Week 7 Graded

Practice questions

Question 1

Let X1 and X2 are two i.i.d. Random variables with mean μ and variance σ^2 Given below are three different formulations of sample mean:

$$A = \frac{X_1 + X_2}{2}$$

$$B = 0.1X_1 + 0.9X_2$$

$$C = X_2$$

Choose the correct options from the following:

1. Var(A) < Var(B) < Var(C)
2. Var(A) = Var(B) = Var(C)
3. Var(A) > Var(B) > Var(C)
4. Var(C) > Var(A) > Var(B)

Answer:

Q1. $X_1, X_2 \sim \text{i.i.d } X$

$E[X] = \mu, \quad Var(X) = \sigma^2$

Given: $A = \dfrac{X_1 + X_2}{2}$

$B = 0.1\, X_1 + 0.9\, X_2$

$C = X_2$

$$Var(A) = Var\left(\frac{X_1 + X_2}{2}\right)$$
$$= \frac{1}{2^2} Var(X_1 + X_2)$$
$$= \frac{1}{4}\left[Var(X_1) + Var(X_2)\right]$$
$$= \frac{1}{4}\left[\sigma^2 + \sigma^2\right] = \frac{\sigma^2}{2}$$

$$Var(B) = Var(0.1\, X_1 + 0.9\, X_2)$$
$$= (0.1)^2\, Var(X_1) + (0.9)^2\, Var(X_2)$$
$$= (0.01 + 0.81)\, \sigma^2$$
$$= 0.82\, \sigma^2$$

$$Var(C) = Var(X_2)$$
$$= \sigma^2$$

$$\therefore \boxed{Var(A) < Var(B) < Var(C)}$$

Question 2

Let $X_1, \ldots, X_{50}$ ~ Exp(0.4). Let Y = $\sum_{i=1}^{50} X_i$. Use CLT to approximate P(Y > 150).

Use: F_Z(1.41) = 0.92073

Answer:

Q2. $X_1, X_2, \ldots, X_{50} \sim$ i.i.d $Exp(0.4)$

$$Y = \sum_{i=1}^{50} X_i$$

To find: $P(Y > 150)$

$X_i \sim Exp(0.4)$

$$E[X_i] = \frac{1}{0.4} = 2.5$$

$$Var[X_i] = \frac{1}{(0.4)^2} = 6.25$$

Using CLT, $\dfrac{Y - n\mu}{\sqrt{n}\,\sigma} \sim N(0,1)$

$$\Rightarrow \frac{Y - 50(2.5)}{\sqrt{50 \times 6.25}} = \frac{Y - 125}{17.67} \mathrel{\dot\sim} N(0,1)$$

$$P(Y > 150) = P\left(\frac{Y-125}{17.67} > \frac{150-125}{17.67}\right)$$

$$= P(Z > 1.41)$$

$$= 1 - F_Z(1.41)$$

$$= 1 - 0.92073$$

$$= \boxed{0.07927}$$

Question 3

Let $X_1, \ldots, X_{100}$~ i.i.d. N(0, 2). Evaluate: $P\left(X_1^2 + \ldots + X_{100}^2 > 220\right)$ approximately using CLT.

Use F_Z(0.7071) = 0.76115

Answer:

Q3. $X_1, \ldots, X_{100} \sim$ i.i.d. $N(0, 2)$

To find: $P(X_1^2 + \ldots + X_{100}^2 > 220)$

$X_i \sim N(0,2)$, $\forall i$

$X_i^2 \sim \text{Gamma}\left(\frac{1}{2}, \frac{1}{2 \cdot 2}\right)$, $\mu = E[X_i^2] = \frac{1}{2} \times 4 = 2$, $\sigma^2 = Var(X_i^2) = \frac{1}{2} \times 16 = 8$

$X_1^2 + \ldots + X_{100}^2 \sim \text{Gamma}\left(\frac{100}{2}, \frac{1}{4}\right)$

$\therefore P(X_1^2 + \ldots + X_{100}^2 > 220)$

$$= P\left(\frac{X_1^2 + \ldots + X_{100}^2 - 100\mu}{\sigma\sqrt{100}} > \frac{220 - 100\mu}{\sigma\sqrt{100}}\right)$$

$$= P\left(\frac{X_1^2 + \ldots + X_{100}^2 - 200}{20\sqrt{2}} > \frac{220 - 200}{20\sqrt{2}}\right) = P\left(Z > \frac{20}{20\sqrt{2}}\right) = P\left(Z > \frac{1}{\sqrt{2}}\right)$$

$$= 1 - F_Z(1/\sqrt{2})$$

$$= 1 - F_Z(0.7071)$$

$$= 1 - 0.76115 = \boxed{0.23885}$$

Question 4

Let X be a random variable having the gamma distribution with parameters α = n and β = 1.

