

Sample Size and Its Role in Central Limit Theorem (CLT)

Mohammad Rafiqul Islam

BRAC University, 66 Mohakhali, Dhaka, Bangladesh

Email: mrafiq@bracu.ac.bd

Abstract--It is very important to determine the proper or accurate sample size in any field of research. Sometimes researchers cannot take the decision that how many numbers of individuals or objects will they select for their study purpose. Also, a set of survey data is used to verify that central limit theorem (CLT) for different sample sizes. From the data of 1348 students, we got the average weight for our population of BRAC University students is 62.62 kg with standard deviation 11.79 kg. We observed that our sample means became better estimators of the true population mean. In addition, the shape of the distribution became more Normal as the sample size increased. So it is concluded that our simulation results were consistent with the central limit theorem.

Keywords---confidence interval (CI), inference, population and sample, precision, sample size.

Introduction

Statistical Inference

If we want to know the average education level in Bangladesh, one way we could do that is gone, and find out the education level of every single person in Bangladesh. And then you could calculate a mean, and then we have the average educational level. But, this is extremely impractical. So instead, we just ask some people, what their education level is, and then we find out the average education level. We shall also try to figure out whether or not that value that we got from asking those people is actually accurate in terms of representing what is in the entire population. When we make calculation from the sample, we cannot be 100% sure that we are 100% right or 100% wrong, but we are going to figure out what the chances are, that we are right or wrong. To get the chances of doing error, we truly need a random sample of the population to represent our sample. The closer we get to that, the better off that we will be. But even if we get them perfectly, randomly selected there is always a chance that, by random chance, we are going to end up with some mistakes. And we are trying to minimize the chances that we are doing wrong. That is what we are all about in inferential statistics is figuring out what the chances are that we are going to be wrong. According to Cox (1958), an inference can be considered as answering the question: ‘What do these data entitle us to say about a particular aspect of the populations that interest us’. Cox (1958), also states, “Two things mark out statistical inferences. First, the information on which they are based is statistical, i.e. consists of observations subject to random fluctuations. Secondly, we explicitly recognize that our conclusion is uncertain, and attempt to measure, as objectively as possible, the uncertainty involved.”

A statistical inference carries us from observations to conclusions about the populations sampled (Cox, 1958). The main challenge is to select a representative sample. Marshall (1996) stated that the size of the sample is determined by the optimum number necessary to enable valid inferences to be made about the population. The larger the sample size, the smaller the chance of a random sampling error, but since the sampling error is inversely proportional to the square root of the sample size, there is usually little to be gained from studying very large samples (Marshall, 1996). Within a quantitative survey design, determining sample size and dealing with nonresponse bias is essential (Barlett *et al.*, 2001). Holton & Burnett (1997), stated: “one of the real advantages of quantitative methods is their ability to use smaller groups of people to make inferences about larger groups that would be prohibitively expensive to study”. Then the question is, how large the sample will be taken from the population to infer about the

research finding? any researchers could benefit from a real-life primer on the tools needed to properly conduct research, including, but not limited to, sample size selection (Barlett *et al.*, 2001).

This paper will be very helpful especially for the beginners to decide their sample size initially. First of all the factors those influence the size of the sample are discussed briefly. Secondly, various formulas for determining sample are explained with examples depending on the field of study. In this paper, mainly the simplified form of Cochran's (1977), sample size formula for both continuous and categorical data will be taken into considerations. Besides, Krejcie & Morgan's (1970), formula for determining sample size for categorical data will be presented as it provides identical sample sizes in all cases. Also, there are many other formulas for determining sample size but these two are the most widely used in the field of research (Barlett *et al.*, 2001).

