Question 41.2023

EE22BTECH11051

Question:

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

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

where $\bar{X} = \frac{1}{10}\Sigma_{i=1}^{10} X_i$ 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;

$$X_i \sim N_2(\mu, \Sigma) \tag{2}$$

$$\tag{3}$$

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

The mean of $\bar{X}$ will be the average of the means of X1, X2, ..., X10, which is:

$$\mu_{\bar{X}} = \frac{\mu + \mu + \mu + ... + \mu}{10} = \frac{10\mu}{10} = \mu \tag{4}$$

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

$$\text{Cov}(\bar{X}) = E[(\bar{X} - \mu)(\bar{X} - \mu)^T] \tag{5}$$
$$= E[(\frac{1}{n}\Sigma_{i=1}^{n} X_i - \mu)(\frac{1}{n}\Sigma_{j=1}^{n} X_j - \mu)^T] \tag{6}$$
$$= E[\frac{1}{n^2}\Sigma_{i=1}^{n}\Sigma_{j=1}^{n}(X_i - \mu)(X_j - \mu)^T] \tag{7}$$
$$= \frac{1}{n^2}\Sigma_{i=1}^{n}\Sigma_{j=1}^{n}E[(X_i - \mu)(X_j - \mu)^T] \tag{8}$$

Since X_1,X_2, ...,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;

$$\text{Cov}(\bar{X}) = \frac{1}{n^2} \times n \times \Sigma \tag{9}$$
$$\bar{X} \sim N_2\left(\mu, \frac{\Sigma}{n}\right) \tag{10}$$

let's define Z as follows:

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

Now, express U in terms of Z:

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

Now, the distribution of Z is $N_2(0, I)$ (a multivariate normal distribution with mean 0 and covariance matrix I). This is given as:

$$\text{Let}\bar{Z} = \Sigma^{-1/2}(\bar{X} - \mu) \tag{13}$$

This transformation standardizes $\bar{X}$ to a standard multivariate normal distribution

$$\bar{Z} \sim N_2(0, I) \tag{14}$$
$$Z = \sqrt{n}\bar{Z} \tag{15}$$
$$\tag{16}$$

Now, we scale $\bar{Z}$ to get Z

$$Z = \sqrt{\bar{Z}} \tag{17}$$

This scaling does not change the fact that Z follows a multivariate normal distribution; it only changes the spread (variance) of the distribution. The mean of Z becomes 0, and the covariance matrix becomes nI due to the scaling by n
Hence we get;

$$Z \sim N_2(0, I) \tag{18}$$

Therefore, the expression for U becomes:

$$U = \frac{1}{1 + ||Z||^2} \tag{19}$$

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

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

$$\Pr\left(\frac{1}{1+\|Z\|^2} \leq \frac{1}{2}\right) = \Pr(\|Z\|^2 \geq 1) \tag{20}$$

The term $\|Z\|^2$ follows a χ^2_2 distribution, and we are looking for $\log_e \Pr(\|Z\|^2 \geq 1)$.

Since $k = 2$ and $n = 10$ for Y, we get;

$$= \int_{10}^{\infty} \frac{1}{2} e^{-\frac{y}{2}}\, dy \tag{21}$$

$$= e^{-5} \tag{22}$$

Hence;

$$\log_e \Pr\left(U \leq \frac{1}{2}\right) = -5 \tag{23}$$

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 adn take the natural log of that value

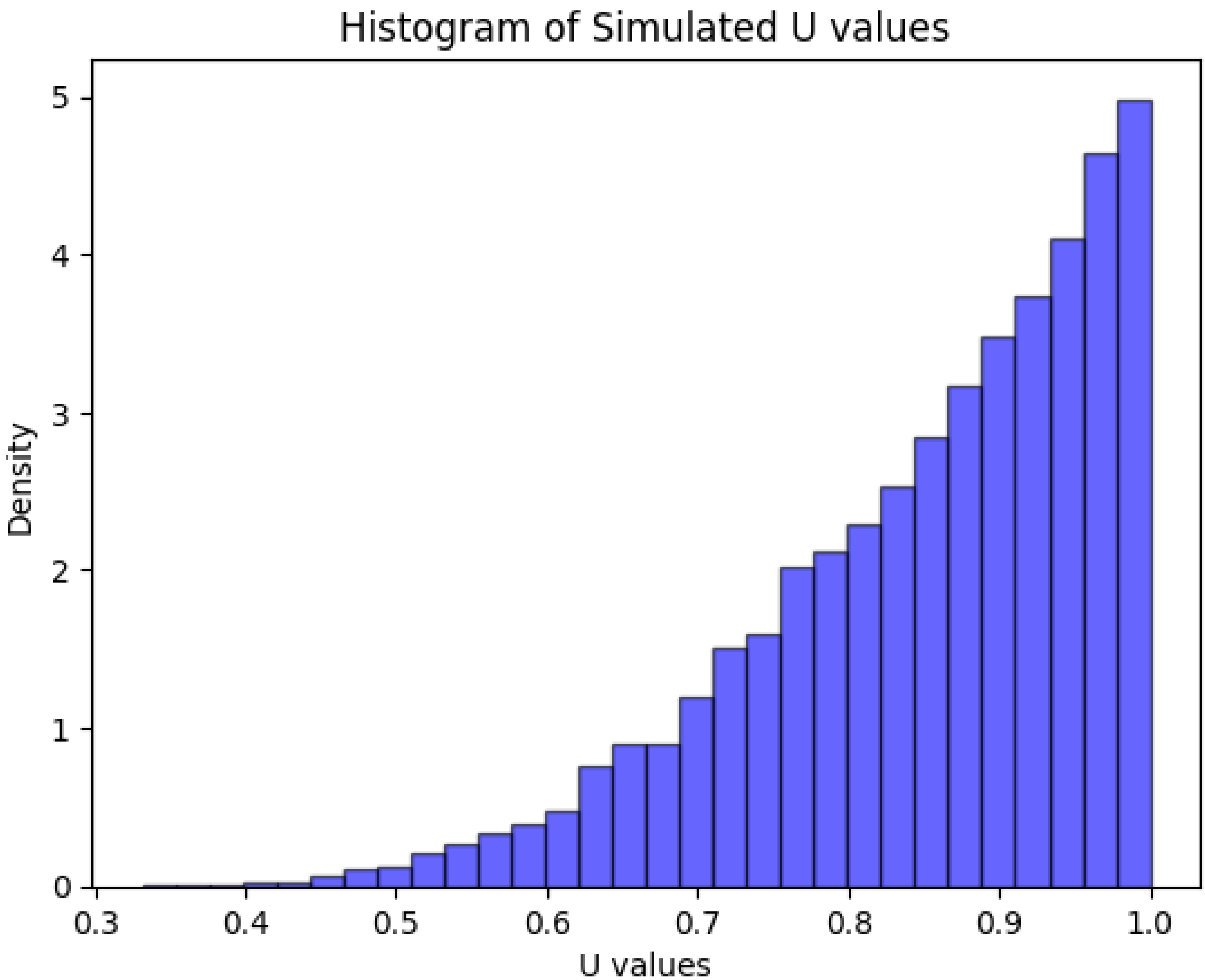

Fig. 1. Plot of $p_X(n)$.Simulations are close to the analysis.