International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor (2018): 7.426

Sampling Error in Survey Research

Dr. Jiban Khadka

Abstract: ***In survey research, all deficiencies or weaknesses are caused by sampling or non-sampling issues that engender the discrepancy between sample value and population value. When the sample values deviate from the population value due to sampling, the deviation is termed as sampling error. It is statistically measured in term of standard error that is used to determine confidence interval to quantify the accuracy of the sample estimates. Unlikely the standard error from a sample distribution, estimated standard error can be calculated bases on the standard deviation derived from a random sample. An appropriate size of a sample can be a better estimator than a large size sample if it consists of similar characteristics of a population. Well defined target population, an exhaustive list of sample frame, effective sample design, reasonable sample size, acceptable level of confidence level are common measures to balance the sampling error.***

Keywords: Survey Error, Sampling Error, Margin of Error, Confidence Level, Sample Size

1. Introduction

In survey research, all deficiencies or weaknesses are caused by sampling or non-sampling issues that engender the discrepancy between sample value and population value. The discrepancy is termed as 'error' although it is explained by several terms. The research analysts or statisticians explain the terms 'mistakes', 'inaccuracies' or 'bias' that seem synonymous to error (Keeves, Johnson & Afrassa, 2000).But, the terms do not represent the actual meaning of the term error. Error is defined as the discrepancy between the observed value (sample statistic/value) derived from the sample survey and the expected value (population parameter/value) that may occur due to sampling or non-sampling errors.

Regarding the sources and types of errors, they vary by research types. As stated by Eurostat (2010) and others, random sampling variation, faulty frame of sampling units, unavailable/non-responsive/refusal/vested interested cases, item non-response case, lacking accurate response due to fear and prestige, ineffective measurement devices or techniques are the causal factors of survey errors. FCSM (2001) have listed the various errors such as sampling error, non-response error, coverage error, measurement error, and processing error that affect the data quality of the research findings. Keeves et al. (2000) state four types of errors in educational research. Out of them, three errors: intrinsic errors, instrumental errors and observational errors are described as common errors that occur in both census and sample survey whereas the fourth error: sampling error occurs only in sample survey research. In similar way, Groves (1989) states three types of errors: Measurement error, non-response error and sampling error. Most of the researches show that a researcher encounters the survey error by two types of errors: sampling error and non-sampling error even though they are described in various ways.

Despite the various types of errors presented in different ways, none are accepted as the best typology of survey error (Eurostat, 2010). From the study of errors, the errors can be categorized into two forms: sampling and non-sampling errors. To make common understanding about survey errors, the categorization of errors drawn from Verma (1981), Eurostat (2010) and the concepts of errors stated by Groves (1989) is presented in the following diagram:

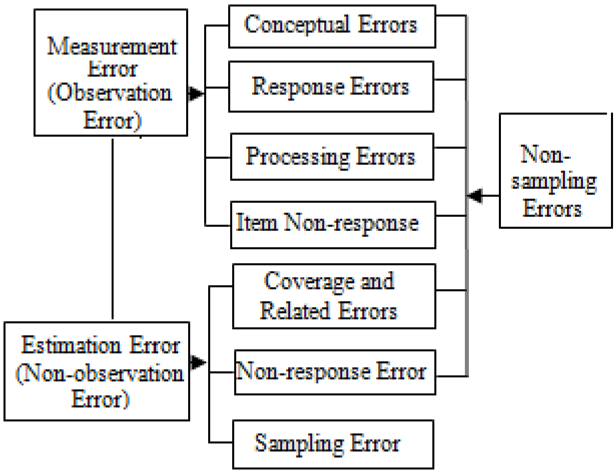

Figure 1: Typology of Errors

Even if all non-sampling errors are controlled, there might have been chances of differing the sample value from population value due to sampling error (Fritz, 2004). In this article, the author has delimit the discussion only in sampling error as it has crucial role to estimate the population parameter and can be measured statistically (FCSM, 2001. The author has incorporated many ideas from the research articles, theoretical or thematic literature to discuss about the different aspects of sampling error. Sample technique, sample size, confidence level, sources of sampling errors and alternatives to reduce the error, and many other aspects of sampling errors are focused in the paper. Two theories: probability and estimation theory are also discussed as they govern the different aspects of sampling error.