In addition, a set of survey data is used to verify that central limit theorem (CLT) for different sample sizes. We examined how sample data can be used to discover the truth about a population. Our population data consists of ages, weights, and heights of 1358 undergraduate students of BRAC University at Dhaka. The data are collected by the students of statistics (summer 2015) of BRAC University students for their coursework assignment. We run few simulations on this data by R programming to see if we can replicate the samples what the Central Limit Theorem tells us about sampling

Research Methods

Factors influencing sample size

One of the most burning questions in any type of quantitative survey is to determine the sample size which is a very important part of research. Spaeth (1992), describes, "Unless there is one variable that you are interested in beyond all others, this is one of the hardest questions that you can ask a survey researcher". Spaeth (1992), also states that one answer is: "How much time or money do you have?" Another is: "It depends." It depends on how accurate you want your estimate to be. That is the easy part. It depends on what kinds of comparisons you want to make. Peers (1996), mentioned that sample size is one of the four inter-related features of a study design that can influence the detection of significant differences, relationships or interactions. Generally, these survey designs try to minimize both alpha (α) error (finding a difference that does not actually exist in the population) and beta (β) error (failing to find a difference that actually exists in the population (Peers, 1996). There are many factors which influence the sample size including the purpose of the study, population size, the risk of selecting a "bad" sample, and the allowable sampling error (Israel, 1992). Miaoulis & Michener (1976), described that in addition to the purpose of the study and population size, three criteria usually need to be specified to determine the appropriate sample size: the level of precision, the level of confidence or risk, and the degree of variability in the attributes being measured. Each of these criteria is elaborated shortly below. It has to be mentioned that sample size also depends on the type of data to be collected, size of the population, time and budget except the three factors mentioned earlier.

Precision or accuracy (Margin of error)

As sample deals with a part of a population, one must accept a risk of being wrong when inferring something about a population based on the basis of sample information. This is why before taking a sample; one should identify the amount of risk to be allowed (or willing to take). This amount of risk directly relates to the size of the sample.

The risk is specified by two interrelated factors: the precision (reliability) range desired and the confidence level. According to Johnson (1959), sampling error or the precision of a sample, result is meant how closely we can reproduce from a sample the results which could be obtained if a complete count of the population were made under the same conditions. Johnson (1959), also states that the difference between the sample result and the true value (population parameter) is called the accuracy of the sample survey which is also known as precision that is most frequently measured. Thus precision is the maximum allowable error expressed in percentage when the sample is taken. This implies the maximum allowable difference between the sample estimate (which is supposed to calculate from the sample) and the true population value. In other words, this is the maximum allowable sampling error. If the difference is reduced, the level of desired precision or accuracy will be higher. Or if the difference widens, the level of desired precision or accuracy will be lower. Thus, 1% precision is greater than the 5% precision as in 1% precision the difference (error) is lesser than the 4% precision. This is why for getting a high degree of precision (accuracy), it is required larger sample size than the relatively low degree of precision (accuracy).

Precision is denoted by 'E' which means sample estimate will be within $\pm E\%$ of the population parameter. Precision level of 5% means that the actual value of the population (parameter) lies within an interval (+0.05 or -0.05) around the sample estimate. Thus, if a researcher finds that 70% of students of a university in the sample have adopted the practice of using English language as a method of communication in the campus with a precision rate of $\pm 5\%$, then the researcher can conclude that between 65% and 75% of students in the population (that is all the students) have adopted the practice. Precision is also known as margin of error.

The general rule relative to acceptable margins of error in educational and social research is as follows: For categorical data, 5% margin of error is acceptable, and, for continuous data, 3% margin of error is acceptable (Krejcie and Morgan, 1970). Most survey organizations use 3%, 5% or 10% precision level as the minimum. If there are too many variables in a research study, the researcher must make decisions as to which variables will be incorporated into formula calculations. Cochran (1977) addressed this issue by stating that "One method of determining sample size is to specify margins of error for the items that are regarded as most vital to the survey. An estimation of the sample size needed is first made separately for each of these important items". Researchers may increase these values when a higher margin of error is acceptable or may decrease these values when a higher degree of precision is needed (Barlett *et al.*, 2001).

