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

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

2. What do you mean by WLLN?

Answer:

Suppose $\{X_i\}$ is an infinite sequence of i.i.d. random variables with common mean μ, If we define Y_n as the r.v. equal to the mean of the first n X_is,. Then, for any ε, the probability for a realization of Y_n to fall more than ε away from μ tends to 0 as n grows without limit.

or No matter how small ε, all you have to do ,to make the probability for the mean of the first n terms to differ from the mean μ by more than ε to be as small as you wish, is to make n large enough.

In the vocabulary of Estimation, the WLLN states that the sample mean is a consistent estimator of the population mean. That is a sample mean converges in probability to the population mean

3. What are the conditions required for the existence of WLLN?

Answer:

For the existence of the law we assume the following conditions

i) E (Xi) exists for all i

ii) Bn= V(X1+X2+…. +Xn) exists and

iii) $\frac{B_n}{n^2} \to 0$ As $n \to \infty$

4. If xi's have only two values with equal probabilities i^{α} and $(-i)^{\alpha}$. Show that a law of large numbers can be applied to the independent variables x1,x2,..,xn only if $\alpha<1/2$

Answer:

Given Xi takes only two values i^{α} and $(-i)^{\alpha}$

For the given problem let $P[Xi=i^{\alpha}]=P(xi=(-i)^{\alpha})=1/2$

$E(Xi) = \sum xi\ p(x\ i)= i^{\alpha} *1/2+(-i)^{\alpha}*1/2=0$

$V(xi)=E(Xi^2)-[E(Xi)]^2$

$E(Xi^2) = \sum xi^2\ p(x\ i)= i^{2\alpha}*1/2+(-i)^{2\alpha}*1/2=i^{2\alpha}$

$V(xi)=E(Xi^2)-[E(Xi)]^2 = i^{2\alpha}$

$Bn=V(X1+X2+\ldots+Xn)=1^{2\alpha}+2^{2\alpha}+\ldots.+n^{2\alpha}=\int_0^n i^{2\alpha}di=\left.\frac{i^{2\alpha+1}}{2\alpha+1}\right| with$ limits 0 to n

$Bn=\frac{n^{2\alpha+1}}{(2\alpha+1)}$ (From Euler Maclaurin's formula)

$$\lim_{n\to\infty}\frac{B_n}{n^2}=\lim_{n\to\infty}\frac{n^{2\alpha+1}}{(2\alpha+1)n^2}=\lim_{n\to\infty}\frac{n^{2\alpha+1-2}}{(2\alpha+1)}=\lim_{n\to\infty}\frac{n^{2\alpha-1}}{(2\alpha+1)}$$

For WLLN to hold $\lim_{n\to\infty}\frac{n^{2\alpha-1}}{(2\alpha+1)}\to 0$ as $n\to\infty$

For which $2\alpha-1<0$ which implies $2\alpha<1$ implies $\alpha<1/2$

Hence the result

5. The mean expenditure per customer at a tire store is \$85.00, with a standard deviation of \$9.00. If a random sample of 40 customers is taken, what is the probability that the sample average expenditure per customer for this sample will be \$87.00 or more?

Answer:

Because the sample size is greater than 30, the central limit theorem says the sample means are normally distributed.

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

Given the mean expenditure per customer at a tire store is $85.00, μ=$85

Standard deviation of $9.00, σ=$9.00 and n=40. Then

$$Z = \frac{\$87 - \$85}{\frac{\$9}{\sqrt{40}}}$$

Z = $2.00 / $1.42 = 1.41

For Z = 1.41 in the standard normal distribution table, the probability is .4207.

This represents the probability of getting a mean between $87.00 and the population mean $85.00.

But we need P(X>$87) = P (Z>1.41) = 0.0793

This is the probability of average expenditure ≥ $87.00.

Therefore, 7.93% of the time, a random sample of 40 customers from this population will yield a mean expenditure of $87.00 or more.

OR

From any random sample of 40 customers, 7.93% of them will spend on average $87.00 or more.

6. Suppose that during any hour in a large department store, the average number of shoppers is 448, with a standard deviation of 21 shoppers. What is the probability that a random sample of 49 different shopping hours will yield a sample mean between 441 and 446 shoppers?

Answer:

Given the average number of shoppers is 448 i.e., μ=448

Standard deviation is of 21, σ=21 and n=49. Then

We have to find $P[441 < \overline{x} < 446]$

$$\Rightarrow P[\frac{441-\mu}{\frac{\sigma}{\sqrt{n}}} < Z < \frac{446-\mu}{\frac{\sigma}{\sqrt{n}}}] = P[\frac{441-448}{\frac{21}{\sqrt{49}}} < Z < \frac{446-448}{\frac{21}{\sqrt{49}}}]$$

$$\Rightarrow P[\frac{-7}{3} < Z < \frac{-2}{3}] = P[-2.33 < Z < -0.67] = 0.152399$$

From the standard normal distribution table, the probability is 0.152399

Therefore, 15% of the time, a random sample of 49 different shopping hours will yield a sample mean between 441 and 446 shoppers

OR

From any random sample of 49 hours, 15% of them will have an average number of shoppers between 441 and 446

7. 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… 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 $P(-2 \leq Z \leq 2) = 0.9545$

So the probability that Jim's records are off by at least $10 is 95%.