Although sampling error is important factor for determining precision and accuracy level of a survey research outputs, it is less descriptive in the research report. Regarding sampling error, only 45 percent of the publications specified the method they used for calculating the sampling errors (FCSM, 2001). Many other researchers pay little or no attention to this important aspect of survey research. A large number of articles on the survey errors has been published and a very limited of them has focused on the sampling error and its different aspects. The available literature is either insufficient or more complex mathematical patterns to incorporate the ideas for a novice researcher. In this article, the author has focused to illuminate about sampling errors and its related factors that the researchers, especially novice researchers, can be beneficial for improving the quality of their survey researches.

Sampling Error

Although sample survey was first used by English merchant John Graunt (1620 – 1674) to infer about the population, Neyman (1930s) is credited who established the superiority of random sampling (probability sampling) based on the concept of confidence intervals (Bethlehem, 2009). In this research approach, a small portion as a part of a population is used to draw findings based on the information gathered from the small portion of the population. The intension of a researcher is to generalize the findings to the whole population. Without considering upon the sampling error, the strength of the findings would be distrustful as it impacts on the estimation of population parameters. Because the sample survey was found as an error generating process, and was first put forth by Pierre Gy in the 1950s (Minnitt, Rice & Spangenberg, 2007). The error cause the distortion of survey results due to compromising representativeness of sample (Visser et. al, n. d.), and the sample value varies sample to sample that deviates from the population value. An example is presented to make the clear concept about sampling error as follows:

Let X_i be the units of population and x_i be the units selected for a sample. The value like mean (μ) obtained from the population containing X_i units is said to be a population parameter whereas the value like mean ($\bar{x}$) obtained from the sample having x_i units is said to be sample statistic. If the mean of a population is 100 (expected value) and the sample mean is 91 (observed value), then, in common sense, the error is 100-91 = 9 and this is lower as compared to the truth. If sample mean is 109, the error is also 9 but this time over estimation of 9 points. Conceptually, any difference between the parameter and the statistic is an error. As much number of samples are selected from a population as many different sample values may occur. When the sample values deviate from the population value due to sampling, the deviation is termed as sampling errors. The error can be measured statistically as follows:

Statistical Measurement of Sampling Error

According to FCSM (2001), the sampling error is probably the best-known source of survey error. The errors also occur as a result of calculating the estimators such as mean, median, proportion, standard error, variance, correlation etc. based on a sample data rather than the entire population data. The sampling error is measured in term of standard error based on simple random sampling (S. E.) (Huck, 2012) and the standard error of sample values from a sample distribution. It is guided by many theories like probability theory, central limit theorem, laws of large numbers or estimation theory.

The standard error of the statistic t is measured by: S. E. (t) = $\sqrt{Var(t)} = \sqrt{\frac{1}{k}\sum_{i=1}^{k}\left(t_i - \bar{t}\right)^2}$ in sample distribution where 't' is as a function of the sample observations, k as the possible samples of size 'n' , and $\bar{t}$ as the average mean of all the sample values. The distribution of *'t'*, according to central limit theorem, is a normal distribution called sample distribution that connects the estimator to the parameter of interest (Joseph & Reinhold, 2003).

In survey research, unlikely sample distribution of sample values, it is not possible to draw all possible samples of a fixed size from a population. In such situation, SE is derived from a single sample extracted from a population and the SE so estimated is named as estimated standard error (Huck, 2012). The estimated standard erro (*SE*) is calculated bases on the standard deviation of a sample. For a single sample from an infinite population, as stated by Gupta (2007) and Visser et. al (n.d.), the equation of *SE* is:

$$\sqrt{\frac{\text{Sample Variance}}{\text{Sample Size}}} = \sqrt{\frac{s^2}{n}} \text{ ....(i)}$$

But, in the case of finite population, the equation for the *SE* based on a sample without replacementis:

$$\sqrt{\frac{\text{Sample Variance}}{\text{Sample Size}}\left[\frac{\text{(Population Size - Sample Size}}{\text{Population Size}}\right]} = \sqrt{\frac{s^2}{n}\left[\frac{N-n)}{N}\right]} \text{ ....(ii)}$$

Both equations are based on simple random sampling and the software program like SPSS also produces the SE based on simple random sampling (Joseph & Reinhold, 2003).Further, they state that the standard error is used to determine confidence interval to estimate the population parameter. The standard error is always estimated based on random sampling technique and the technique possesses some important properties of probability theory.