More commonly, there is a sufficient variation among the sample size n 's so that we are reluctant to choose the largest, either from budgetary considerations or because this will give an over-all standard of precision substantially higher than originally contemplated. In this event, the desired standard of precision may be relaxed for certain of the items, in order to permit the use of a smaller value of the sample size n (Cochran, 1977). "The precision of the results procured from the sample survey is contingent not only on the size of the sample but also on other aspects of the sample design, such as the way the sample is chosen and the process of calculating the estimates from the survey results" described by Johnson (1959).

Level of Confidence Interval (CI)

The confidence level means, how much we are confident that the sample estimate is as accurate as we desired. The confidence or risk level is based on ideas encompassed under the Central Limit Theorem (Israel, 1992). Central Limit Theorem states that when a population is repeatedly sampled, the average value of the attribute obtained by those samples is equal to the true population value and distributed normally about the true value (Israel, 1992). To minimize the risk one should have high confidence. This risk is reduced for 99% confidence levels and increased for 90% (or lower) confidence levels. In most cases, the 95% confidence level is specified. The 99% confidence level (an alpha level of .01) may be used in those cases where decisions based on the research are critical and errors may cause substantial financial or personal harm, e.g., major programmatic changes (Barlett *et al.*, 2001). If a 95% confidence level is chosen, 95 out of 100 samples will have the true population value within the range of precision specified earlier (Israel, 1992). Thus, the desired 3% precision with 95% confidence interval means that we are 95% confident to get 3% precision that is we are 95% confident to make maximum 3% error when we will take the sample size. For the most common confidence levels 90%, 95% and 99% of the Z Scores are 1.645, 1.96 and 2.57 respectively.

Degree of Variability

The distribution of attributes or characteristics in the population is known as the degree of variability in the attributes being measured (Israel, 1992). The desired attributes or a character in the population also is an important factor for determining the sample size. For the more variability in the population, which is known as heterogeneous population, the larger sample size is required to obtain a given level of precision. On the other hand, the smaller sample size is needed when there is less variability (more homogeneous) in a population. Usually, the variability is measured by the variance. The estimation of the variance of the primary variables of interest is very important for calculating the sample size as the researcher does not have direct control over variance and must incorporate variance estimates into research design (Barlett *et al.*, 2001). There are four ways Cochran (1977), of estimating population variances for sample size determinations: (1) take the sample in two steps, and use the results of the first step to determine how many additional responses are needed to attain an appropriate sample size based on the variance observed in the first step data; (2) use pilot study results; (3) use data from previous studies of the same or a similar population; or (4) estimate or guess the structure of the population assisted by some logical mathematical results. The first three ways are logical and produce valid estimates of variance. However, in many educational and social research studies, it is not feasible to use any of the first three ways and the researcher must estimate variance using the fourth method.

When estimating the variance of a categorical (proportional) variable such as gender, [Krejcie & Morgan \(1970\)](#), recommended that researchers should use 0.50 as an estimate of the population proportion. This proportion will result in the maximization of variance, which will also produce the maximum sample size. This proportion can be used to estimate variance (0.25) in the population.

Sample size for estimating Proportion (or percentage)

(Population N is large or unknown)

Standard textbook authors and researchers offer tested methods that allow studies to take full advantage of statistical measurements, which in turn give researchers the upper hand in determining the correct sample size ([Barlett et al., 2001](#)). If the population size N is large (or unknown), the first approximation developed by [Cochran \(1977\)](#), of minimum sample size n_0 is needed to estimate a population proportion p to within the margin of error E at $100(1-\alpha)$ % confidence is:

$$n_0 = \frac{\hat{p}(1 - \hat{p})z_{\alpha/2}^2}{E^2}$$

The number $z_{\alpha/2}$ is the tabulated value of z (for standard normal distribution) is determined by the desired $100(1-\alpha)$ % level of confidence. To say that we wish to estimate the population proportion to within a certain number of percentage points means that we want the margin of error ‘ E ’ to be no larger than that number (expressed as a proportion).