8. From a population with Standard deviation 2 a sample of size 50 is drawn. What is the probability that the sample mean will differ from the population mean by not more than 2?

Answer:

Let x1, x2... xn are a random sample independently drawn. Let mean of the population be μ and σ=2

Let x1+x2+....+xn=X

Since the sample is large, by Central limit theorem

$$\frac{\frac{X}{n}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

To find $P[|X-\mu| \le -2]$

$$= P[\left|\frac{\frac{X}{n}-\mu}{\sigma/\sqrt{n}}\right| < \frac{-2}{\sigma/\sqrt{n}}]$$

$$= P[|Z| < \frac{-2}{2/\sqrt{50}}] = P[|Z| < -0.71]$$

$\Rightarrow P[-0.71 < Z < 0.71] = \int_{-0.71}^{0.71} \phi(z)dz = 2 * 0.2611 = 0.5222$ (From the standard normal distribution table)

9. A distribution with an unknown mean μ has variance =1.5. By using Central Limit Theorem find how large a sample should be taken in order that the probability will be at least 0.95.that the sample mean will be within 0.5 of the population mean

Answer:

It is given that the probability will be at least 0.95.that the sample mean will be within 0.5 of the population mean, then

$$P[|\bar{x}-\mu| < 0.5] \ge 0.95$$

From the standard Normal distribution table $P[|Z| < 1.96] = 0.95$

Therefore

$$P\left[\left|\frac{\bar{x}-\mu}{\sqrt{1.5}/\sqrt{n}}\right| < 1.96\right] = 0.95 \Rightarrow P\left[\left|\bar{x}-\mu\right| < \frac{1.96\sqrt{1.5}}{\sqrt{n}}\right] = 0.95$$

But as per question

$$P[\left|\bar{x}-\mu\right| < 0.5] \geq 0.95$$

Therefore

$$\frac{1.96\sqrt{1.5}}{\sqrt{n}} \leq 0.5 \Rightarrow \frac{(1.96)^2(1.5)}{(0.5)^2} \leq n \Rightarrow n \geq 24$$

Therefore sample size should be at least 24

10. A pooling agency has to take a sample of voters in a given state large enough that the probability is only 0.01 that they will find the proportion favouring to certain candidate to be less than 50% when in fact it is 52%. How large a sample should be taken?

Answer:

The proportion of voters favouring to certain candidate is 52%=p=0.52

Hence μ=0.52, σ^2 =pq/n= 0.52*0.48/n=0.2496/n

Then we know that

$$\frac{p - E(p)}{S.D(p)} \sim N(0,1)$$

Therefore $\dfrac{p-0.52}{\sqrt{\dfrac{0.2496}{n}}} \sim N(0,1)$

But as per question the probability is only 0.01 that they will find the proportion of voters favouring to certain candidate to be less than 50%. Then

$$P[p < 0.50] = 0.01$$

P $\left[\dfrac{p-E(p)}{S.D(p)} < \dfrac{0.50-0.52}{\sqrt{\dfrac{0.2496}{n}}}\right] = 0.01$

P $[Z < \frac{p-0.52}{\sqrt{\frac{0.2496}{n}}}] = 0.01$. Then from the table of standard normal probabilities Z=-2.33

Then

$$\frac{0.50-0.52}{\sqrt{\frac{0.2496}{n}}} = -2.33 \Rightarrow -0.02 = -2.33\sqrt{\frac{0.2496}{n}} \Rightarrow \frac{(0.2496)(2.33)^2}{(-0.02)^2} = n$$

n=3388. Therefore sample size should be 3388

11. Let {Xn} be a sequence of independent random variables with $P[Xi = \pm i] = \frac{1}{2}$. Examine whether the WLLN holds for the sequence

Answer:

For the given problem we have P [Xi=i] =P (xi=-i] =1/2

E (Xi) = ∑xi p(x i) = i*1/2+ (-i)*1/2=0

V (xi) =E $(Xi^{2)}$-[E (Xi)] ^{2}

E (Xi^2) = ∑xi^2 p(x i) = i^2*1/2+ (-i) ^{2}*1/2=i^2

V (xi) =E $(Xi^{2)}$-[E (Xi)] ^{2} = i^2

Bn=V(X1+X2+…+Xn) =1^2+2^2 +…. +n^2 =n (n+1) (2n+1)/6

$$\lim_{n\to\infty} \frac{B_n}{n^2} = \lim_{n\to\infty} \frac{n(n+1)(2n+1)}{6n^2} \to \infty \text{ As } n \to \infty$$

The sufficient condition of WLLN is not satisfied by the given random variables Hence we cannot draw a conclusion whether WLLN holds or not.

