ECE528 S18: Homework – MATLAB Assignment

Assigned on February 27; Due on March 20, 2018

Useful MATLAB commands:

lookfor, help, rand, randn, unidrnd, binornd, poissrnd, nchoosek, mean, var, cov, function, hist, plot

To generate random variables, two MATLAB functions are useful:

- "rand" generates one or a sequence of N random numbers that are uniformly distributed between 0 and 1, $\mathcal{U}[0, 1]$.
- "randn" generates one or a sequence of N random numbers that are Gaussian distributed with zero mean and unit variance, $\mathcal{N}(0, 1)$.

There are other random variable generators in MATLAB, such as "binornd" and "poissrnd" for generating Binomial and Poisson random variables respectively.

Set 1: Generating Random Variables of Certain Distributions

1. Let X be a random variable obeying Bernoulli distribution. In a general form, it takes on two possible values a and b with probabilities p and $1-p$ respectively, with $0 \le p \le 1$, as follows (w.p. stands for with probability):

$$X = \begin{cases} a, & \text{w.p. } p \\ b, & \text{w.p. } (1-p) \end{cases}$$

   (a) Let $a = 1$, $b = 0$, and $p = 0.6$. Use MATLAB to generate a sequence of $N = 20$ independent realizations of X. Write down these N numbers.

   (b) Plot the histogram of these N numbers? What do you observe?

   (c) Write a MATLAB function "*function X=my-bernoulli*(p, N, a, b)" that yields a sequence of N outcomes of the above Bernoulli random variable. You may use this function to conveniently simulate Bernoulli random variables in the future.

   Hint: Suppose that Y is uniformly distributed between $[0, 1]$, which can be generated using "*rand*". Let

$$Z = \begin{cases} 1, & 0 \le Y < 0.6 \\ 0, & 0.6 \le Y \le 1 \end{cases}$$

   What is the probability $P_Z(Z = 1) =$? What is the probability $P_Z(Z = 0) =$?

   Can you extend this idea to generate a Bernoulli distribution with any two possible values a and b, and any parameter $p \in [0, 1]$? The above RV Z is a special case of the RV X with $a = 1, b = 0, p = 0.6$.

   Can you extend this idea to generate the outcomes of a discrete random variable with a finite set of values? (see the next problem.)

Set 2. Computing mean and variance via simulations.

Let $\{x_n\}_{n=1}^N$ be N samples/realizations of a random variable X that obeys certain distribution. The sample mean $\bar{X}$ and sample variance s_X^2 are given by

$$\bar{X} = \langle X \rangle_N = \frac{1}{N} \sum_{n=1}^{N} x_n;$$

$$s_X^2 = \frac{1}{N-1} \sum_{n=1}^{N} (x_n - \bar{X})^2.$$

1. Let X be a random variable obeying Bernoulli distribution. It takes on two possible values a and b with probabilities p and $1-p$ respectively, as follows (w.p. stands for with probability):

$$X = \begin{cases} a, & \text{w.p. } p \\ b, & \text{w.p. } (1-p) \end{cases}$$

Let $a = 1$, $b = 0$, and $p = 0.6$.

(a) What are the analytical results for the mean m_X and variance σ_X^2 of X?

(b) Use MATLAB to generate a sequence of $N = 20$ independent realizations of X. Compute the sample mean $\bar{X}$ and variance s_X^2 using the formulae above. Compute the sample mean and variance using the MATLAB functions "mean" and "var". Do you get the same results?

(c) Repeat (b) for $N = 100$ and $N = 1000$. Do the sample mean and variance change? Are they closer to the analytical results? Document your observations concisely.

Set 3: Generating RVs of certain distributions as an inverse problem

The inverse problem:

For a function of one random variable, its inverse problem was explained in the last class. Solution to the inverse problem can be quite useful in system design, and in generating random numbers with arbitrary distributions for Monte Carlo simulations. This assignment provides such exercises.

1. The inverse problem above can be used to generate RVs of desired probabilistic distributions. We wish to generate a random number (RN) sequence x_i, $i = 1, \ldots, N$, with Rayleigh distribution. That is, $\{x_i\}_i$ are random realizations of a RV X that obeys Rayleigh distribution. In this case,

$$F_X(x) = 1 - e^{-x^2/(2\sigma^2)}; \qquad F_X^{(-1)}(u) = \sqrt{-2\sigma^2 \ln(1-u)}.$$

When U is uniformly distributed in $[0, 1]$, $(1 - U)$ is uniform in $[0, 1]$ as well. According to the inverse problem above, if we start with a RN sequence $\{u_i\}_{i=1}^N$ generated from a uniform distribution in $[0, 1]$, then $x_i = \sqrt{-2\sigma^2 \ln u_i}$, $i = 1, \ldots, N$, is a RN sequence with Rayleigh distribution.

