The International Journal Of Engineering And Science (IJES)
|| Volume || 5 || Issue || 5 || Pages || PP -59-62 || 2016 ||
ISSN (e): 2319 – 1813 ISSN (p): 2319 – 1805

Application of Central Limit Theorem to Study the Student Skills in Verbal, Aptitude and Reasoning Test

[1]Dr.P.MEGALADEVI, [2]Ms.M.PRIYA , [3]Dr. R.VASUKI

[1]***DEAN*** */ Department of Management Studies*
Jayshriram Group of Institutions, Tirupur, Tamil Nadu, India
[2]***AP*** */ Department of Science & Humanities*
Jayshriram Group of Institutions, Tirupur, Tamil Nadu,India
[3]***Professor/*** */ Department of Management Studies*
Jayshriram Group of Institutions, Tirupur, Tamil Nadu, India

ABSTRACT

Through this paper we analyses the application of the central limit theorem to study the Verbal, Apptitude and Reasoning skills of students. The planning of teaching is based on the mathematical knowledge about the theorem. The different meanings of this theorem were analyzed using the history of its development and previous research studies related to this theorem. Results at the end of this work will serve to improve the correct application of different elements of meaning for central limit theorem when solving the selected problem and to prepare new proposals to teach statistics to students. The central limit theorem forms the basis of inferential statistics and it would be difficult to overestimate its importance. In a statistical study, the sample mean is used to estimate the population mean. However, the number of different samples (of a given size) that could be taken is extremely large and these different samples would have different means. Some would be lower than the mean of the population and some would be higher.The central limit theorem states that, for samples of size n from a normal population, the distribution of sample means is normal with a mean equal to the mean of the population and a standard deviation equal to the standard deviation of the population divided by the square root of the sample size. (For suitably large sample sizes, the central limit theorem also applies to populations whose distributions are not normal.)

Keywords: *CLT, Population mean, Sample mean, Confidence intervals, Hypothesis.*

Date of Submission: 02 May 2016 Date of Accepted: 10 May 2016

I. INTRODUCTION

The practical significance of the central limit theorem is twofold. First, as the sample size increases the standard deviation of the sample means decreases. Consequently, we can be assured that larger samples tend to yield more accurate estimates of the population mean than smaller samples.Second, it is very unlikely that a sample mean will exactly equal the population mean. Even if it did, we wouldn't know it. Consequently, the sample mean is used to create a range of values (called a confidence interval) that is likely to contain the population mean. The purpose of descriptive statistics is to allow us to more easily grasp the significant features of a set of sample data. However, They tell us little about the population from which the sample was taken. Inferential statistics is the branch of statistics that deals with using sample data to make valid judgments about the population from which the data came.

Inferential statistics is a powerful technique used by researchers and practioners for a wide array of purposes such as testing the falsehood of theories and identifying important factors that may influence relevant outcome. The (CLT), one of the most important theorems in statistics, implies that under mot distributions, normal or non-normal, the sampling distribution of the sample mean will approach normality as the sample size increases(Hayes,1994).Without the CLT, inferential statistics that rely on the assumption of normality(e.g., two-sample t-test,ANOVA) would be nearly useless, in the social sciences where most of the measure are not normally distributed (Miceri,1989).Researchers like Mendez, 1991 carried out assessment on how students understand some of basic features of simple versions of TCL. Later del Mas, Garfield and chance (1991) organized teaching experiments to see how students could develop an intuitive understanding of sampling distribution deducted from TCL. In our research we take in to consideration of nearly 150 students verbal, aptitude and reasoning scores that they obtained to study the application of CLT.

II. DESIGNING A STUDY PROCESS OF THE CENTRAL LIMIT THOREM

A Brief note on History of CLT

Abraham De Moivre published "Doctrine of chances", where he included a proof for the normal approximation to the Binomial distribution.2. In 1809 Laplace gave a proof of CLT for mutually independent, identically distributed discrete random variables with finite mean and variance; in 1824 Poisson extended the result to continuous symmetric bounded random variables:Chebyscheff(1887) generalized the theorem for distributions with infinite range, provided some moments are finite; Later the theorem was expanded to non-identically distributions or dependent variables.3. Some researchers like Dirichlet & Cauchy worried about the precision of the estimation.4. Linderberg,Levy,Feller and others provided sufficient and complete condition for the theorem.

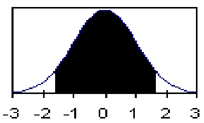
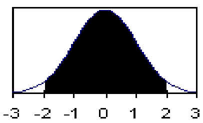
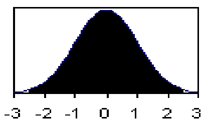
In this way the TCL developed gradually and paved way for developing statistical and mathematical tools by above mentioned authors. The latest CLT is used in a variety of indirect applications, such as finding asymptotic sampling distribution for different parameters, building confidence intervals or testing hypotheses for these parameters or finding an adequate sample size to carry out an inference with a given precision.