12. Examine whether the WLLN is applicable to the following sequence of uncorrelated random variables

Answer:

$$P[Xk = \pm\sqrt[3]{k}] = \frac{1}{2}$$

E (Xk) = ∑x p (xk) = $k^{1/3} * 1/2 + (-k)^{1/3} * 1/2 = 0$

V (xk) =E (Xk$^{2)}$-[E (Xk)]^{2}

E (Xk^2) = ∑x^2 p(x) = $k^{2/3} * 1/2 + (-k)^{2/3} * 1/2 = k^{2/3}$

V (xi) =E (Xi$^{2)}$-[E (Xi)]^{2} = i$^{2\alpha}$

Bn=V(X1+X2+…+Xn) = $1^{2/3} + 2^{2/3} + ... + n^{2/3}$ = $\int_0^n k^{2/3} dk = \left.\frac{k^{(2/3)+1}}{(2/3)+1}\right| with$

limits 0 to n

Bn= $\frac{n^{(2/3)+1}}{(2/3)+1} = \frac{3}{5} n^{\frac{5}{3}}$ (From Euler Maclaurin's formula)

For WLLN to hold $\lim_{n \to \infty} \frac{B_n}{n^2}$ =0

$$\lim_{n \to \infty} \frac{B_n}{n^2} = \lim_{n \to \infty} \frac{3}{5} \frac{n^{\frac{5}{3}}}{n^2} = \lim_{n \to \infty} \frac{3}{5} n^{-1/3} \to 0 \text{ as } n \to \infty$$

Therefore WLLN holds for the given sequence of random variables

13.

In a game involving repeated throws of a balanced die, a person receives Rs. 3 if the resulting number is greater than or equal to three and loses Rs 3 otherwise. Use CLT to find the probability that in 25 trials his total earnings exceeds Rs 25

Answer:

Let Xi be the earnings in the ith game . Then

x	Probability	Xp	X^2 p
3	2/3	2	6
-3	1/3	-1	3

E(Xi)= ∑Xi p =1

E(Xi^2)= ∑Xi^2 p =9

$V(Xi) = E(Xi^2) - [E(Xi)]^2 = 8$

Let X=X1+X2+...+X25

Then E(X) = n E(Xi) = 25X1= 25

V(X) = V(∑Xi) = ∑V(Xi)= 25 X 8= 120

By the CLT

$Z = \frac{X - 25}{\sqrt{120}} \sim N(0,1)$ Asymptotically

Consider $P(X > 25) = P[\frac{X - 25}{\sqrt{120}} > 0] = P[Z > 0] = 0.5$

From the table of standard Normal probabilities

Therefore the probability that the total earnings exceeds Rs.25 is equal to 0.5

14. An insurance company has 10,000 automobile policy holders. If the expected yearly claim per policy holder is $260 with an Standard deviation of $800 , approximate the probability that the total yearly claim exceeded $2.8 million

Answer:

Let Xi denote the yearly claim of policy holder i, i=1, 2... 10,000

By CLT X= ∑Xi will have approximately Normal distribution with mean 10,000*260 = $2.6 * 10^6$ and Standard deviation 800 $\sqrt{10^4} = 8*10^4$

Hence $P[X>2.8 * 10^{6}] =$

$P [(X-2.6 * 10^6)/ 8*10^4 > (2.8 * 10^6 -2.6 * 10^6)/ 8*10^4$

=P [Z>20/8] = P [Z>2.5] =0.0062

That is there are only 6 chances out of 1000 that the total yearly claim will exceed $2.8 million

15. Let {Xn} be a sequence of independent random variables with Xi having an Exponential distribution with mean $2^{-i/2}$. Examine whether the WLLN holds for the sequence {Xn}

Answer:

Given Xi ~ Exponential with mean $2^{-i/2}$

Given E (Xi) = $2^{-i/2}$

Therefore V (Xi) = $(2^{-i/2})^2 = 2^{-i}$

Because when Xi~Exponential with mean θ the variance of Xi is θ^2

Bn=V(X1+X2+…+Xn) = $2^{-1} + 2^{-2} + 2^{-3} + \ldots + 2^{-n} = 1 - (1/2^n)$

$$\lim_{n\to\infty} \frac{B_n}{n^2} = \lim_{n\to\infty} \frac{1-(\frac{1}{2^n})}{n^2} \to 0 \text{ as } n \to \infty$$

Therefore WLLN holds for the given sequence of random variables.