Probability Theory

Theoretically, sampling error is measured by standard error and the equation of standard error $\sqrt{\frac{1}{k}\sum_{i=1}^{k}\left(t_i - \bar{t}\right)^2}$ is based on simple random sampling. Lavrakas (2013) states that simple or other type of random sampling reduces the sample variances. If there is sampling bias, the error can not be avoided by the repeated selection of samples form a population (Lavrakas, 2013).In simple random sampling technique, all units of population of size 'N' have equal chance of being selected in the sample of size 'n' and each unit has the equal probability (=n/N) , as classical definition of probability, being selected in the sample. In repeated sampling process with replacement, the samples of size n can be selected from the same population of size n in $^{N}c_n$ in ways by the method of combination. In repeated number of selection of random samples with the same size n from the same population N, the sample value like mean, median, proportion, standard deviation and other statistics occur in a certain pattern or in normal form. This pattern is governed by two theories "laws of large numbers and central limit theorem that are as fundamental theories of probability. The distribution of these statics like mean or others formed is called sample distribution in bell-shaped curve and then the standard deviation of sample means about the average sample mean is the standard error of this distribution. The mean of this distribution is equal to the population mean. It is possible only by efficient samples that are similar to each other and have low variance and low sampling errors (Lavrakas, 2013).

Laws of Large Numbers. What result would come if a certain size of sample is drawn at a large number? The behaviour of sample statistic can be described by probability ideas. According to the mathematical law of large number, the sample value (say $\bar{x}$) approaches the population value (say μ), symbolically, $\bar{x} \rightarrow \mu$ as the number of samples increases infinitely(Sedlmeier & Gigerenzer, 1997). The distribution of the sample statistics forms a normal distribution as the sample size increases. But, the laws of mathematical laws of large number does not concern with the population distribution. But, according to the empirical laws of large numbers, as sample size increases the sample distribution approaches the population distribution that concerns with frequency distribution unlikely the sampling distribution as governed by the mathematical laws of large numbers (Sedlmeier & Gigerenzer, 1997).

In social science research, mathematics laws of large numbers is not applicable as it does not concern with the size of samples rather it concerns with the number of samples and the property of central limit theorem is applied in this law. Contrary, the empirical laws concerns that as the sample size increases, sample value better estimates the population parameter. In survey research, the accuracy is estimated from frequency distribution rather than sample distribution (Sedlmeier & Gigerenzer, 1997).

Central Limit Theorem. Gupta (2007) states that if x_1, x_2,x_3,...,x_n is a random sample of size n from any population with size N, mean μ and variance σ^2, than the sample mean of the n samples $(\bar{x})$ is normally distributed with mean μ and variance σ^2/n provided n is sufficiently large. In exhaustive case of sample means, the sample mean approaches population mean that is unbiased estimate of population mean (Gupta, 2007).In other words, a number of samples with same size through same method can be drawn from the same population. If a sample statistic is calculated from these samples, the statistics differ from sample to sample and the average of the statistic is supposed to be equal with population parameter value (Gupta, 2007). The standard deviation of sample means from the sample distribution standard error.

In practice, a researcher can not compute the SE based on repeated sampling from the same population. Other alternative approaches are employed like Taylor series linearization and jack knife replication of samples from the parent sample (FCSM, 2001) that can be used as alternatives. In central limit theorem, n is interpreted in two ways, first number of samples and second size of samples. When an experiments is conducted on a number times with samples of size n, the sample values like mean form a normal distribution and the average sample mean equals to population mean. Confidence intervals can be determined from each sample value but it is not sure that population value lies in each of the intervals. In contrast, if a sample of size n increases from a population, this experiment may not satisfies the property of central limit theorem. But, the distribution of sample units is identical to the population distribution that is main concern of a researcher in a survey. Whatever be the size of sample, the estimation from the sample is necessary to be best estimator to the population parameter. According to estimation theory, the estimators from a sample can be determined that is discussed as follows:

Estimation theory

In the entire discussion of sampling error, the central concern of a survey researcher is how to find such sample value that describes the characteristics of population value. There always remains sampling error until the sample size is less than population size. The researcher's intension is to calculate the better estimator to describe the population characteristics.

