# Solution to GATE ST 2023.40

Devansh Jain - EE22BTECH11018

Question: Let $X_1, X_2, ..., X_{10}$ be a random sample of size 10 from a population having $\mathcal{N}\left(0, \theta^2\right)$ 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 $\theta^2$, 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,

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

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

Since

$$E\left(X^2\right) = V(X) + (E(X))^2 \tag{3}$$

Using (3)

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

$$= \frac{1}{10}\left(10\theta^2\right) \tag{5}$$

$$= \theta^2 \tag{6}$$

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

$$V(T) = E\left(T^2\right) - E(T)^2 \tag{7}$$

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

But

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

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{E}\left(X_i^2 X_j^2\right) \tag{10}$$

and

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

$$= \sum_{i=1}^{10} \sum_{j=1}^{10} \mathrm{E}\left(X_i\right) \mathrm{E}\left(X_j\right) \tag{12}$$

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

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

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

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

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

Now,

$$V\left(T\right) = \frac{1}{100} \sum_{i=1}^{10} \mathrm{cov}\left(X_i^2, X_i^2\right) \tag{16}$$

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

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

$$V\left(X_i^2\right) = E\left(X_i^4\right) - E\left(X_i^2\right)^2 \tag{18}$$

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

$$M_X\left(t\right) = E\left(e^{tX}\right) \tag{19}$$

$$= e^{\frac{1}{2}\theta^2 t^2} \tag{20}$$

Differentiating it with respect to $t$,

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

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

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

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

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

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

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

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

Now,

$$E\left(X^4\right) = \frac{d^4 M_X(t)}{dt^4}|_{t=0} \tag{29}$$
$$= 3\theta^4 \tag{30}$$
$$E\left(X^2\right) = \frac{d^2 M_X(t)}{dt^2}|_{t=0} \tag{31}$$
$$= \theta^2 \tag{32}$$

$$V\left(X_i^2\right) = 3\theta^4 - \left(\theta^2\right)^2 \tag{33}$$
$$= 2\theta^4 \tag{34}$$
$$V(T) = \frac{1}{100}\sum_{i=1}^{10} 2\theta^4 \tag{35}$$
$$= \frac{1}{5}\theta^4 \tag{36}$$

The mean square error of $cT$ as an estimator of $\theta^2$ is given by

$$f = E\left(\left(cT - \theta^2\right)^2\right) \tag{37}$$

Since

$$E\left(X^2\right) = V(X) + (E(X))^2 \tag{38}$$

Now,

$$f = V\left(cT - \theta^2\right) + \left(E\left(cT - \theta^2\right)^2\right) \tag{39}$$
$$= V(cT) + \left(E(cT) - \theta^2\right)^2 \tag{40}$$
$$= c^2 V(T) + \left(cE(T) - \theta^2\right)^2 \tag{41}$$
$$= \frac{c^2}{5}\theta^4 + \theta^4 (c-1)^2 \tag{42}$$

