

Question 41.2023

EE22BTECH11051

Question:

Suppose that X_1, X_2, \dots, X_{10} are independent and identically distributed random vectors each having $N_2(\mu, \Sigma)$ distribution, where Σ is non-singular. If

$$U = \frac{1}{1 + (\bar{X} - \mu)^T \Sigma^{-1} (\bar{X} - \mu)}, \quad (1)$$

where $\bar{X} = \frac{1}{10} \sum_{i=1}^{10} X_i$ then the value of $\log_e \Pr\left(U \leq \frac{1}{2}\right)$ equals

- 1) -5
- 2) -10
- 3) -2
- 4) -1

GATE ST 2023

Solution:

Parameter	Values	Description
n	10	Number of random vectors
μ		mean
Σ		Covariance

We are given a bivariate distribution;

$$X_i \sim N_2(\mu, \Sigma) \quad (2)$$

(3)

The distribution of \bar{X} can be given as:

The mean of \bar{X} will be the average of the means of X_1, X_2, \dots, X_{10} , which is:

$$\mu_{\bar{X}} = \frac{\mu + \mu + \mu + \dots + \mu}{10} = \frac{10\mu}{10} = \mu \quad (4)$$

And since the distributions X_1, X_2, \dots, X_{10} are independent, the covariance between them is zero. Hence we can find the new covariance as:

$$\text{Cov}(\bar{X}) = E[(\bar{X} - \mu)(\bar{X} - \mu)^T] \quad (5)$$

$$= E\left[\left(\frac{1}{10} \sum_{i=1}^{10} X_i - \mu\right)\left(\frac{1}{10} \sum_{j=1}^{10} X_j - \mu\right)^T\right] \quad (6)$$

$$= E\left[\frac{1}{100} \sum_{i=1}^{10} \sum_{j=1}^{10} (X_i - \mu)(X_j - \mu)^T\right] \quad (7)$$

$$= \frac{1}{100} \sum_{i=1}^{10} \sum_{j=1}^{10} E[(X_i - \mu)(X_j - \mu)^T] \quad (8)$$

Since X_1, X_2, \dots, X_{10} are independent and identically distributed random vectors, we get the covariance term $E[(X_i - \mu)(X_j - \mu)^T]$ to be Σ hence;

$$\text{Cov}(\bar{X}) = \frac{1}{100} \times 10 \times \Sigma \quad (9)$$

$$\bar{X} \sim N_2\left(\mu, \frac{\Sigma}{n}\right) \quad (10)$$

let's define Z as follows:

$$Z = \sqrt{10}\Sigma^{-1/2}(\bar{X} - \mu) \quad (11)$$

Now, express U in terms of Z :

$$U = \frac{1}{1 + (\bar{X} - \mu)^T \Sigma^{-1}(\bar{X} - \mu)} = \frac{1}{1 + Z^T Z} \quad (12)$$

Now, the distribution of Z is $N_2(0, I)$ (a multivariate normal distribution with mean 0 and covariance matrix I). Therefore, the expression for U becomes:

$$U = \frac{1}{1 + \|Z\|^2} \quad (13)$$

where $\|Z\|^2$ is the squared Euclidean norm of Z .

Now, the probability $\Pr(U \leq \frac{1}{2})$ is equivalent to

$$\Pr\left(\frac{1}{1 + \|Z\|^2} \leq \frac{1}{2}\right) = \Pr(\|Z\|^2 \geq 1) \quad (14)$$

The term $\|Z\|^2$ follows a χ_2^2 distribution, and we are looking for $\log_e \Pr(\|Z\|^2 \geq 1)$.

