

Solution to GATE ST 2023.40

Devansh Jain - EE22BTECH11018

Question: Let X_1, X_2, \dots, X_{10} be a random sample of size 10 from a population having $\mathcal{N}(0, \theta^2)$ distribution, where $\theta > 0$ is an unknown parameter. Let $T = \frac{1}{10} \sum_{i=1}^{10} X_i^2$. If the mean square error of cT ($c > 0$) as an estimator of θ^2 , is minimized at $c = c_0$, then the value of c_0 equals

- (i) $\frac{5}{6}$
- (ii) $\frac{2}{3}$
- (iii) $\frac{3}{5}$
- (iv) $\frac{1}{2}$

(GATE ST 2023)

Solution: The mean of T is given by,

$$E(T) = E\left(\frac{1}{10} \sum_{i=1}^{10} X_i^2\right) \quad (1)$$

$$= \frac{1}{10} \sum_{i=1}^{10} E(X_i^2) \quad (2)$$

Since

$$E(X^2) = V(X) + (E(X))^2 \quad (3)$$

Using (3)

$$E(T) = \frac{1}{10} \sum_{i=1}^{10} V(X_i) + E(X_i)^2 \quad (4)$$

$$= \frac{1}{10} (10\theta^2) \quad (5)$$

$$= \theta^2 \quad (6)$$

Similarly, using (3), the variance of T is given by,

$$V(T) = E(T^2) - E(T)^2 \quad (7)$$

$$= \frac{1}{100} \left(E\left(\left(\sum_{i=1}^{10} X_i^2\right)^2\right) - \left(E\left(\sum_{i=1}^{10} X_i^2\right)\right)^2 \right) \quad (8)$$

But

$$E\left(\left(\sum_{i=1}^{10} X_i^2\right)^2\right) = E\left(\sum_{i=1}^{10} \sum_{j=1}^{10} X_i^2 X_j^2\right) \quad (9)$$

$$= \sum_{i=1}^n \sum_{j=1}^n E(X_i^2 X_j^2) \quad (10)$$

and

$$\left(E \left(\sum_{i=1}^{10} X_i \right) \right)^2 = \left(\sum_{i=1}^{10} E(X_i) \right)^2 \quad (11)$$

$$= \sum_{i=1}^{10} \sum_{j=1}^{10} E(X_i) E(X_j) \quad (12)$$

Using (10) , (12), and the definition of covariance,

$$V(T) = \frac{1}{100} \left(\sum_{i=1}^{10} \sum_{j=1}^{10} (E[X_i^2 X_j^2] - E(X_i^2) E(X_j^2)) \right) \quad (13)$$

$$= \frac{1}{100} \left(\sum_{i=1}^{10} \sum_{j=1}^{10} \text{cov}(X_i^2, X_j^2) \right) \quad (14)$$

As all the variables are i.i.d's and are thus uncorrelated,

$$\text{cov}(X_i^2, X_j^2) = \begin{cases} 0 & \text{if } i \neq j \\ V(X_i^2) & \text{if } i = j \end{cases} \quad (15)$$

Now,

$$V(T) = \frac{1}{100} \sum_{i=1}^{10} \text{cov}(X_i^2, X_i^2) \quad (16)$$

$$= \frac{1}{100} \sum_{i=1}^{10} V(X_i^2) \quad (17)$$

Using (3) to find $V(X_i^2)$,

$$V(X_i^2) = E(X_i^4) - E(X_i^2)^2 \quad (18)$$

Using moment generating function to find $E(X_i^4)$ and $E(X_i^2)$,

$$M_X(t) = E(e^{tX}) \quad (19)$$

$$= e^{\frac{1}{2}\theta^2 t^2} \quad (20)$$

Differentiating it with respect to t ,

$$\frac{dM_X(t)}{dt} = E(Xe^{tX}) \quad (21)$$

$$\frac{d^n M_X(t)}{dt^n} = E(X^n e^{tX}) \quad (22)$$

$$\frac{dM_X(t)}{dt} = \theta^2 t e^{\frac{1}{2}\theta^2 t^2} \quad (23)$$

$$\frac{d^2 M_X(t)}{dt^2} = \theta^2 e^{\frac{1}{2}\theta^2 t^2} + \theta^4 t^2 e^{\frac{1}{2}\theta^2 t^2} \quad (24)$$