In planning studies, investigators should also consider attrition or loss to follow-up. The formula above gives the number of participants needed with complete data to ensure that the margin of error in the confidence interval does not exceed E .

Note that ‘ p ’ may be actual (from census) or estimated from the past experience. The formula for estimating how large a sample to take contains the number \hat{p} , which we know only after we have taken the sample. There are two ways out of this dilemma. Typically the researcher will have some idea as to the value of the population proportion p , hence of what the sample proportion \hat{p} is likely to be. For example, if last month 37% of all voters thought that state taxes are too high, then it is likely that the proportion with that opinion this month will not be dramatically different, and we would use the value 0.37 for \hat{p} in the formula.

The second approach to resolving the dilemma is simply to replace \hat{p} in the formula by 0.5 ([Krejcie & Morgan, 1970](#)). This is because if \hat{p} is large then $1 - \hat{p}$ is small, and vice versa, which limits their product to a maximum value of 0.25, which occurs when $\hat{p}=0.5$.

Example: In the absence of estimated proportion (p) we assumed that the estimated population proportion p is 50% i.e. $p=0.50$, and for 95% confidence level the value of $z=1.96$; and 4% margin of error ($E=4\%=0.04$) we need a sample size:

$$n_0 = \left(\frac{1.96 \times 0.50}{0.04} \right)^2$$

$n_0 = 600$

For $E=5\%$, $n_0=385$ and for $E=6\%$, $n_0=267$ and so on. As margin of error ‘ E ’ increases sample size ‘ n ’ decreases

The sample size for estimating Proportion (N is known and small)

If we know or can estimate the population size N , which is small, first of all, we calculate an initial sample size n_0 as before. Then the final sample size ‘ n ’ which is more precise is calculated as ([Cochran, 1977](#)).

Note: For a big population the difference between n_0 and n is negligible but for a small population the difference is appreciable.

Example: In the absence of estimated proportion (p) we assumed that the estimated population proportion p is 50% i.e. $p=0.50$, and for 95% confidence level the value of $z=1.96$ and margin of error $e=4\%=0.04$ then we need an initial sample size:

$$n_0 = \left(\frac{1.96 \times 0.50}{0.04} \right)^2 = 600$$

If we know the population N = 2000 then the final sample size n which is calculated as

$$n = \frac{600}{1 + \frac{600}{2000}} = 462$$

If N = 5000, n = 538

Note: For a big population the difference between n_0 and n is negligible but for a small population the difference is appreciable.

Minimum Sample Size for Estimating a Population Mean (Continuous Data)

If the population size N is large (or unknown), the minimum sample size n_0 is needed to estimate a population mean μ within the margin of error E at $100(1-\alpha)$ % confidence is (Cochran, 1977).

$$n_0 = \frac{z_{\alpha/2}^2 \sigma^2}{E^2} \text{ (rounded up)}$$

Where, σ = population standard deviation

- 1) σ may be actual or is estimated from the past experience. If σ not available, one could take a preliminary sample size, $n \geq 30$ to provide an estimate of σ .
- 2) If σ cannot be guessed at all or estimated otherwise, a rule of thumb for estimating σ is to take one-sixth of the range of the values the researcher expects. If researcher expects that mean age will be between 30 to 35 years then an estimate of σ

$$\sigma = 1/6(35-30) = 0.83$$

- 1) The estimated σ can be considered as $\sigma = 0.5$ (Krejcie & Morgan, 1970)
- 2) According to Israel (1992) "The disadvantage of the sample size based on the mean is that a "good" estimate of the population variance is necessary. Often, an estimate is not available. Furthermore, the sample size can vary widely from one attribute to another because each is likely to have a different variance. Because of these problems, the sample size for the proportion is frequently preferred."
- 3) Adjusted Sample size: If the population size N is known or not negligible (Cochran, 1977), the required sample size for estimating the population mean will be

