretical and Applied Statistics*, 2(5), 122-127.
- [7] Barnabas, A.F. Sunday, A.O. 2014. Comparison of allocation procedures in a stratified random sampling of skewed populations under different distributions and sample sizes. *International Journal of Innovative Science, Engineering & Technology*, 1(8), 218-225.
- [8] Winkler, W.E. 2009. Sample allocation and stratification. Research report series of U.S. Census Bureau (Statistics #2009-8).
- [9] Cochran, W.G. 1977. Sampling Techniques (p. 596). New York: John Wiley & Sons.
- [10] Lohr, S. 2010. Sampling: design and analysis (2nd ed.) (p. 596). Boston: Brooks/Cole.
- [11] Hájek, J. 1958. Some contributions to the theory of probability sampling. *Bulletin of the institute of international statistics*. 36(3): 127-134.
- [12] Hájek, J. 1964. Asymptotic theory of rejective sampling with varying probabilities from a finite population. *Annals of Mathematical Statistics*, 35, 1491-1523.
- [13] Williams, M.S., Schreuder, H.T., Terrazas, G.H. 1998. Poisson sampling: the adjusted and unadjusted estimator revisited. Research Note of United States Department of Agriculture (RMRS-RN-4).
- [14] Aires, N. 1999. Algorithms to find exact inclusion probabilities for conditional poisson sampling and pareto pps sampling designs. *Methodology and Computing in Applied Probability*, 1 (4), 457-469.
- [15] Grafström, A. 2010. On unequal probability sampling designs, Umeå University, Doctoral dissertation, p. 31, Umeå.
- [16] Ghosh, D., Vogt, A. 1999. A modification of poisson sampling. *Proceedings of the American Statistical Association, Survey Research Methods Section*, 198-199.
- [17] Saavedra, P.J. 2000. Estimation strategies using variants of poisson sampling discussion. *Proceedings of the American Statistical Association International Conference of Establishment Surveys, Session 4*, 300-301.
- [18] Williams, M.S., Ebel, E.D., Wells, S.J. 2009. Poisson sampling: A sampling strategy for concurrently establishing freedom from disease and estimating population characteristics. *Preventive Veterinary Medicine*, 89, 34-42.
- [19] Brewer, K.R.W., Early, L.J., Hanif, M. 1984. Poisson, modified poisson and collocated sampling. *Journal of Statistical Planning and Inference*, 10, 15-30.
- [20] Lundquist, A. 2009. Contributions to the theory of unequal probability sampling. Umeå University, Doctoral dissertation, p.28, Umeå.

- [21] Särndal, C.E., Swenson, B. Wretman, J. 1992. Model assisted survey sampling. (p.695), New York: Springer-Verlag.
- [22] Ardilly, P. Tillé, Y. 2006. *Sampling methods: exercises and solutions* (p. 382), New York: Springer.
- [23] TÜİK. 2016. Annual Industry and Service Statistics (n.d.) In TURKSTAT Central Dissemination System-MEDAS. <https://biruni.tuik.gov.tr/medas> (Erişim Tarihi: 17.12.2016)