When a researcher is unable to study entire population units by any reasons, they need to study a part of population (sample) to infer about the population. Estimation theory deals with an estimated point or interval derived from the sample. The real value is expected to be equal to the estimated point or to be laid in the interval. Sampling error has important role in both types: point or interval estimation (SCNCC, 2003). As the point estimation doesn't take level of confidence, it simply computes the sample value that is assumed to be equal to the unknown population value (Huc, 2012). In this paper, only interval estimation is discussed as it is logical and useful in social survey research.

Interval Estimation: Standard errors are usually used to quantify the accuracy of the estimates. According to sample distribution theory, it is indicated that about 68 percent of the estimates lie within one standard error or standard deviation of the mean ($\mu \pm \sigma$), 95 percent lie within two standard deviations ($\mu \pm 2\sigma$) and all estimates are supposed to lie within three standard deviations ($\mu \pm 3\sigma$) in normal or symmetrical distribution. In normal distribution, the observations or values are equally divided from the central value (mean, median or mode) and the curve of such distribution is bell-shaped in this distribution. As stated before, it is not possible to take all possible samples or other alternative methods to find standard error. Then, the researcher needs to derive an interval from a sample to estimated the population parameter. The interval derived from a sample of a size at a certain level of confidence is known as confidence interval (Huck, 2012). The confidence interval is estimated by the equation:

$$\bar{X} - z\frac{s}{\sqrt{n}} \le \mu \le \bar{X} + z\frac{s}{\sqrt{n}}. \quad \text{......................................(iii)}$$

The equation (iii) is derived from three concepts: standard errors, margin of error and confidence level. In the equation, the three concepts are $\bar{x}$ as sample mean, $\frac{s}{\sqrt{n}}$ as standard error and z score at specified level of confidence where n is the size of sample.

To determine confidence interval, sample size has determinate roles that increases or decreases the size of margin of error or confidence interval. Margin of errors, simply, explains the sampling error quantitatively and it is defined by $2\sigma/\sqrt{n}$ where σ is the standard deviation of the observations distributed normally. If a sample size is of 30 or more for a population with the regardless of distribution, the distribution of sample means approaches approximately normal. In common words, Bartlett, Kotrlick & Higgins (2001) state, "the risk the researcher is willing to accept in the study, commonly called the margin of error." With the help of margin of error (*me*), confidence interval is determined and calculated at different levels of confidence such as 90%, 95%, 99% etc. where "actual value is assumed to be laid in the interval" (Gilliland & Melfi, 2010). Bartlett et al. (2001) prefer to take 3% in continuous and 5% in categorical data as the acceptable level of margin of error in survey research. An example is presented in the following diagram that shows how sample size affects the errors at a certain confidence level.

According to Israel (2013), as the accurate value is not possible to get from a sample, a range, some times called sampling error, is necessary to be derived from the sample value in which the population parameter is assumed to be laid. The range is sometimes called precision level. And confidence level or level of risk is that is encompassed by central limit theorem. According to this theorem, when a sample of a definite size is repeatedly drawn from a population, the average value of the sample values is equal to the population value. The claim is made based on confidence level. In 95% confidence level, 95 sample values out of 100 sample values fall within the the range of precision level(Israel, 2013). In normal distribution, 95% of the sample values are expected to lie within two times the standard deviations from the population value.

Let, population size be N, different sizes of samples n_1, n_2 and n_3 are selected from the same population where $n_1 < n_2 < n_3$. The results of respective margin of errors at a certain level of confidence interval can be calculated by Cochran's formula in the order: $me_1 > me_2 > me_3$ (Bartlett et al., 2001). The following figure 1 conceptually represents the above information:

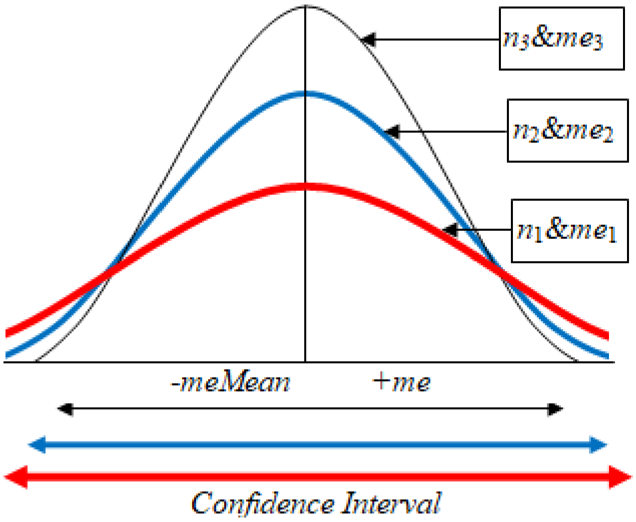

Figure 1: Presentation of Margin of Errors at Different Sizes of Samples

From the observation of diagram, in large size of sample such as n_3, sample statistic is closer to central value because the length of confidence interval is shorter than in the small size of sample. The diagram shows that the sampling error measured in terms of margin of error decreases as the sample size increases and justifies as having the inverse relation between the sampling error and sample size. Hence, a researcher should consider the statistical role of margin of error at the stage of making decision while determining sample size. Size of population, sampling technique and sample size affect the margin of error. As stead by Fritz (2003), the size of population has little impact on the margin of error but the random sampling is essential to make sample representative to ensure equal probability for being selected in a sample.

Sample Size and Sampling Error

As stated by Lee et al. (2009), it is always not possible to increase the size of sample as the researchers wish. Further, say that no researchers can claim that a large size sample is always truly representative of an entire population. So, the researchers' need is to take a reasonable size of sample that keeps balance between the inputs: effort, time and cost, and also to produce a better estimation for the population parameter. An example is discussed to observe effects of sample size on sampling error as follows:

For example, units of a Population = $\{2,4,6,8,10,12,14,16,18,20\}$, N = 10,mean (μ) =11.00 SE = 1.91, SD= 6.06

Table 1: Means, Standard Errors, Standard Deviations based on different Sample Size

Sample	Units	Size(n)	Mean($\overline{X}$)	Min	Max	SE	SD(σ)
I	6, 8, 14, 16, 10	5	10.8	6.00	16.00	1.85	4.15
II	6, 8, 14, 16, 10, 18	6	12.0	6.00	18.00	1.93	4.73
III	6, 8, 14, 16, 10, 18, 12	7	12.0	6.00	18.00	1.63	4.32
IV	6, 8, 14, 16, 10, 18, 12, 20	8	13.0	6.00	20.00	1.73	4.90
V	6, 8, 14, 16, 10, 18, 12, 4	8	11.0	4.00	18.00	1.73	4.90

In sample I of size 5, the mean of the sample (=10.8) is differed by .2 (=11-10.8) from population mean (=11) with *SE* =1.85 and SD = 4.15. In sample II, when the size of sample is increased by one unit 18 (n =6) that falls outside the range of sample I, mean (= 12) is differed by -1 (=11-12), and *SE* (=1.93) and SD (=4.73) increased in comparison to sample I. In sample III, size of sample is increased by one unit 12 (n=6) which falls within the range of sample II, the mean remains the same but both *SE* (=1.63) and SD (=4.32) are decreased. In sample IV, one unit 20 is increased that falls outside the range of sample III have followed the trend of sample II. But, size of sample in sample V is unchanged but a unit 20 is replaced by 4 that reduces the range of sample IV and the mean equals to population mean keeping the *SE* and SD the same. From this observation, it seems that *SE*, in all cases, does not decrease as the sample size increases. The trend of change in *SE* and SD are in the same direction in all sample sizes. When sample size increases and range decreases that yields the good estimator like sample V in comparison to sample IV. If the additional units are outliers that increases both *SE* and SD and also increases the difference between estimator and actual value. Appropriate small size can be the better estimator than a large size sample if the sample size sample has similar characteristics of a population like sample I in comparison to II, III and IV. In general trend of *SE*, as sample size increases, the sampling error decreases but cannot be estimated *SE* only looking at the size of sample. Because it also depends upon the sample mean that is affected by the value of sample units. Sample IV and V also justifies that the sample error occur by chance of selection of units of population in the sample. When unit 20 is replaced by 4, then the sample mean equals to population mean where *SE* and SD remains unchanged. From the analysis of sample sizes have inverse effect on the margin of error.