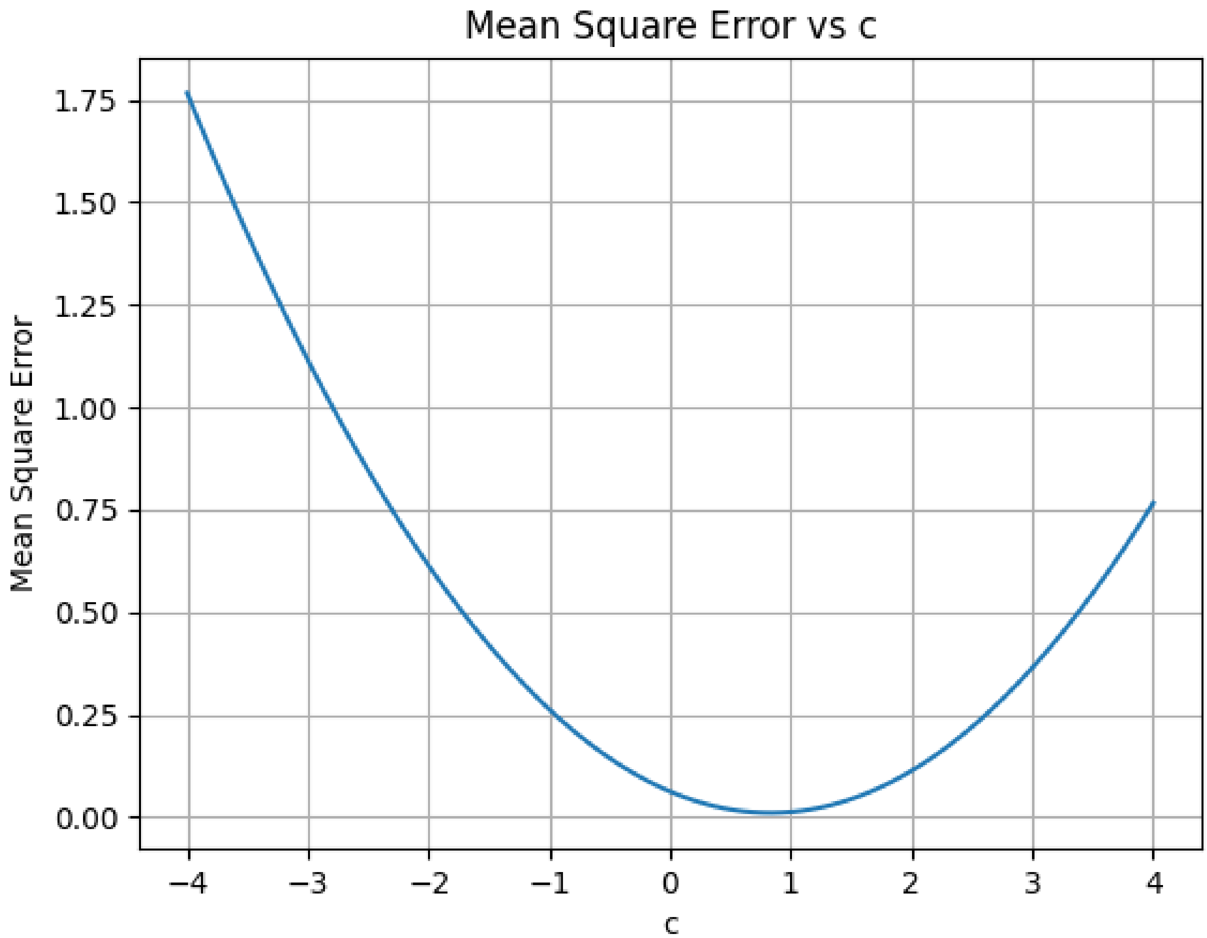
Minimizing the mean square error,

$$\frac{df}{dc} = 0 \tag{43}$$
$$\frac{df}{dc} = \frac{2c}{5}\theta^4 + 2\theta^4 (c-1) \tag{44}$$
$$\frac{c}{5} + c - 1 = 0 \tag{45}$$
$$6c = 5 \tag{46}$$
$$c = \frac{5}{6} \tag{47}$$

To verify that $cT$ is a good estimate for $\theta$, let

$$\theta = 0.5 \tag{48}$$
$$E\left(\left(cT - \theta^2\right)^2\right) = \frac{1}{5}\left(\frac{25}{36}\right)(0.5)^4 + \left(\frac{-1}{6}\right)^2 (0.5)^4 \tag{49}$$
$$= \frac{1}{96} \tag{50}$$
$$= 0.0104 \tag{51}$$

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




Using laplace transform to find CDF,

$$M_T(s) = M_{\frac{1}{10}\sum_{i=1}^{10} X_i^2}(s) \tag{52}$$

$$= \prod_{i=1}^{10} M_{\frac{X_i^2}{10}}(s) \tag{53}$$

Now,

$$M_{\frac{X^2}{10}}(s) = E\left(e^{\frac{-sX^2}{10}}\right) \tag{54}$$

$$= \int_{-\infty}^{\infty} e^{\frac{-sx^2}{10}} \frac{e^{\frac{-x^2}{2\theta^2}}}{\theta\sqrt{2\pi}} dx \tag{55}$$

$$= \frac{1}{\theta\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-x^2}{2}\left(\frac{1}{\theta^2}+\frac{s}{5}\right)} dx \tag{56}$$

Using the substitution,

$$u = x\sqrt{\frac{1}{\theta^2} + \frac{s}{5}} \tag{57}$$

$$du = \left(\sqrt{\frac{1}{\theta^2} + \frac{s}{5}}\right)dx \tag{58}$$

$$M_{\frac{X^2}{10}}(s) = \frac{1}{\theta\sqrt{\frac{1}{\theta^2} + \frac{s}{5}}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{\frac{-u^2}{2}}du \tag{59}$$