Confidence Interval used in CLT

A Confidence interval is a range of values based on the sample mean, the sample size and either the sample or the population standard deviation that is likely to contain the population mean. The confidence level is the proportion of samples that will yield a confidence interval that actually contains the population mean .For example, if the confidence level is 95% then the confidence interval generated using the techniques described below ill contain the population mean. The remaining 5% of the samples will result in confidence intervals that do not contain the population meanwhile we can set the confidence level at any value we wish, the most common confidence levels are 90%, 95% and 99%.It stands to reason that larger intervals are more likely to include the population mean than smaller ones. Consequently, higher confidence levels are associated with wider intervals.

Population Standard deviation

A confidence interval can be derived from the distribution of sample means using CLT if the population standard deviation is given. It is derived by the formulaWhere x-bar is the sample mean, sigma is the population standard deviation, n is the sample size, and z* is the critical z value. The critical value z and its negative delimit a central area under the standard normal curve equal to the desired confidence level.

Confidence Level		
90%	95%	99%
z* = 1.645	z* = 1.960	z* = 2.576
		
The area between -1.645 and 1.645 is 0.90	The area between -1.960 and 1.960 is 0.95	The area between -2.576 and 2.576 is 0.99

For 90% the confidence level z = 1.645
For 95% the confidence level z = 1.960
For 99% the confidence level z = 2.576

The above formula is changed as follows if the population standard deviation is unknown.Where x-bar is the sample mean, s is the sample standard deviation, n is the sample size, and "t" is the theoretical t value. The critical value of "t" can be taken from Student's t distribution.

The table below illustrates a 90%, a 95%, and a 99% confidence interval. Notice that the only thing that changes in the calculation is the critical value z*.

III. HYPOTHESIS

In inferential statistics, a study is often performed to allow the researcher to investigate two possible hypotheses about a population. The null hypothesis states that a population parameter has some specific value that is assumed to be correct. The alternate hypothesis challenges this assumption. The alternate hypothesis challenges this assumption. The statistical study results in a decision to accept or reject the null hypothesis. The null hypothesis asserts that the population mean is some specific value. This value is often based on previous statistical studies.

Null & Alternate Hypothesis

In inferential statistics, a study is often performed to allow the researcher to investigate two possible hypotheses about a population. The null hypothesis states that a population parameter has some specific value that is assumed to be correct. The alternate hypothesis challenges this assumption. The statistical study results in a decision to accept or reject the null hypothesis. (Statisticians do not accept the alternate hypothesis, they reject the null hypothesis. This is primarily because the null hypothesis is specific while the alternate hypothesis is vague.)

One such inference test involves the mean of a population. The null hypothesis asserts that the population mean is some specific value. This value is often based on previous statistical studies. The alternate hypothesis can have one of three forms as population mean is not equal to a specific value, population mean is less than a specific value and population mean is more than this specific value. Null Hypothesis is denoted by H_0.Here the null hypothesis is stated that the mean skill of the selected students skill is equal to seventy five whereas the Alternate Hypothesis which is represented by H_1 is stated as the mean skill of the selected students is not equal to seventy five.

In the study data were collected from 150 students who appeared for Verbal test, Reasoning test, Aptitude test. The survey is carried out in an Engineering College among 150 graduates who came for personality development for Managerial Skills. The Verbal test is conducted for thirty marks. The Reasoning test is conducted for thirty marks and the Aptitude test is conducted for fourty marks. The average and standard deviation for twenty five students(I Batch), then for fifty students(II Batch), hundred students(III Batch) and then for one fifty students(IV Batch) is arrived and analyzed.

IV. RESULTS AND DISCUSSIONS

The results of the study are presented in the following tables.

Table-1.1 Average and Standard Deviation of Respondents

Sample size	**Mean of Verbal test**	**S.D for Verbal test**	**Mean of Reasoning test**	**S.D for Reasoning test**	**Mean of Aptitude test**	**S.D for Aptitude test**
25	20.4	9.265	18.24	9.265	32.68	8.806
50	20.46	9.12	18.46	9.120	32.8	9.303
100	20.14	8.40	18.14	8.409	33.33	8.757
150	20.26	7.93	18.53	7.884	32.98	8.885

Table 1.1 gives the Mean and Standard deviation of Verbal, Reasoning and Aptitude test for Set I(25 Students), Set II (50 Students), Set III (100 Students) and for Set IV (150 Students) are listed in the above table.