$$n = \frac{n_0}{1 + \frac{n_0}{N}}$$

A Simplified Formula for Proportions

Yamane (1967) provides a simplified formula to calculate sample sizes. This formula was used to calculate the sample sizes in Tables 1 and 2 and is shown below. A 95% confidence level and $p = 0.5$ are assumed

$$n = \frac{N}{1 + N(E^2)}$$

Table 1
The sample size for big population

Size of Population(N)	Sample Size (n) for Precision (E) of:			
	±3%	±5%	±7%	±10%
500	A	222	145	83
600	A	240	152	86
700	A	255	158	88
800	A	267	163	89
900	A	277	166	90
1,000	A	286	169	91
2,000	714	333	185	95
3,000	811	353	191	97
4,000	870	364	194	98
5,000	909	370	196	98
6,000	938	375	197	98
7,000	959	378	198	99
8,000	976	381	199	99
9,000	989	383	200	99
10,000	1,000	385	200	99
15,000	1,034	390	201	99
20,000	1,053	392	204	100
25,000	1,064	394	204	100
50,000	1,087	397	204	100
100,000	1,099	398	204	100
>100,000	1,111	400	204	100

A = Assumption of normal population is poor (Yamane, 1967). The entire population should be sampled.

Table 2
Sample size for small population

Size of Population(N)	Sample Size (n) for Precision (E) of:		
	±5%	±7%	±10%
100	81	67	51
125	96	78	56
150	110	86	61
175	122	94	64
200	134	101	67
225	144	107	70
250	154	112	72
275	163	117	74
300	172	121	76
325	180	125	77
350	187	129	78
375	194	132	80
400	201	135	81
425	207	138	82
450	212	140	82

Additional information

In addition, an adjustment in the sample size may be needed to accommodate a comparative analysis of subgroups (e.g., such as an evaluation of program participants with nonparticipants).

- a) [Sudman \(1976\)](#), suggests that a minimum of 100 elements is needed for each major group or subgroup in the sample and for each minor subgroup, a sample of 20 to 50 elements is necessary.
- b) When the attribute is present 20 to 80 percent of the time (i.e., the distribution approaches normality) that 30 to 200 elements are sufficient said by [Kish \(1965\)](#).
- c) On the other hand, skewed distributions can result in serious departures from normality even for moderate size samples (Kish, 1965). Then a larger sample or a census is required.
- d) Frequently, researchers will add a buffer of 5-20% to the necessary sample sizes to achieve a desired level of power to allow for some dropout/non-participation. The sample size also is often increased by 30% to compensate for nonresponse ([Israel, 1992](#)).
- e) In telephone and face-to-face interviewing, response rates of 70 to 80 percent are common. Response rates to self-administered surveys range much more widely than this, although it is possible to achieve a response rate of 70 percent or even higher ([Spaeth, 1992](#)).
- f) To correct for the difference in design, the sample size is multiplied by the design effect (deff) which range is 1 to 2
- g) For more complex designs, e.g., stratified random samples, must take into account the variances of subpopulations, strata, or clusters before an estimate of the variability in the population as a whole can be made ([Israel, 1992](#)).
- h) [Israel \(1992\)](#), describes that virtually the entire population would have to be sampled in small populations (e.g., 200 or less) to achieve a desirable level of precision. As A census eliminates sampling error and provides data on all the individuals in the population. In addition, some costs such as questionnaire design and developing the sampling frame are "fixed," that is, they will be the same for samples of 50 or 200 ([Israel, 1992](#)).