The integral evaluates to 1 since it is the pdf of a normal distribution,

$$M_{\frac{X^2}{10}}(s) = \frac{1}{\sqrt{1 + \frac{s\theta^2}{5}}} \tag{60}$$

$$M_T(s) = \frac{1}{\left(1 + \frac{s\theta^2}{5}\right)^5} \tag{61}$$

Taking inverse laplace transform gives us the pdf,

$$p_T(t) = L^{-1}\left[M_T(s)\right] \tag{62}$$

$$= L^{-1}\left[\frac{1}{\left(1 + \frac{s\theta^2}{5}\right)^5}\right] \tag{63}$$

$$= \frac{3125t^4e^{\frac{-5t}{\theta^2}}}{24\theta^{10}}u(t) \tag{64}$$

ROC of laplace transform:

$$Re(s) > \frac{-5}{\theta^2} \tag{65}$$

CDF of T:

$$F_T(t) = \int_{-\infty}^{t} p_T(t)\,dt \tag{66}$$

$$= \int_{-\infty}^{t}\frac{3125t^4e^{\frac{-5t}{\theta^2}}}{24\theta^{10}}u(t)\,dt \tag{67}$$

$$= \int_{0}^{t}\frac{3125t^4e^{\frac{-5t}{\theta^2}}}{24\theta^{10}}dt \tag{68}$$

Applying integration by parts to solve the integral,

$$F_T(t) = \frac{3125}{24\theta^{10}}\left(\frac{-\theta^2}{5}t^4e^{\frac{-5t}{\theta^2}} + \int_0^t \frac{4\theta^2}{5}t^3e^{\frac{-5t}{\theta^2}}dt\right) \tag{69}$$

$$= \frac{3125}{24\theta^{10}}\left(-e^{\frac{-5t}{\theta^2}}\left(\frac{\theta^2}{5}t^4 + \frac{4\theta^4}{25}t^3\right) + \int_0^t \frac{12\theta^4}{25}t^2e^{\frac{-5t}{\theta^2}}dt\right) \tag{70}$$

$$= \frac{3125}{24\theta^{10}}\left(-e^{\frac{-5t}{\theta^2}}\left(\frac{\theta^2}{5}t^4 + \frac{4\theta^4}{25}t^3 + \frac{12\theta^6}{125}t^2\right) + \int_0^t \frac{24\theta^6}{125}te^{\frac{-5t}{\theta^2}}dt\right) \tag{71}$$

$$= \frac{3125}{24\theta^{10}}\left(-e^{\frac{-5t}{\theta^2}}\left(\frac{\theta^2}{5}t^4 + \frac{4\theta^4}{25}t^3 + \frac{12\theta^6}{125}t^2 + \frac{24\theta^8}{625}t\right) + \int_0^t \frac{24\theta^8}{625}e^{\frac{-5t}{\theta^2}}dt\right) \tag{72}$$

$$= \frac{3125}{24\theta^{10}}\left(-e^{\frac{-5t}{\theta^2}}\left(\frac{\theta^2}{5}t^4 + \frac{4\theta^4}{25}t^3 + \frac{12\theta^6}{125}t^2 + \frac{24\theta^8}{625}t + \frac{24\theta^{10}}{3125}\right) + \frac{24\theta^{10}}{3125}\right) \tag{73}$$

$$= 1 - \frac{e^{\frac{-5t}{\theta^2}}}{24\theta^8}\left(625t^4 + 500\theta^2t^3 + 300\theta^4t^2 + 120\theta^6t + 24\theta^8\right) \tag{74}$$

Let $\theta = 0.5$ for simulation,

$$F_T(t) = 1 - \frac{e^{-20t}}{3}\left(20000t^4 + 4000t^3 + 600t^2 + 60t + 3\right) \tag{75}$$

Simulation procedure:

(i)

$$u_1 = (double)\frac{rand()}{RAND\_MAX} \tag{76}$$

$$u_2 = (double)\frac{rand()}{RAND\_MAX} \tag{77}$$

Generates a uniform distribution between 0 and 1.