Table-1.2 – Range of Confidence interval for Verbal, Reasoning and Aptitude for Set I

Type	**Sample size**	**Average**	**"t"value**	**Standard deviation**	**Level of significance**	**Confidence interval**
Verbal	25	20.4	2.145	9.265	0.95	(16.425,24.374)
Reasoning	25	18.24	2.145	9.265	0.95	(14.276,22.224)
Aptitude	25	32.68	2.145	8.806	0.95	(28.903,36.457)

From the Table 1.2, we may come to know that the sample size is below thirty. So we apply Students "t" distribution instead of Z distribution. The result may be concluded that for 95% Confidence limit and sample size 25, the limit value lies between (16.425,24.374) for the selected population mean in case of verbal test whereas for Reasoning ,the average Confidence limit lies between (14.276,22.224) while the limits range from 28.90 and 36.46 in case of Aptitude.

Table-1.3 - Range of Confidence interval for Verbal, Reasoning and Aptitude for Set II

Type	Sample size	Average	Z value	Standard deviation	Level of significance	Confidence interval
Verbal	50	20.46	1.96	9.120	0.95	(17.94,22.980)
Reasoning	50	18.46	1.96	9.120	0.95	(15.93,20.987)

Aptitude	50	32.8	1.96	9.303	0.95	(35.37,30.230)

When we change the sample size from 25 to 50, we use Z value in Table 1.3 and at we may see that at 95% confidence limit the average limit of Verbal test lies between (17.94, 22.98). Applying the same procedure for Reasoning , we get the limits range from 15.93 to 20.987.It is observed that the 95% confidence limit for Aptitude for a sample size 50 lies between 35.37 and 30.23.

Table-1.4 - Range of Confidence interval for Verbal, Reasoning and Aptitude for Set III

Type	Sample size	Average	Z value	Standard deviation	Level of significance	Confidence interval
Verbal	100	20.46	1.96	8.409	0.95	(21.788,18.492)
Reasoning	100	18.46	1.96	8.409	0.95	(16.492,19.780)
Aptitude	100	33.33	1.96	8.757	0.95	(31.613,35.046)

From the Table 1.4, we may come to know that the sample size is 100. So we apply Z distribution. The result may be concluded that for 95% Confidence limit and sample size 100, the limit value lies between (21.788,18.492) for the selected population mean for verbal test whereas for Reasoning ,the Confidence limit lies between (16.492,19.780) while the limit range lies in (31.613,35.046) for Aptitude.

Table-1.5 - Range of Confidence interval for Verbal, Reasoning and Aptitude for Set IV

Type	Sample size	Average	Z value	Standard deviation	Level of significance	Confidence interval
Verbal	150	20.26	1.96	7.938	0.95	(18.990,21.530)
Reasoning	150	18.53	1.96	7.884	0.95	(17.268,19.792)
Aptitude	150	32.98	1.96	8.885	0.95	(34.402,31.558)

When we change the sample size to 100 , we use Z value in Table 1.5 and at we may see that at 95% confidence limit the average limit of Verbal test lies between (18.990,21.530).Applying the same procedure for Reasoning , we get the limits range from 17.268 to 19.792. It is observed that the 95% confidence limit for Aptitude for a sample size 100 lies between 34.402 and 31.558.

V. CONCLUSION

The CLT is an essential tool in statistical inference. The technological development helps the students to understand and perform it in graphical, simulation capabilities and internet applets. In Our study if the confidence interval includes 75 (the mean under the null hypothesis), we accept the null hypothesis. If the confidence interval does not contain the mean given in the null hypothesis, we would reject the null hypothesis. The next stage in our research is experimenting this proposal for normal and non-normal size factors in Management Investment sectors. This may help the students to segregate the data for the various distribution selection based on data.

REFERENCES

[1] Arsham, H. (2005). System Simulation: The Shortest Route to Applications, Version 9.Retrieved 5/31/06 from http://home.ubalt.edu/ntsbarsh/simulation/sim.htm.

[2] Fleishman, A. I. (1978). A Method For Simulating Non-Normal Distributions.Psychometrika, 43 (4), 521-532.

[3] Hays, W. L. (1994). Statistics (5th ed.). New York: Holt, Rinehart and Winston. Micceri, T. (1989). The Unicorn, The Normal Curve, and Other Improbable Creatures.Psychological Bulletin, 105 (1), 156-166.

[4] Fischer, Hans (2011), A History of the Central Limit Theorem: From Classical to Modern Probability Theory, Sources and Studies in the History of Mathematics and Physical Sciences, New York: Springer, doi:10.1007/978-0-387-87857-7,ISBN 978-0-387-87856-0, MR 2743162, Zbl 1226.60004 (Chapter 2: The Central Limit Theorem from Laplace to Cauchy: Changes in Stochastic Objectives and in Analytical Methods, Chapter 5.2: The Central Limit Theorem in the Twenties) http://onlinestatbook.com/s

[5] Fleishman, A. I. (1978). A Method For Simulating Non-Normal Distributions. Psychometrika, 43 (4), 521-532.

[6] Micceri, T. (1989). The Unicorn, The Normal Curve, and Other Improbable Creatures.*Psychological Bulletin*, *105* (1), 156-166.