Frequently Asked Questions

1. State Central Limit Theorem.

Answer:
Suppose X1 ,X2,...Xn, be n independent random variables having the same probability density function each with E(Xi)=μ and V(Xi) = σ^2, for i=1,2 ,...,n then Sn = X1+X2+....+Xn is approximately Normal with mean nμ and variance nσ^2. Also

$$Z = \frac{\overline{X} - n\mu}{\sqrt{n}\sigma} \text{ is asymptotically N(0,1)}$$

2. When do we say that a random variable is asymptotically Normally distributed?

Answer:

If the distribution of a random variable Y depends on a parameter n and if there exists two quantities μ or σ (which may or may not depend on n such that

$$\lim_{n\to\infty} P\left[\frac{Y-\mu}{\sigma} \le t\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} \exp(-x^2/2)dx$$

Then, we say that Y is asymptotically normally distributed with mean μ and variance σ^2.

3. State the general assumptions of Central Limit Theorem

Answer:

General assumptions of CLT are:

- The variables are independent
- All the variables follow same distribution
- The mean and variance of all the variables exists
- The mean and variance of all the variables are same
- Number of variables ate very large (that is n tends to infinity)

4. What is the implication of Normal convergence?

Answer:

The important implication of the Normal convergence is that for large n each percentage point of the distribution of a R.V Y will be close to the value computed from a Normal distribution with mean μ and variance σ^2, independent of whether Y has these moments or not.

In Normal convergence the limiting distribution is always continuous (in fact Normal) no matter whether Y is continuous or not.

5. Briefly explain the History of CLT.

Answer:

Historically, the first CLT was obtained in the first part of 18^{th} century by De Moivre who sharpened Bernoulli's theorem and stated that the variables of a Bernoulli sequence obey the Central limit law. The present form is due to Laplace and hence it goes by the name of both being called De Moivre- Laplace Theorem.

Laplace working for about 20 years in this problem was the first to suggest that the Normal distribution is the limiting law of a large variety of cases. The first turning point for the Central Limit problem was the popular Liapounov's theorem given in 1901. There have been many studies of the problem since then all aimed at improving it.

The next significant step in this direction came in 1922 when Linderberg gave the sufficient condition, which was later in 1945 shown by Feller to be necessary too. The case with dependent variables has also been studied. But here we shall be concerned with the classical problem where variables are independent

6. What is the significance of De Moivre–Laplace theorem?

Answer:

In probability theory, the D**e Moivre–Laplace theorem** is a normal approximation to the Binomial distribution. It is a special case of the central limit theorem. It states that the binomial distribution of the number of "successes" in *n* independent Bernoulli trials with probability *p* of success on each trial is approximately a normal distribution with mean *np* and standard deviation $\sqrt{npq}$, if *n* is very large and some conditions are satisfied.

The theorem appeared in the second edition of *The Doctrine of Chances* by Abraham de Moivre, published in 1738. The "Bernoulli trials" were not so-called in that book, but rather de Moivre wrote about the probability distribution of the number of times "heads" appears when a coin is tossed 3600 times.

7. State De Moivre–Laplace theorem.

Answer:

The sum of *n i.i.d*. Bernoulli RV's with success probability *p* is a binomial RV. As $n \to \infty$, does the CLT also say this sum approaches a Gaussian (Normal) with mean *np* and variance npq

As *n* grows large, for *k* in the neighbourhood of *np* we can approximate

$$\binom{n}{k} p^k q^{n-k} \simeq \frac{1}{\sqrt{2\pi npq}} e^{-(k-np)^2/(2npq)}, \quad p+q=1,\ p>0,\ q>0$$

in the sense that the ratio of the left-hand side to the right-hand side converges to 1 as $n \to \infty$.

8. Narrate the history of CLT, which moved from De Moivre's to Laplace.

Answer:

History of CLT from De Moivre's to Laplace

- De Moivre investigated the limits of the binomial distribution as the number of trials increases without bound and found that the function exp($-x^2$) came up in connection with this problem.
- The formulation of the normal distribution, $(1/\sqrt{2\pi})\exp(-x^2/2)$, came with Thomas Simpson.
- This idea was expanded upon by the German mathematician Carl Friedrich Gauss who then developed the principle of least squares
- Independently, the French mathematicians Pierre Simon de Laplace and Legendre also developed these ideas. It was with Laplace's work that the first inklings of the Central Limit Theorem appeared.
- In France, the normal distribution is known as Laplacian Distribution; while in Germany it is known as Gaussian.

9. Briefly explain the initial work that lead to CLT.

Answer:

Laplace was calculating the probability distribution of the sum of meteor inclination angles. He assumed that all the angles were r.v's following a triangular distribution between 0 and 90 degrees. He faced with the following problems

- The deviation between the arithmetic mean which was inflicted with observational errors, and the theoretical value
- The exact calculation was not achievable due to the considerable amount of celestial bodies

Hence, he decided to obtain solution to the problems through an approximation, which lead to CLT.