(ii)

$$X_i = \sqrt{\theta^2}\left(\sqrt{-2\log u_1}\cos 2\pi u_2\right) + \mu \tag{78}$$

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

(iii)

$$T = \frac{1}{10}\sum_{i=1}^{10} X_i^2 \tag{79}$$

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 $E\left(\left(cT - \theta^2\right)^2\right)$ for a range of values of $c$ and choosing that value of $c$ which gives the minimum value for the expression.

(v) The cdf is simulated by counting the number of samples less than a certain $t$ and dividing it by the total number of samples. This gives the CDF of $T$ at $t$.

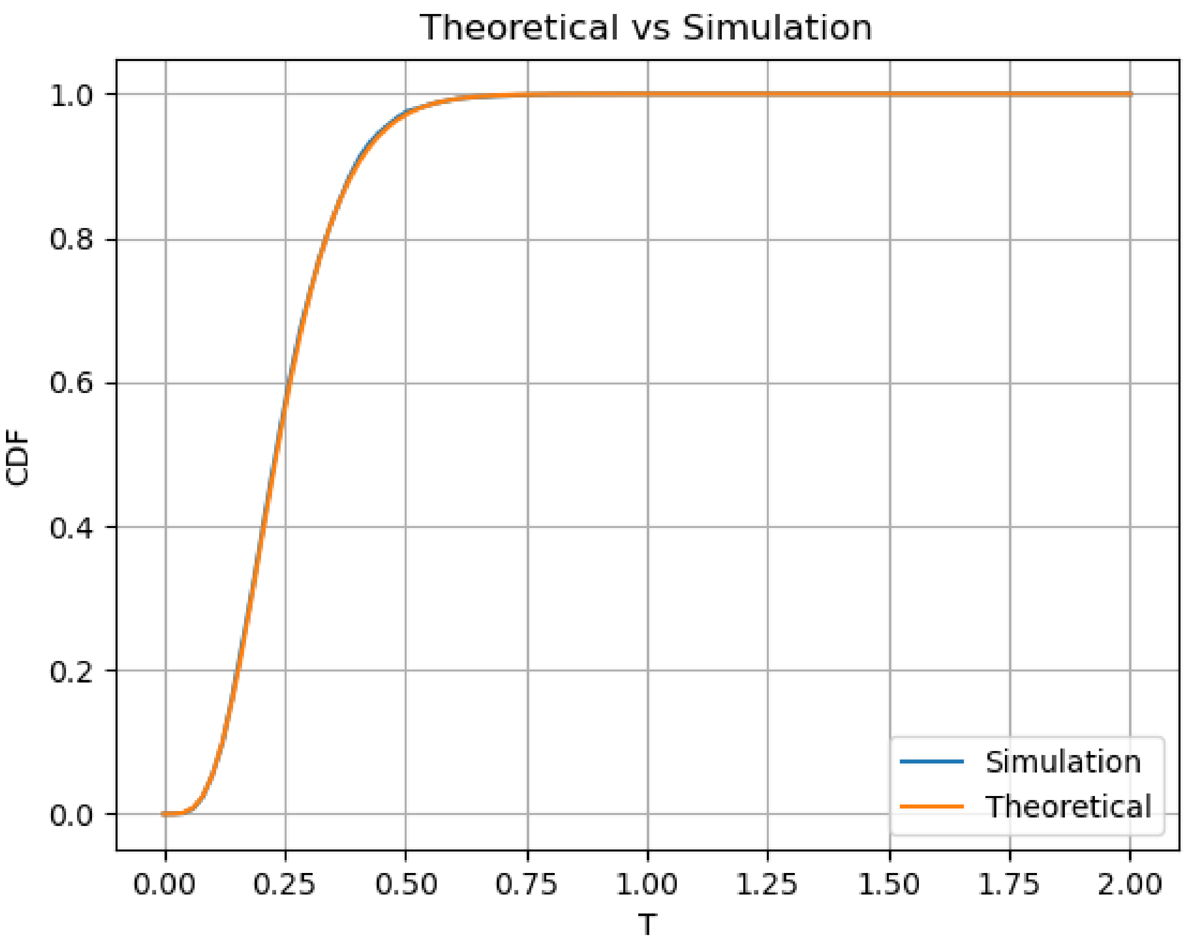



Fig. 4. Theoretical CDF vs Simulation CDF