Hint:

1. If X ~ Gamma(α , β),
   $$E[X] = \frac{\alpha}{\beta},\ Var[X] = \frac{\alpha}{\beta^2}$$

1. Use WLLN.

Find the value of n such that

$$P\left(\left|\frac{X}{n} - 1\right| > 0.01\right) < 0.01$$

Answer:

Q4. $X \sim \text{Gamma}(n, 1)$

To find n: $P\left(\left|\frac{X}{n} - 1\right| > 0.01\right) < 0.01$

Let $X = X_1 + \dots + X_n$, where $X_i \sim \text{Gamma}(1, 1)$

Recall: $X_1 + \dots + X_n \sim \text{Gamma}(n\alpha, \beta)$

Now, $\mu = E[X_i] = \frac{\alpha}{\beta} = 1$

$\sigma^2 = \text{Var}(X_i) = \frac{\alpha}{\beta^2} = 1$

WLLN: $P\left(\left|\frac{X}{n} - 1\right| > 0.01\right) < 0.01$

$\left\{ E\left(\frac{X_1 + \dots + X_n}{n}\right) = 1 \right\}$

$\Rightarrow P\left(\left|\frac{X_1 + \dots + X_n}{n} - 1\right| > 0.01\right) < \frac{\sigma^2}{n(0.01)^2}$

$\therefore \frac{1}{n(0.01)^2} < 0.01 \Rightarrow \boxed{n > 10^6}$

Question 5

Let X be a random variable having the gamma distribution with parameters α = n and β = 1.

Hint:

1. If X ~ Gamma(α, β),

$$E[X] = \frac{\alpha}{\beta}, Var[X] = \frac{\alpha}{\beta^2}$$

1. Use CLT.
2. Use F_Z(2.58) = 0.995

Find the value of n such that

$$P\left(\left|\frac{X}{n} - 1\right| > 0.01\right) < 0.01$$

1. 66560
2. 66565
3. 66575
4. 66500

Q5. Using CLT :

$$P\left(\left|\frac{X}{n} - n\right| > 0.01\right) < 0.01$$

$$\Rightarrow P\left(\left|\frac{X_1 + \dots + X_n}{n} - 1\right| > 0.01\right) < 0.01$$

$$\Rightarrow P\left(\left|\frac{X_1 + \dots + X_n - n}{\sqrt{n}}\right| > 0.01\sqrt{n}\right) < 0.01$$

$$\Rightarrow P(|Z| > 0.01\sqrt{n}) < 0.01$$

$$\Rightarrow 2P(Z > 0.01\sqrt{n}) < 0.01$$

$$\Rightarrow 1 - F_Z(0.01\sqrt{n}) < \frac{0.01}{2}$$

$$\Rightarrow F_Z(0.01\sqrt{n}) > 0.995$$

$$\Rightarrow F_Z(0.01\sqrt{n}) > F_Z(2.58) \Rightarrow 0.01\sqrt{n} > 2.58 \Rightarrow n > 258^2 \Rightarrow \boxed{n > 66564}$$

Question 6

Suppose speed of vehicles on a particular road are normally distributed with mean 25 mph and standard deviation 3 mph.

Find the probability that the mean speed $\overline{X}$ of 30 randomly selected vehicles is between 24 and 27 mph.

Write your answer in terms of F_Z

Answer:

Q6: Let X denote the speed of vehicles on particular road.

$n = 30,\ \mu = 25,\ \sigma = 3$

$X_1, \dots, X_{30} \sim$ i.i.d. X

To find: $P(24 < \bar{X} < 27)$

$$= P\left(24 < \frac{X_1 + \dots + X_{30}}{30} < 27\right)$$

$$= P\left(24 < \frac{Y}{30} < 27\right)$$

$$= P\left(24 - 25 < \frac{Y}{30} - 25 < 27 - 25\right)$$

$$= P\left(-1 < \frac{Y - 25(30)}{30} < 2\right)$$

$$= P\left(-\frac{\sqrt{30}}{3} < \frac{Y - 25(30)}{3\sqrt{30}} < \frac{2\sqrt{30}}{3}\right) = P\left(-\frac{\sqrt{30}}{3} < Z < \frac{4\sqrt{5}}{3}\right) = \boxed{F_Z(3.65) - F_Z(-1.82)}$$