(a) Find the mean and variance of the Rayleigh distributed X by analysis.

(b) Let $\sigma^2 = 1$. Generate a RN sequence of length $N = 1000$ with Rayleigh distribution. Show the pdf by plotting the histogram. Compute the sample mean $\bar{X} = \frac{1}{N}\sum_{i=1}^N x_i$ and sample variance $s_X^2 = \frac{1}{N-1}\sum_{i=1}^N (x_i - \bar{X})^2$.

(c) Repeat (b) for $\sigma^2 = 0.5, 5, 10$ respectively.

(d) Rayleigh distributed samples can also be generated from functions of Gaussian samples. Specifically, R is Rayleigh distributed with parameter σ^2 if $R = \sqrt{X_R^2 + X_I^2}$, where X_R and X_I are independent Gaussian RVs with the same distribution $\mathcal{N}(0, \sigma^2)$. That is, for a complex variable $\tilde{X} = X_R + jX_I$, its amplitude R is Rayleigh distributed. Hence, you may use Gaussian samples X_R and X_I to generate Rayleigh distributed RV R. Repeat (b) using this alternative approach.

(e) Briefly explain your observations.

2. We wish to generate a RN sequence y_i, $i = 1, \ldots, N$, with exponential distribution, that is,

$$F_Y(y) = 1 - e^{-y/\lambda}.$$

(a) Find the mean and variance of the exponential distributed Y by analysis.

(b) Generate a RN sequence of length $N = 1000$ with exponential distribution, for $\lambda = 1$. Show the pdf by plotting the histogram. Compute the sample mean and sample variance.

(c) Repeat (b) for $\lambda = 0.5, 5$ respectively.

(d) Briefly explain your observations.

Set 4. Law of Large Numbers.

1. Let X_i, $i = 1, \ldots, N$, be a sequence of independent identically distributed (i.i.d.) RVs, each following Rayleigh distribution with $\sigma = 1$. Consider a new RV given by the arithmetic mean

$$X = \frac{1}{N} \sum_{i=1}^{N} X_i.$$

(a) Suppose that $\{X_i\}_{i=1}^N$ are i.i.d. RVs with Rayleigh distribution ($\sigma = 1$) and $N = 100$. Generate $L = 1000$ realizations of X and show the pdf of X by plotting its histogram. Based on the histogram, do you think the pdf of X can be well approximated by a Gaussian pdf?

(b) Suppose that $\{X_i\}_{i=1}^N$ are i.i.d. RVs with Rayleigh distribution ($\sigma = 1$) and $N = 10000$. Generate $L = 1000$ realizations of X and show the pdf of X by plotting its histogram. Do you think the pdf of X can be well approximated by a Gaussian pdf?

(c) Suppose that $\{Y_i\}_{i=1}^N$ are i.i.d. RVs with exponential distribution ($\lambda = 1$) and $N = 100$. Generate $L = 1000$ realizations of $Y = \frac{1}{N} \sum_{i=1}^N Y_i$ and show the pdf of Y by plotting its histogram. Do you think the pdf of X can be well approximated by a Gaussian pdf?

(d) Suppose that $\{Y_i\}_{i=1}^N$ are i.i.d. RVs with exponential distribution ($\lambda = 1$) and $N = 10000$. Generate $L = 1000$ realizations of $Y = \frac{1}{N} \sum_{i=1}^N Y_i$ and show the pdf of Y by plotting its histogram. Do you think the pdf of X can be well approximated by a Gaussian pdf?

(e) Let

$$Z = \frac{X_1 + \cdots + X_N}{Y_1 + \cdots + Y_N}$$

where $\{X_i\}$ are i.i.d. Rayleigh distributed RVs and $\{Y_i\}$ are i.i.d. exponentially distributed RVs. $N = 100$. Generate $L = 1000$ random realizations of Z and compute the sample mean. [Note: see Problems 5.13 and 5.14 on Page 292 of the textbook]

(f) Repeat (e) for $N = 10000$.

(g) Document your observations concisely.