Margin of error explains the population parameter from the laws of probability and laws of chances (Fritz, 2003). The confidence interval for the population mean from sample mean is: sample mean $\pm$ Z value $\times$ SE or $\left(\overline{X} \pm z\frac{s}{\sqrt{n}}\right)$. In the example, in the interval, the population mean is expected to be laid. The required interval at 95% confidence level is (10.8-1.96X1.85 $\leq \mu \leq$ 10.8-1.96X1.85) = (7.714 $\leq \mu \leq$ 14.426) for sample I. The margin of error is least in sample III that justifies size of sample and chance of units of selection determine the sampling error. So, the sample III provides the best estimation for the population mean although each interval in above five samples contain the population mean. In all sample surveys, the population values may not lie in the intervals (Huck, 2012).

Sources of Sampling Error

Why does the error, especially sampling error, occur in survey research? It is not appropriate to answer standing at an extreme position as there is no single answer to the question. According to Batenburg (2011), four challenges of sample survey are relevant to the stated question. They are: to get enough and the right respondents, to construct a valid and reliable questionnaire, to realize sufficient response, and to get the most out of the data. The common understanding about the non-observational error (coverage error, lacking randomness in sampling, non-response case) that occurs due to the lack of representativeness of sample, and the other factors of respondents such as non-response: unwillingness to respond, unavailable respondents and unable to respond or non-response error by unit and item affect the estimation of sampling errors (FCSM, 2001). Keeves, Johnson, & Afrassa (2000) state that sampling errors are caused by two reasons. First, it is caused by lacking enough representativeness of the population called sample bias and secondly, it is caused by the variability between cases and the number of cases. Gupta (2007) states primarily five reasons of sampling error: faulty selection of sample, substitution, faulty demarcation of sampling units, error due to bias in the estimation method and variability of the population. These all statements regarding sampling errors indicate that procedure of sampling, representativeness of population, exclusion and inclusion criteria of sampling units mainly cause the sampling error.

The ten different sources of sampling errors were split by Gy (1982) into two components: random error and bias. Walther & Moore (2005) state that lacking representativeness of the target population in sampling causes the bias whereas the presence of random error causes the degree of precision that depends on the observed data. The sampling bias occurs due to sample selection that consequently distorts the measurement process and lessens the degree of generalization (Sica, 2006). In contrast, Sica (2006) doesn't agree with the argument 'the bias can be corrected', and says that it is impossible to eliminate completely. Also, Cohen et al. state that sampling error may occur due to the chance of selection of population units besides the result of mistakes made in sampling process. Another source of sampling error is also attributed to the population heterogeneity creates random differences between sample and population characteristics. Lack of well-defined sample frame causes the sampling error because all units may not have equal chance to be selected in the sample. For example, in sample survey of private schools conducted by the author, the author first prepared a sample frame and selected sample of secondary schools. In field visit, unfortunately, some schools were closed, transferred or out of defined population. In the list of schools published by the district office, lower secondary schools were listed in the list of secondary schools. The substituted schools in replace of the schools may bring, to some extent, differences to expected sample values.

In summary, lacking detail information about population, well-defined sample frame, random error and bias, sampling technique and reasonable size of sample mainly cause the sampling error.

Alternatives to Reducing Sampling Error

Several factors cause sampling error, and the factors need to treated at right stage of research process. One of the alternative measures is the sampling technique that determines the representativeness of a sample. As random sampling technique provides the chances to all unit of population to be selected in a sample, the technique reduces the sampling error. Visser et. al (n. d.) state that sampling error is comparatively higher in cluster sampling than the simple random sampling and stratified random sampling, but it is lower in stratified random sampling than simple random sampling . The stratified sampling technique avoids bias, and is possibly a best way to make sample representative even for the highly heterogeneous units of population. If a population size is small and located in a narrow territory, simple random can be effective. If the population is more heterogeneous, stratified random sampling can be effective. In stratified sampling technique, strata should be non-overlapping as far as possible. In very big size population, scatter units and need more time, labour and cost, cluster sampling can be appropriate. In cluster sampling technique, the clusters are necessarily to be homogeneous.

In survey research, it is necessary to specify the actual units of analysis from whom the information collected can address research issues. And then the issue of sample frame occurs that affects the selection of sample units. Many statisticians state that the difference may not occur only by sampling method rather they occur due to the mismatch between a target population and the sample frame. Well defined sampling frame is a way to reduce sampling error otherwise defective sampling frame creates over or under coverage error that cause the sample bias. Under coverage and over coverage both are not appropriate for the accuracy and precision level in the sample value. It ultimately brings variation between sample and population values. Careful consideration upon the detail information of population, time, cost or labour for data collection, updated list of population units can help to prepare the an effective and efficient sample design.