Sample size in general (Qualitative) used in Survey

Let us determine the sample size by using a sound statistical formula for the first stage sampling. The calculation of sample size is complicated by the fact that some of the factors vary by indicators. To calculate the proper sample size, using the appropriate mathematical formula, several factors be specified and values for others be assumed or taken from previous or similar surveys. The following formula is used to calculate the sample size:

$$n_0 = \frac{\hat{p} \times (1 - \hat{p}) \times z_{\alpha/2}^2 \times (R) \times (\text{deff})}{E^2}$$

Where:

- a) n_0 is the required sample size, for the key indicator
- b) \hat{p} is the predicted or anticipated prevalence for the indicator being estimated. When prevalence for the indicator is unknown is replaced by $\hat{p}=0.5$ in the formula. This is the most conservative estimate, since it gives the largest possible estimate of n
- c) The value of $z_{\alpha/2}$ is 1.96 to achieve the 95 per cent level of confidence and the value of $z_{\alpha/2}$ is 2.57 to achieve the 99 per cent level of confidence.
- d) R is the factor necessary to raise the sample size by 5 to 20 percent for non-response
- e) The sample size is multiplied by the design effect (deff) which range from 1 to 2 if a multi-stage sampling is done instead of a simple random sample.
- f) E is the margin of error (level of accuracy) to be tolerated at the given percent level of confidence. The higher value of 'E' will yield lower sample size and a smaller value of 'E' will yield higher sample size.
- g) If the population size N is known or not negligible ([Cochran, 1977](#)), the required sample size n_0 will be adjusted to get the final sample size n as follows:

$$n = \frac{n_0}{1 + \frac{n_0}{N}}$$

Results and Analysis

A case study: How sample data can be used to discover the truth about a population

We will examine how *sample* data can be used to discover the truth about a *population*. Our population data consists of ages, weights, and heights of 1358 undergraduate students of BRAC University at Dhaka. The data are collected by the students of statistics (Summer, 2015), of BRAC University students for their coursework assignment. We will run few simulations on this data to see if we can replicate what the Central Limit Theorem tells us about sampling. We are pretending that we don't know the "true" population parameters, but in fact, we do. We will use 'R' programming for our data analysis and simulation.

Central Limit Theorem

The central limit theorem (CLT), one of the most important theorems in statistics, implies that under most distributions, normal or non-normal, the sampling distribution of the sample means will approach normality as the sample size increases (Hays, 1994). One of the simplest versions of the theorem says that it is a random sample of size n (say, n larger than 30) from an infinite population, finite standard deviation, then the standardized sample mean converges to a standard normal distribution or, equivalently, the sample mean approaches a normal distribution with mean equal to the population mean and standard deviation equal to standard deviation of the population divided by the square root of sample size n (Arsham, 2005). Without the CLT, inferential statistics that rely on the assumption of normality (e.g., two sample t-test, ANOVA) would be nearly useless, especially in the social sciences where most of the measures are not normally distributed (Micceri, 1989). It is often suggested that a sample size of 30 will produce an approximately normal sampling distribution for the sample mean from a non-normal parent distribution.

There is little to no documented evidence to support that a sample size of 30 is the magic number for non-normal distributions. Arsham (2005), claims that it is not even feasible to state when the central limit theorem works or what sample size is large enough for a good approximation, but the only thing most statisticians agree on is "that if the parent distribution is symmetric and relatively short-tailed, then the sample mean reaches approximate normality for smaller samples than if the parent population is skewed or long-tailed". It seems that "normality is a myth; there never was, and never will be, a normal distribution" (Geary, 1947). Still, normality is an assumption that is needed in many statistical tests. Determining the best way to approximate normality is the only option since true normality does not seem to exist. This theorem, also described briefly below, only implies that the sampling means are approximately normally distributed when the sample size is large enough. Micceri (1989) suggests conducting more research on the robustness of the normality assumption based on the fact that real-world data are often contaminated.

Observing the Population

We found that the average weight for our population of BRAC University students (1358 students) was 62.32 kg with standard deviation 11.79 kg. The shape of weights of the students can be visualized from the following histogram (Fig 1):

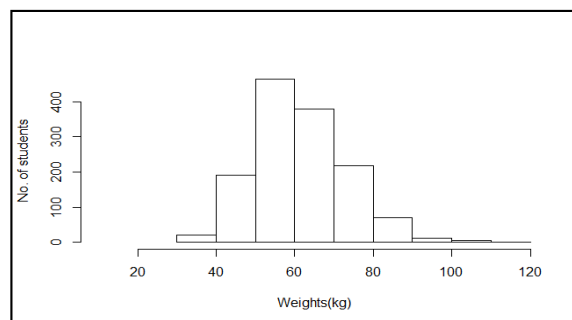


Figure 1. Histogram of weights

From the Figure 1, we can say that the weights of the students are not normally distributed which is slightly positively skewed.