10. Obtain CLT as a generalization of large numbers.

Answer:

If $\overline{X}n = \sum_{k=1}^{n} \frac{Xk}{n} = \frac{Sn}{n}$ **and σ (Sn) = S.D(Sn)= sn**

$$\frac{\overline{X}n - E(\overline{X}n)}{\sqrt{V(\overline{X}n)}} = \frac{Sn - E(Sn)}{sn} = Sn^*$$ is a Standard Normal variate with mean 0 and variance unity. If Xk's are independent and identically distributed, $V(Sn)=n\sigma^2$ where $\sigma^2 = V(Xk)$, $\mu = E(Xk)$ so that

$$Sn^* = \frac{Sn - n\mu}{\sigma\sqrt{n}} = \frac{\overline{X}n - \mu}{\frac{\sigma}{\sqrt{n}}}$$

If $P\left[\left|Sn^*\right| \le x\right] \to \frac{1}{\sqrt{2\pi}} \int_{-\infty} e^{-\frac{y^2}{2}} dy = \phi(x)$

$$P\left[\left|Sn^*\right| \le x\right] \to \frac{1}{\sqrt{2\pi}} \int_{-x}^{x} e^{-\frac{y^2}{2}} dy = \phi(x) - \phi(-x) \text{ or}$$

$$P\left[\left|\overline{X}n - \mu\right| \le (x\sigma)/\sqrt{n}\right] \to \frac{1}{\sqrt{2\pi}} \int_{-x}^{x} e^{-\frac{y^2}{2}} dy = \phi(x) - \phi(-x)$$

By taking $(x\sigma)/\sqrt{n} = \varepsilon$ arbitrary that is, $x = \varepsilon\sqrt{n}/\sigma$ we see that as $n \to \infty$

$P\left[\left|\overline{X}n - \mu\right| \le \varepsilon\right] \to 1$. Thus

$\overline{X}n \to \mu$ in probability which is the statement of the Weak Law of Large Numbers. Thus CLT gives the probability bound for $\left[\left|\overline{X}n - \mu\right|\right]$ while Law of large numbers (LLN) gives only the limiting value. For example, Bernoulli Law of large numbers says that Sn/n tends to p in probability. De Moivre Laplace CLT gives us the bound for the values of

$\frac{Sn - np}{\sqrt{npq}}$ in terms of Normal probability law., thus providing an estimate of the rate of convergence. When Xk's are independent but not identically distributed CLT might hold but LLN might not hold or LLN might hold but CLT might not hold

11. What is Bernoulli's idea behind CLT, which inspired De Moivre?

Answer:

- In 1777, Johann Bernoulli took a different approach.
- He thought that scientists were more likely to arrive at a measurement near the middle of a reasonable range or error than near the outer edges.

The true error curve

- The true curve representing errors, the normal curve, was actually discovered in 1733 by French mathematician Abraham de Moivre, but it was not recognized as the curve representing errors until Gauss' 1809 published it what we know as Normal curve

12. Explain De Moivre's contribution to Central Limit Theorem.

Answer:

- Consider a Bernoulli Trial - An experiment with only 2 possible outcomes: success or failure

- De Moivre explored the number of successes that occurred in a given number of trials, focusing only on experiments that had equal chances of success and failure.
- He looked at the probability of achieving 1, 2… n-1, n successes.
- De Moivre discovered that as he graphed the probability of each number of successes, it formed a bell-shaped curve.
- He then found the probability density function of the Normal distribution to represent the curve:
- Where a random variable, x = number of successes, mean is represented by μ and the standard deviation by σ
- De Moivre then used this ***continuous*** function to approximate the probability of the ***discrete*** number of successes.
- To do this, he had to integrate the function over the appropriate limits.
- To make the integration easier, De Moivre standardized the variable and also needed to adjust the limits.
- He standardized the variable using the following formula:

  $X-\mu /\sigma = X^*=(X- n/2)/ \sqrt{n}/2$

Where, n/2 is the mean and $\sqrt{n}/2$ is the standard deviation for the specific Bernoulli Trials he studied.

- So for his approximation, De Moivre used the following formula:
- $$\lim_{n\to\infty} P[a^* \le X^* \le b^*] = \frac{1}{\sqrt{2\pi}} \int_{a^*}^{b^*} \exp(-t^2/2)dt$$

  Where a* and b* are the appropriate adjusted limits of integration.

13. Write a note on Laplace's contribution to Central Limit Theorem

Answer:

Laplace contribution

- He expanded De Moivre's approximation to Bernoulli trials with any probability of success.
- More importantly, the first Central Limit Theorem was credited to him.
- This theorem allowed scientists to find how many samples they needed to make their average reasonably accurate.
- Laplace recognized that the sum of independent, identically distributed random variables follows a normal curve as the number of trials, or summands, increases.
- For our purposes, "independent, identically distributed random variables" can simply translate to repeating an identical experiment.

14. Give some examples for the application of De Moivre-Laplace CLT.

Answer:

- Playing repeated games of blackjack at a casino
- Taking many measurements of the distance between Earth and a given star
- Performing quality tests on many samples of a product at a factory

15. Each time Jim charges an item to his credit card, he rounds the amount to the nearest dollar in his records. If he has used his credit card 300 times in the past 12 months, what is the probability that his record differs from the total expenditure by, at most, $10?

Answer:

Let the total expenditure be $\sum X_i$, $i=1\ldots 300$, $n=300$
We want to find:
$P(-10 \leq \sum X_i \leq 10)$

An error of any amount is equally likely over the interval –0.50 to 0.50
So $\mu=0$ and S.D :$\sigma=1/\sqrt{12}$

$\sum X_i \sim N(n\mu, n\sigma^2)$ which implies $\sum X_i \sim N(300*0, 300 * 1/12)$

First standardize: $Z= \sum X_i- (300) *0/\sqrt{300}/\sqrt{12}$

$P(-10 \leq \sum X_i \leq 10)$ implies

$P(-10-300(0)/\sqrt{300}/\sqrt{12} \leq Z \leq 10-300(0)/\sqrt{300}/\sqrt{12})$

Simplifying $P(-2 \leq Z \leq 2)=0.9545$
Therefore, the probability that Jim's records are off by at least $10 is 95%.