Sample size is a core factor to determinate sampling error. Mathematically, sampling error can be reduced by increasing the size of sample or confidence level. When the size of a sample increases, the ratio of sample size to population size also increases that lessens the sampling error. Coe (1996) states three factors: accuracy, precision and cost that should be on balance while determining a sample size. According to him, accuracy concerns with sampling error, non-sampling error and bias, and the precision with the standard error. Eng (2003) states the five study design parameters that determine an appropriate sample size. They are: minimum expected difference (also known as the effect size), estimated measurement variability, desired statistical power, significance criterion, and a one- or two-tailed statistical analysis. The nature of data in survey research is another pertinent factor to determine sample size in which continuous data require comparatively less number of sample units in comparison to categorical data (Bartlett et al., 2001). In survey research, if a researcher has no easy access to all population units due to various problems like transportation, communication, nature of respondents, social and cultural factors or any other factors related to inputs, an appropriate size of a sample can be better alternative rather than a large size of sample. Many approaches, choices or techniques are available to determine a sample size such as historical research, percentage samples, mean samples or formulae such as Yamane (1967), Krejcie & Morgan (1970), Cochran (1977) formulae etc.

Another alternative is to reduce the sampling error by reducing estimated *SE*. In this method, sampling error can be reduced by increasing confidence level such as from 90% to 95% or to 99%. Both sample size and confidence level determines sampling error. In mathematical model , as the sample size of sample increases to size of population, sample error approaches to zero. Symbolically, as $n \rightarrow N$, $\bar{x} \rightarrow \mu$ (in the case of average value). The size of sample inversely affects the *SE* by $n^{1/2}$ but change in population size moderately affects the *SE* (Wingersky & Lord, 1984). Further, say that the number of common items have little effects on *SE* and then even a few number of sample items may be enough to estimate population parameter. In the table 1, sample II with size 6 is better estimator than sample V of size 8 in regard with the population value. If first selected respondents are interested to participate in the study that is appropriate in controlling sampling error. In the case of non-responsive cases, there may be chance of occurring outliers and they may not match to the other respondents. The best fit possible cases can be substituted in the unit non-responsive situation.

2. Conclusion

In sample survey research, the intent of the study is to generalize the findings to the whole population. The sampling error indicates the research quality as it is derived from a sample. Is a target population well defined? How is the sample frame prepared? What are effects of sample design? Is the sample size appropriate? What is the level of confidence? All these questions concern with sampling error. The study of sampling error can be the better understanding about population parameters. However, non-sample errors have also determinant role to estimate the population parameters. In all situation, repeated sampling process is not possible and then, some aspects like estimated standard error, use of empirical law of large numbers, random sampling technique, sample size, acceptable margin of error are most common considerable factors. Because they have significant role to estimate the population value. Well defined target population, an exhaustive list of sample frame, effective sample design with the use of random sampling technique, reasonable sample size, acceptable level of confidence level are major alternatives to control the sampling error. Careful attention on the components of sampling error can make a survey research trustworthy.

References

[1] Bartlett, J. E., Kotrlick, J. W. & Higgins, C. C. (2001). Organizational research: Determining appropriate sample size in survey research. *Information Technology, Learning, and Performance Journal*, 19 (1).

[2] Batenburg, R. (2011). *Survey research: challenges and practices*. Wo-college survey underzoek.

[3] Best, J. W. & Kahn, J. V. (2008). *Research in education* (10th ed.). New Delhi: Prentice Hall.

[4] Bethlehem,J., Cobben, F. & Schouten, B. (2009). Indicators for the representativeness of survey response. *Statistics canada's international symposium series*. Statistics Canada. Retrieved from *www.statcan.gc.ca/pub/11-522-x/2008000/article/10976-eng.pdf* on 20 November, 2013.

[5] Cochran, W. G. (1977). *Sampling techniques* (3rd ed.). New York: John Wiley and Sons. ISBN: 0-471-16240-X.

[6] Coe, R. (1996). Sampling size determination in farmer surveys. *Transforming Lives and Landscapes*. ICRAF Research Support Unit Technical Note No 4. ICRAF World Agroforestry Centre Nairobi, Kenya. 11 pp.