Observing the Sampling Distributions

We drew 1000 samples of sizes $n=15$, $n=30$ and $n=50$ respectively from our population of students' weights. Then we calculate mean of each sample of sizes 15, 30 and 50; and got the following histograms (Fig. 2) of 1000 sample means for three different sample sizes:

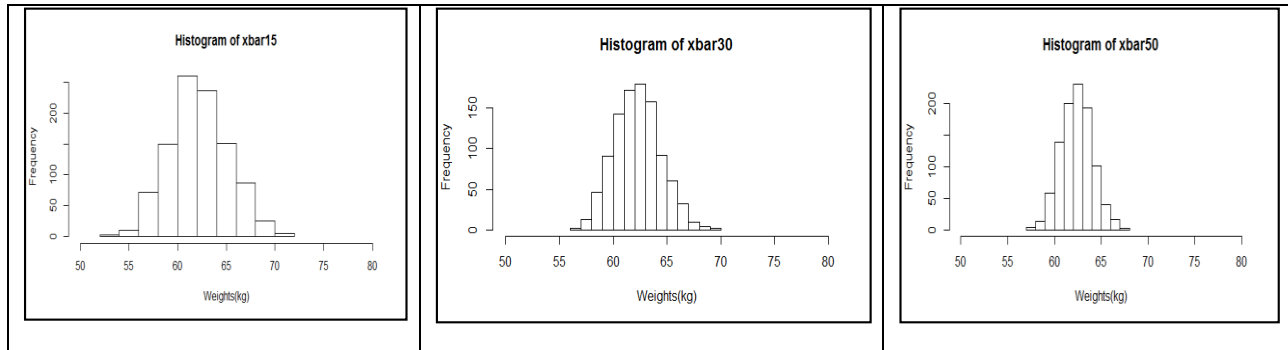


Figure 2. Histogram of weights for different sample sizes

Later we calculate the mean and standard deviation of 1000 sample means for the sample size $n=15$, $n=30$ and $n=50$ respectively. The variability of the sample means is predicted by the standard error (SE), which is obtained by dividing the population standard deviation by the square root of the sample size. The results of the sampling distribution for each sample size are summarized below table 3:

Table 3
Simulation result of different sample sizes

Results of simulation (100 times)	Sample Size		
	$n=15$	$n=30$	$n=50$
Mean	62.22	62.27	62.37
Standard deviation	3.047	2.16	1.665
Standard error	3.045	2.153	1.668

According to the Central Limit Theorem (CLT) as sample size increases the sampling distribution becomes more normal, the mean of the sampling distribution will be same as the population mean; and the variability of the sample means is predicted by the standard error (Arsham, 2005), which will be less variable as the sample size increases. As James & Berg Sjeff (2002), stated that the sample size is inversely related to the variability of sample means: the greater the sample size, the narrower the range of sample means.

From the above table (Table 3) of the sampling distribution, we can say that the sample means of weight is approximately same as the population means (62.32 kg) for all three sampling distributions and the size of standard error decreased as the sample size increased from 15 to 50.

It is also observed from the histograms (Fig. 2) of sampling means which are derived by the simulation that the if the sample size increase the shape of the sampling distribution will tend to be Normal no matter what the shape of the population.

