Question 41.2023

EE22BTECH11051

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

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

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:
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:
The variance of a sum of i.i.d random variables is calculated as

$$\operatorname{var}(Y_n) = \text{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2\right] - \left(\text{E}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right]\right)^2 \tag{5}$$

$$= \frac{1}{n^2}\left\{\text{E}\left[\left(\sum_{i=1}^{n} X_i\right)^2\right] - \left(\text{E}\left[\sum_{i=1}^{n} X_i\right]\right)^2\right\} \tag{6}$$

But

$$\text{E}\left[\left(\sum_{i=1}^{n} X_i\right)^2\right] = \text{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{n} X_i X_j\right] \tag{7}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \text{E}\left[X_i X_j\right] \tag{8}$$

and

$$\left(\mathrm{E}\left[\sum_{i=1}^{n} X_i\right]\right)^2 = \left(\sum_{i=1}^{n} \mathrm{E}\left[X_i\right]\right)^2 \tag{9}$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \mathrm{E}\left[X_i\right] \mathrm{E}\left[X_j\right] \tag{10}$$

Putting (8) and (10) in (6) , and using the definition of covariance,

$$\operatorname{var}(Y_n) = \frac{1}{n^2}\left\{\sum_{i=1}^{n} \sum_{j=1}^{n} \left(\mathrm{E}\left[X_i X_j\right] - \mathrm{E}\left[X_i\right] \mathrm{E}\left[X_j\right]\right)\right\} \tag{11}$$

$$= \frac{1}{n^2}\left\{\sum_{i=1}^{n} \sum_{j=1}^{n} \operatorname{cov}\left(X_i, X_j\right)\right\} \tag{12}$$

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

$$\operatorname{cov}\left(X_i, X_j\right) = \begin{cases} 0 & \text{if } i \neq j \\ \operatorname{var}(X_i) & \text{if } i = j \end{cases} \tag{13}$$

$$\operatorname{var}(Y_n) = \frac{1}{n^2}\left(\sum_{i=1}^{n} \operatorname{cov}(X_i, X_i)\right) \tag{14}$$

$$= \frac{1}{n^2}\left(\sum_{i=1}^{n} \operatorname{var}(X_i)\right) \tag{15}$$

$$= \frac{1}{n^2}\left(\sum_{i=1}^{n} \Sigma\right) \tag{16}$$

$$= \frac{\Sigma}{n} \tag{17}$$

hence;

$$\bar{X} \sim N_2\left(\mu, \frac{\Sigma}{n}\right) \tag{18}$$

Now let us convert this normal distribution into χ_2^2 distrribution; where χ_k^2 distribution is given as;

$$f(x) = \frac{x^{\frac{k}{2}-1}}{2^{\frac{k}{2}}\Gamma(\frac{k}{2})} \quad x \geq 0 \tag{19}$$

$$\tag{20}$$

hence we get;

$$\frac{(\bar{X}-\mu)^T(\bar{X}-\mu)}{\left(\frac{\Sigma}{n}\right)} \sim \chi_2^2 \tag{21}$$

we can now write this as;

$$n(\bar{X}-\mu)^T \Sigma^{-1} (\bar{X}-\mu) \sim \chi_2^2 \tag{22}$$

Let Y be;

$$Y = n(\bar{X}-\mu)^T \Sigma^{-1} (\bar{X}-\mu) \tag{23}$$

which follows χ_2^2 distrribution
Now;

$$\Pr\left(U \leq \frac{1}{2}\right) = \Pr\left(\frac{1}{1+\frac{Y}{n}} \leq \frac{1}{2}\right) \tag{24}$$

$$= \Pr\left(\frac{Y}{n} \geq 1\right) \tag{25}$$

$$= \Pr\left(Y \geq 10\right) \tag{26}$$

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

$$= \int_{10}^{\infty} \frac{1}{2} e^{-\frac{y}{2}}\, dy \tag{27}$$

$$= e^{-5} \tag{28}$$

Hence;