[7] Cohen, L., Manion, L. & Morrison, K. (2007). *Research methods in education* (6thed.). USA: Rutledge Publication.
[8] Creswell, J. W. (2011). *Research design* (3rded). India: SAGE Publication Pvt. Ltd.
[9] Eng. J. (2003). Sample size estimation: How many individuals should be studied? *Statistical concepts series.* Radiology 2003; 227:309–313, Index terms:Published online 10.1148/radiol.2272012051.
[10] Ercikan, K. & Roth, W. M. (2006). What good is polarizing research into qualitative and quantitative? *Educational researcher*, 35 (5).
[11] Eurostat (Statistical Office of the European Union) (2010). *An assessment of survey errors in EU-SILC.* Luxembourg: Author. ISSN 1977-037.
[12] Federal Committee on Statistical Methodology (2001). *Measuring and reporting sources of error in surveys*. USA: Author.
[13] Gilliland, D. & Melfi, V. (2010). A note on confidence interval estimation and margin of error.*Journal of Statistics Education*, 18 (1).
[14] Groves, R. M. & Lyberg, L. (2010). Total survey error: past, present, and future. *Public Opinion Quarterly*, 74 (5), 2010, pp. 849–879. Oxford University Press.
[15] Groves, R. M. (1989)). *Survey errors and survey costs*. New Jersey: John Wiley & Sons.
[16] Gupta, S. C. (2007). *Fundamentals of statistics* (6th edition). Mumbai: Himalaya Publishing House.
[17] Huck, S. W. (2012). *Reading statistics and research* (6th ed.). USA: Pearson Education, Inc (ebook).
[18] Israel , G. D. (2013). *Determining sample size*. Agricultural Education and Communication Department, UF/IFAS Extension, University of Florida.
[19] Joseph, L. & Reinhold, C. (2003). Introduction to probability theory and sampling distributions. *Fundamentals of clinical researchfor radiologists*. American Roentgen Ray Society.
[20] Keeves, J. P., Johnson, T. G. & Afrassa, T. M. (2000). Errors: What are they and how significant are they? *International Education Journal*, 1 (3).
[21] Lavrakas, P. J. (Ed.) (2013). Sampling Error (Chapter). *Encyclopedia of survey research methods*. Sage Publications, Inc. DOI: http://dx.doi.org/10.4135/9781412963947.n510
[22] Lee, G., Bryan, J. B. & Johnson, T. P. (2009). Survey method in public administration research: *A content analysis of journal publications paper presented at the state policy research symposium*. College of Urban Planning and Public Affairs, University of Illinois, Chicago, March 17-18, 2009.
[23] Minnitt, R. C., Rice, P.M. & Spangenberg, C. (2007). Understanding the components of the fundamental sampling error: a key to good sampling practice. *The Journal of the Southern African Institute of Mining and Metallurgy*, 107, 505-511.
[24] Office of Quality Improvement (2010). Survey fundamentals: A guide to designing and implementing surveys. Madison: University of Wisconsin.
[25] Scheuren, F. (editor) (2004). What is a survey (2^{nd} ed.). American Statistical Association. Retrieved from *www.amstat.org/sections/srms/pamphlet.pdf* on 3 January, 2015.
[26] Schouten, B., Cobben, F. & Bethlehem, J. (2009). Indicators for the representativeness of survey response. *Survey methodology*, 35 (1), 101-113. Statistics Canada, Catalogue No. 12-001-X
[27] Sedlmeier, P. & Gigerenzer, G. (1997). Intuitions about sample size: The empirical law of large numbers. *Journal of behavioral decision making*, Vol. 10, 33±51 (1997)
[28] Sica, G. T. (2006). Bias in research studies. *Radiology*, 238 (3), 780-78.
[29] Statistics Canada's National Contact Centre (SCNCC) (2003). *Survey methodsand practices*. Canada: Minister of Industry.
[30] Vissbr, P. S., Krosnick, J. A., & Lavraws, P. J. (n.d.). Survey Research. Chapter nine. Retrieved on 1 January, 2015 from web.stanford.edu/.../faculty/krosnick/Survey_Research.pdf
[31] Walther, B. A., Moore, J. L. (2005). The concepts of bias, precision and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. *Ecography*, 28:6.
[32] Wingersky,S.& Lord, F. M.(1984).In investigation of methods for reducing sampling error in certain IRT procedure. *Apply psychological measurement*(3), pp. 347-364.