Conclusion

From the data of 1348 students, we got the average weight for our population of BRAC University students is 62.62 kg with standard deviation 11.79 kg. We found that that the weights of the students are not normally distributed which is slightly positively skewed. We drew samples of different sizes from our population to simulate the Central

Limit Theorem. The simulation was done in 100 times with R programming. The sample means of weight is approximately same as the population means (62.32 kg) for all three sampling distributions and the size of standard error decreased as the sample size increased from 15 to 50. As we increased the size of our sample from 15 to 50, the sample means to become less variable and tended to cluster more tightly around the true mean. In other words, our sample means became better estimators of the true population mean. In addition, the shape of the distribution became more Normal as the sample size increased. So it is concluded that our simulation results were consistent with Central Limit Theorem (CLT).

Acknowledgements

I am very much grateful to my students of BRAC University for providing data.

References

- Apedoe, X. S., & Reeves, T. C. (2006). Inquiry-based learning and digital libraries in undergraduate science education. *Journal of science education and technology*, 15(5-6), 321-330.
- Arsham, H. (1998). Algorithms for sensitivity information in discrete-event systems simulation. *Simulation Practice and Theory*, 6(1), 1-22.
- Barlett, J. E., Kotrlik, J. W., & Higgins, C. C. (2001). Organizational research: Determining appropriate sample size in survey research. *Information technology, learning, and performance journal*, 19(1), 43.
- Cochran, W. G. (2007). *Sampling techniques*. John Wiley & Sons.
- Cox, D. R. (1958). Some problems connected with statistical inference. *The Annals of Mathematical Statistics*, 29(2), 357-372.
- De Calderero, R. P., Panchana, M. J. C., Lectong, D. M., & Hernández, E. H. O. (2018). Use of Concrete Debris. *International Journal of Physical Sciences and Engineering (IJPSE)*, 2(1), 1-12.
- Eckhardt, K. W., & Ermann, M. D. (1977). *Social research methods: Perspective, theory, and analysis*. Random House (NY).
- Fleishman, A. I. (1978). A method for simulating non-normal distributions. *Psychometrika*, 43(4), 521-532.
- González, A. E. D., Arauz, W. M. S., Gámez, M. R., & Alava, L. A. C. (2017). Photovoltaic Energy to Face an Earthquake. *International Journal of Physical Sciences and Engineering (IJPSE)*, 1(3), 19-30.
- Holton, E. (1997). Human resource development research handbook.
- Israel, G. D. (1992). Determining sample size.
- Israel, G. D. (1992). *Sampling the evidence of extension program impact*. University of Florida Cooperative Extension Service, Institute of Food and Agriculture Sciences, EDIS.
- Johnson, P. O. (1959). Development of the sample survey as a scientific methodology. *The Journal of Experimental Education*, 27(3), 167-176.
- Kish, L. (1965). Survey sampling.
- Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. *Educational and psychological measurement*, 30(3), 607-610.
- Kusumayanti, G. D., & Dewantari, N. M. (2017). The Influence of Low Purine Diet and Physical Activity on Changing of Uric Acid Levels in Hyperuricemia. *International Journal of Health Sciences (IJHS)*, 1(3), 1-9.
- Marshall, M. N. (1996). Sampling for qualitative research. *Family practice*, 13(6), 522-526.
- Meehan, M. L., Cowley, K. S., Finch, N. L., Chadwick, K. L., Ermolov, L. D., & Riffle, M. J. S. (2004). Special Strategies Observation System-Revised: A Useful Tool for Educational Research and Evaluation. *American Educational Research Association*.
- Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological bulletin*, 105(1), 156.
- Spaeth, J. L. (1992). Perils and pitfalls of survey research. *Allerton Park Institute (33rd: 1991)*.
- Stonehouse, J. M., & Forrester, G. J. (1998). Robustness of the t and U tests under combined assumption violations. *Journal of Applied Statistics*, 25(1), 63-74.
- Sudman, S. (1976). *Applied sampling* (No. 04; HN29, S8.). New York: Academic Press.
- Thompson, B. (2003). Guidelines for authors reporting score reliability estimates. *Score reliability: Contemporary thinking on reliability issues*, 91-102.