|  |  |  |
| --- | --- | --- |
| Related image | **KONERU LAKSHMAIAH EDUCATION FOUNDATION**  (Deemed to be University estd, u/s, 3 of the UGC Act, 1956) (NAAC Accredited “A++” Grade University)  Green Fields, Guntur District, A.P., India – 522502  **Department of Computer Science and Engineering**  (DST - FIST Sponsored Department) |  |

**B.Tech. II CSE(H) PROGRAM**

**A.Y. 2023-24 ODD, Semester-II**

**Course Code: 22MT2005**

**PROBABILITY, STATISTICS AND QUEUING THEORY**

**Course Outcome-2**

**Session 20:** **Central Limit Theorem and its Applications**

1. **Course Description (Description about the subject)**
2. **Aim**

To explain the central limit theory and its applications in real life problems.

1. **Instructional** **Objectives (Course Objectives)**

To understand the central limit theory and its applications.

1. **Learning** **Outcomes (Course Outcome)**

**CO2**: Students will be able to explain the central limit theory and its applications in real life problems in statistical calculations.

1. **Module** **Description** **(CO-2 Description)**

The central limit theorem relies on the concept of a sampling distribution, which is the [probability distribution](https://www.scribbr.com/statistics/probability-distributions/) of a statistic for a large number of [samples](https://www.scribbr.com/methodology/population-vs-sample/) taken from a population.

1. **Session** **Introduction**

The central limit theorem is a statistical theorem that states that, given a sufficiently large sample size, the sampling distribution of the mean will approximate a normal distribution regardless of the distribution of the population. This means that, even if the population distribution is not normal, the distribution of the sample means will be approximately normal as the sample size gets larger.

1. **Session description**

The central limit theorem is a very important theorem in statistics because it allows us to use the normal distribution to make inferences about populations even when we do not know the population distribution. For example, we can use the central limit theorem to calculate confidence intervals and hypothesis tests for the mean of a population.

The central limit theorem has a few assumptions that must be met in order for it to hold. These assumptions are:

* The samples must be independent.
* The samples must be identically distributed.
* The sample size must be sufficiently large.

The central limit theorem is a powerful tool that can be used to make inferences about populations. However, it is important to remember that the central limit theorem only holds under certain conditions. If these conditions are not met, then the central limit theorem may not be accurate.

Some examples of how the central limit theorem can be used:

* to calculate a confidence interval for the mean of a population.
* to conduct a hypothesis test for the mean of a population.
* to compare the means of two populations.
* to estimate the standard deviation of a population.

The central limit theorem is a fundamental theorem in statistics that is used in many different applications. It is a powerful tool that can be used to make inferences about populations even when we do not know the population distribution.

The central limit theorem relies on the concept of a sampling distribution, which is the [probability distribution](https://www.scribbr.com/statistics/probability-distributions/) of a statistic for a large number of [samples](https://www.scribbr.com/methodology/population-vs-sample/) taken from a population. Imagining an experiment may help you to understand sampling distributions:

**Example**: Suppose that you draw a [random sample](https://www.scribbr.com/methodology/simple-random-sampling/) from a population and calculate a [statistic](https://www.scribbr.com/statistics/parameter-vs-statistic/) for the sample, such as the mean.

Now you draw another random sample of the same size, and again calculate the [mean](https://www.scribbr.com/statistics/mean/).

You repeat this process many times, and end up with a large number of means, one for each sample.

The distribution of the sample means is an example of a sampling distribution.

The central limit theorem says that the sampling distribution of the mean will always be **normally distributed**, as long as the sample size is large enough. Regardless of whether the population has a normal, Poisson, binomial, or any other distribution, the sampling distribution of the mean will be normal.

A normal distribution is a symmetrical, bell-shaped distribution, with increasingly fewer observations further from the center of the distribution.

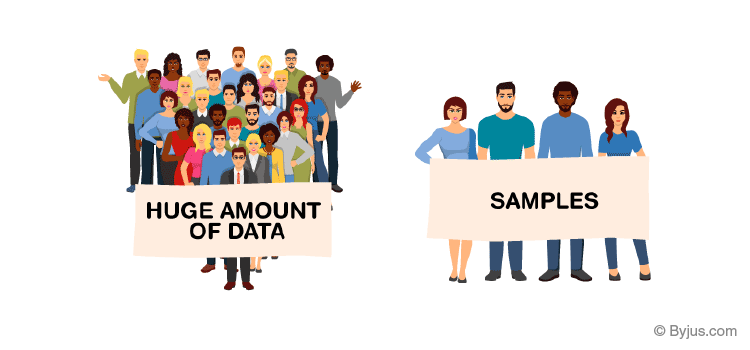
The Central Limit Theorem defines that the mean of all the given samples of a population is the same as the mean of the population(approx.) if the sample size is sufficiently large enough with a finite variation. It is one of the main topics of statistics.

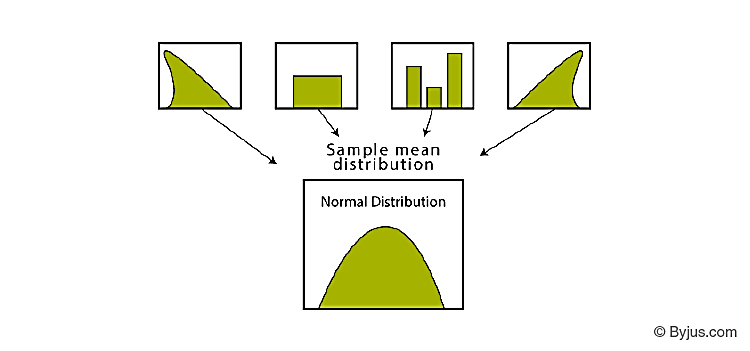
* The **mean** of the sampling distribution is the mean of the population.
* The **standard** **deviation (σx)** of the sampling distribution is the standard deviation of the population divided by the square root of the sample size.