Since $k = 2$ for Y , we get;

$$= \int_{10}^{\infty} \frac{1}{2} e^{-\frac{y}{2}} dy \quad (15)$$

$$= e^{-5} \quad (16)$$

Hence;

$$\log_e \Pr\left(U \leq \frac{1}{2}\right) = -5 \quad (17)$$

Simulation Steps:

- 1) Set the number of simulations

- 2) Run a loop for number of simulations specified
- 3) Generate random vectors from a bivariate normal distribution
- 4) Calculate the sample mean and set the covariance matrix
- 5) Using the given formula, calculate U for each simulation
- 6) Calculate the simulated probability and take the natural log of that value

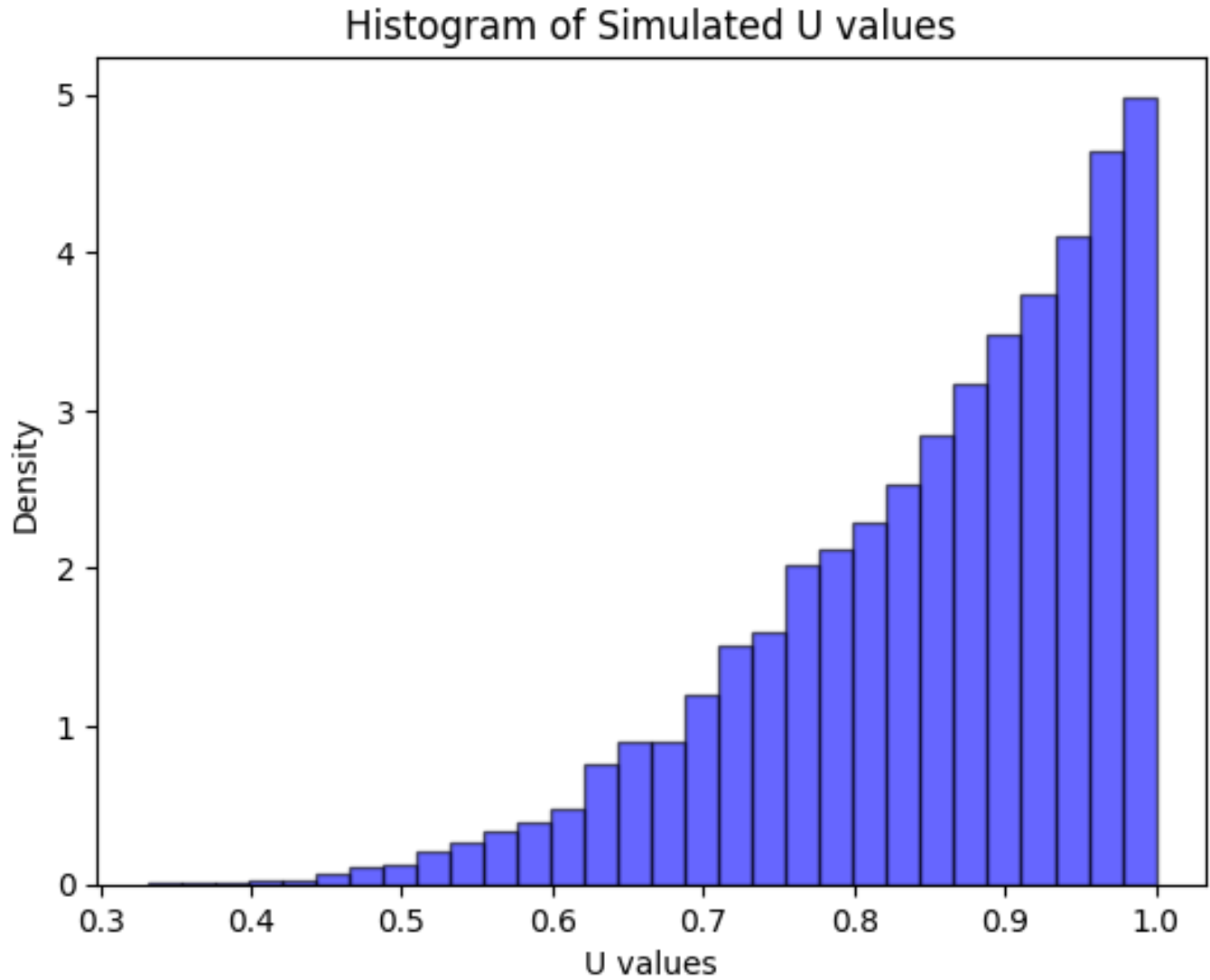


Fig. 1. Plot of $p_X(n)$. Simulations are close to the analysis.