$$\frac{d^3 M_X(t)}{dt^3} = \theta^4 t e^{\frac{1}{2}\theta^2 t^2} + 2\theta^4 t e^{\frac{1}{2}\theta^2 t^2} + \theta^6 t^3 e^{\frac{1}{2}\theta^2 t^2} \quad (25)$$

$$= 3\theta^4 t e^{\frac{1}{2}\theta^2 t^2} + \theta^6 t^3 e^{\frac{1}{2}\theta^2 t^2} \quad (26)$$

$$\frac{d^4 M_X(t)}{dt^4} = 3\theta^4 e^{\frac{1}{2}\theta^2 t^2} + 3\theta^6 t^2 e^{\frac{1}{2}\theta^2 t^2} + 3\theta^6 t^2 e^{\frac{1}{2}\theta^2 t^2} + \theta^8 t^4 e^{\frac{1}{2}\theta^2 t^2} \quad (27)$$

$$= 3\theta^4 e^{\frac{1}{2}\theta^2 t^2} + 6\theta^6 t^2 e^{\frac{1}{2}\theta^2 t^2} + \theta^8 t^4 e^{\frac{1}{2}\theta^2 t^2} \quad (28)$$

Now,

$$E(X^4) = \frac{d^4 M_X(t)}{dt^4} \Big|_{t=0} \quad (29)$$

$$= 3\theta^4 \quad (30)$$

$$E(X^2) = \frac{d^2 M_X(t)}{dt^2} \Big|_{t=0} \quad (31)$$

$$= \theta^2 \quad (32)$$

$$V(X_i^2) = 3\theta^4 - (\theta^2)^2 \quad (33)$$

$$= 2\theta^4 \quad (34)$$

$$V(T) = \frac{1}{100} \sum_{i=1}^{10} 2\theta^4 \quad (35)$$

$$= \frac{1}{5} \theta^4 \quad (36)$$

The mean square error of cT as an estimator of θ^2 is given by

$$f = E\left((cT - \theta^2)^2\right) \quad (37)$$

Since

$$E(X^2) = V(X) + (E(X))^2 \quad (38)$$

Now,

$$f = V(cT - \theta^2) + \left(E(cT - \theta^2)\right)^2 \quad (39)$$

$$= V(cT) + \left(E(cT) - \theta^2\right)^2 \quad (40)$$

$$= c^2 V(T) + \left(cE(T) - \theta^2\right)^2 \quad (41)$$

$$= \frac{c^2}{5} \theta^4 + \theta^4 (c - 1)^2 \quad (42)$$

Minimizing the mean square error,

$$\frac{df}{dc} = 0 \quad (43)$$

$$\frac{df}{dc} = \frac{2c}{5} \theta^4 + 2\theta^4 (c - 1) \quad (44)$$

$$\frac{c}{5} + c - 1 = 0 \quad (45)$$

$$6c = 5 \quad (46)$$

$$c = \frac{5}{6} \quad (47)$$

To verify that cT is a good estimate for θ , let

$$\theta = 0.5 \quad (48)$$

$$E\left((cT - \theta^2)^2\right) = \frac{1}{5} \left(\frac{25}{36}\right) (0.5)^4 + \left(\frac{-1}{6}\right)^2 (0.5)^4 \quad (49)$$

$$= \frac{1}{96} \quad (50)$$

$$= 0.0104 \quad (51)$$

The mean square error is small relative to the value of θ , hence cT is a good estimate.

Simulation procedure:

(i)

$$u_1 = (\text{double}) \frac{\text{rand}()}{\text{RAND_MAX}} \quad (52)$$

$$u_2 = (\text{double}) \frac{\text{rand}()}{\text{RAND_MAX}} \quad (53)$$

Generates a uniform distribution between 0 and 1.

(ii)

$$X_i = \sqrt{\theta^2} \left(\sqrt{-2 \log u_1} \cos 2\pi u_2 \right) + \mu \quad (54)$$

Transforms the uniform distribution into gaussian distribution of desired mean and variance. Ten such random variables are generated.

(iii)

$$T = \frac{1}{10} \sum_{i=1}^{10} X_i^2 \quad (55)$$

The values of the random variables are squared and then averaged together to generate T .

(iv) The value of c which minimizes the mean square error is found by calculating $E\left((cT - \theta^2)^2\right)$ for a range of values of c and choosing that value of c which gives the minimum value for the expression.