$$\log_e \Pr\left(U \leq \frac{1}{2}\right) = -5 \tag{29}$$

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

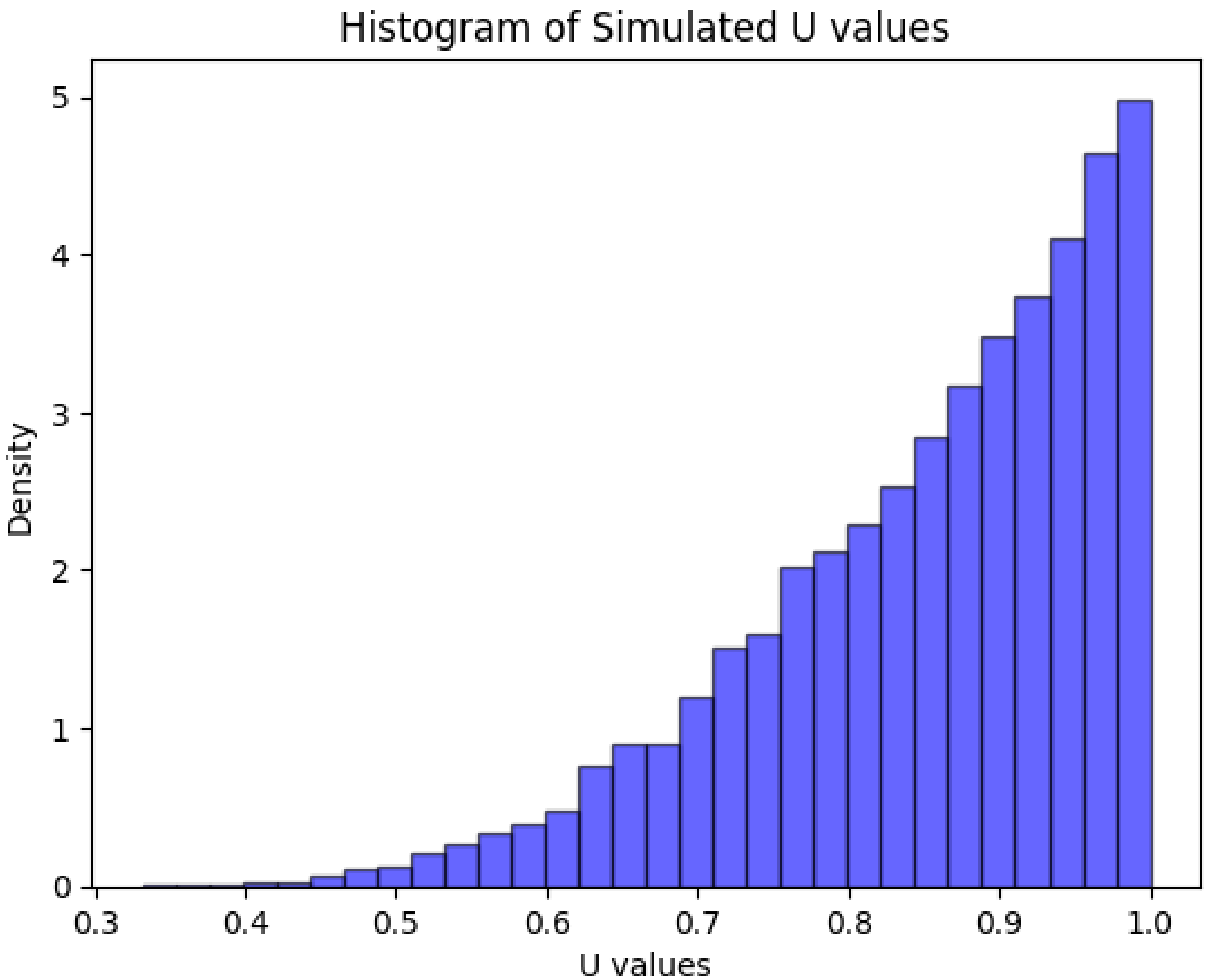

Fig. 1. Plot of $p_X(n)$.Simulations are close to the analysis.