**Central Limit Theorem Example**

Let us take an example to understand the concept of Central Limit Theorem (CLT):

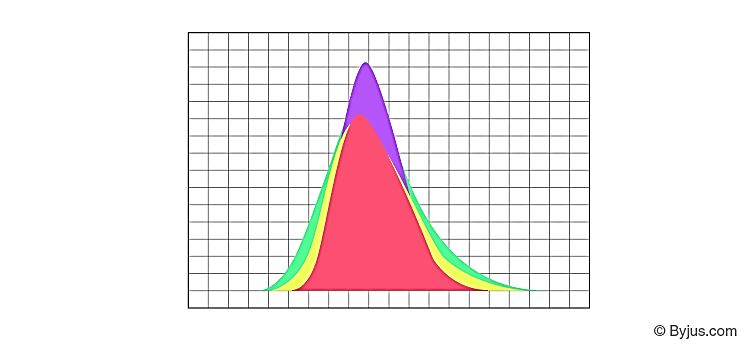
Suppose you have 10 teams in your school (Sports). Each team will have 100 students in it. Now, we want to measure the average height of the students in the sports team. The simplest way to do would be to find the average of their heights. The first step in this would be to measure the weight of all the students individually and then add them. Then, Divide the sum of their weights with the total number of students. This way we will get the average height. But this method will not make sense for long calculations as it would be tiresome and very long.





So, we will use CTL(Central Limit Theorem) to make the calculation easy. In this method, we will randomly pick students from different teams and make a sample. Each sample will include 20 students. Then, we will follow the following steps to solve it.

1. Take all these samples and find the mean for each individual sample.
2. Now, Find the mean of the sample means.
3. This way we will get the approximate mean height of the students in the sports team.
4. We will get a bell curve shape if we will find the histogram of these sample mean heights.



## **Central Limit Theorem Formula**

The central limit theorem is applicable for a sufficiently large sample size (n≥30). The formula for central limit theorem can be stated as follows:

and

Where, µ =population mean, σ = population standard deviation,

µx = sample mean, σx = sample standard deviation, n =sample size

**Applications of Central Limit Theorem**

|  |  |
| --- | --- |
| **Statistical Application of CLT** | **Practical Significance of CLT** |
| If the distribution is not known or not normal, we consider the sample distribution to be normal according to CLT. As this method assumes that the population given is normally distributed. This helps in analyzing data in methods like constructing confidence intervals. | one of the most common applications of CLT is in election polls. To calculate the percentage of persons supporting a candidate which are seen on news as confidence intervals. |
| To estimate the population, mean more accurately, we can increase the samples taken from the population which will ultimately decrease the sample means deviation. | It is also used to measure the mean or average family income of a family in a particular region. |
| To create a range of values which is likely to include the population mean, we can use the sample mean. |  |

1. **Activities/ Case studies/related to the session.**

**NA**

1. **Examples & contemporary extracts of articles/ practices to convey the idea of the Session**
2. **SAQ's-Self Assessment Questions**

1.Explain Central Limit Theorem

2.Explain Graph of Central Limit theorem

1. **Summary**

The central limit theorem can be used to illustrate the law of large numbers. The law of large numbers states that the larger the sample size you take from a population, the closer the sample mean.

1. **Terminal Questions**
2. A study involving stress is conducted among the students on a college campus. *The stress scores follow a uniform distribution* with the lowest stress score equal to one and the highest equal to five. Using a sample of 75 students, find:
3. the probability that the *mean stress score* for the 75 students is less than 2
4. the 90th percentile for the *mean stress score* for the 75 students
5. the probability that the *total of the 75 stress scores* is less than 200
6. the 90th percentile for the *total stress score* for the 75 students
7. Find the 90th percentile for the mean of 75 stress scores. Draw a graph.
8. Suppose that a market research analyst for a cell phone company conducts a study of their customers who exceed the time allowance included on their basic cell phone contract. The analyst finds that for those people who exceed the time included in their basic contract, the **excess time used** follows an exponential distribution with a mean of 22 minutes.

Consider a random sample of 80 customers who exceed the time allowance included in their basic cell phone contract.

1. U.S. scientists studying a certain medical condition discovered that a new person is diagnosed every two minutes, on average. Suppose the standard deviation is 0.5 minutes and the sample size is 100.
2. Find the median, the first quartile, and the third quartile for the sum of sample times of diagnosis in the United States.
3. Find the probability that a diagnosis occurs on average between 1.75 and 1.85 minutes.
4. Find the value that is two standard deviations above the sample mean.
5. Find the IQR for the sum of the sample times.
6. **Case Studies (CO Wise)**

**NA**

1. **Answer Key**

**NA**

1. **Glossary**

**NA**

1. **References of books, sites, links Textbooks:**

**Textbooks:**

1. Probability and Statistics Rukmangad Achari E. and E. Keshava Reddy
2. Probability and Statistics for Engineers and Scientists” Ronald E. Walpole, Sharon L. Myers and Keying Ye 8th Edition Pearson pub
3. Probability & Statistics for Engineers Dr. J. Ravichandran first Edition Wiley-India

**Reference books:**

1. Hossein Pishro-Nik, Introduction to Probability, Statistics, and Random Processes, 2014, by Kappa Research LLC; ISBN-13: 978-0990637202

**Web Resources**

1. https://ncert.nic.in/textbook.php?kemh1=0- 16
2. https://ncert.nic.in/textbook.php?jemh1=ps-15
3. **Keywords**

Central limit theorem, population, sample