Probability and Stochastic Processes (EL6303) November 6, 2018
NYU Tandon School of Engineering, Fall 2018
Instructor: *Dr. Elza Erkip*

Exercise 6

1. Suppose $X_1, X_2, \dots$ are independent and identically distributed random variables that are uniformly distributed in the interval $(0, 2a)$ where $a > 0$.

   (a) We first consider the product of $X_1, \dots, X_n$. Let
   $$Y = \prod_{i=1}^{n} X_i.$$
   Find $E(Y)$.

   (b) Now suppose that instead of taking the product of n X_i's, we flip a biased coin determine how many X_i's we should include in the above product. Let N be the number of coin flips until we see the first Head. We assume P(Head) $= p, 0 < p < 1$.

      i. Show that N has the geometric distribution.

      ii. Let
      $$Z = \prod_{i=1}^{N} X_i.$$
      Find $E(Z)$.
      Hint 1: You can use $E(Z) = E(E(Z|N))$.
      Hint 2: $\sum_{i=0}^{\infty} \alpha^i = \frac{1}{1-\alpha}$ for $\alpha < 1$.

      iii. For what values of a will $E(Z)$ be finite? Explain.

2. Consider $X \sim$ Uniform$(-1, 1)$ and $Y = X^2$.

   (a) Find $E(X)$ and $E(Y)$.

   (b) Find the correlation coefficient between X and Y.

   (c) Are X and Y independent? Are they uncorrelated? Explain.

   (d) Find an estimate $\hat{Y} = h(X)$ of Y such that mean square error $e_1 = E[(Y - \hat{Y})^2]$ is minimized. Find the resulting mean square error e_1.

(e) Find a linear estimate $\tilde{Y} = l(X) = aX + b$ of Ysuch that mean square error $e_2 = E[(Y - \tilde{Y})^2]$ is minimized. Find the resulting mean square error e_2.

(f) Compare e_1 and e_2. Which one is smaller? Comment.

3. A 2×2 multi-input multi-output (MIMO) channel is represented as follows:

$$\mathbf{Y} = \mathbf{H}\mathbf{X},$$

where $\mathbf{X} = (X_1, X_2)^t$ is the channel input, $\mathbf{Y} = (Y_1, Y_2)^t$ is the channel output and

$$\mathbf{H} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

is the channel matrix. Here $\mathbf{X}^t$ denotes transpose of vector $\mathbf{X}$.

We assume the channel $\mathbf{H}$ is a known matrix, and the channel input $\mathbf{X}$ is a random vector. We also assume that

$$E(X_1) = E(X_2) = 0, \mathrm{Var}(X_1) = \sigma_1^2, \mathrm{Var}(X_2) = \sigma_2^2, \mathrm{Corr}(X_1, X_2) = \rho.$$

(a) Find $\mathrm{Var}(Y_1)$.

(b) Find $E(Y_1 Y_2)$.

(c) For $\sigma_1^2 = \sigma_2^2 = \sigma^2$ and $a = b = c = 1$, find d such that Y_1 and Y_2 are uncorrelated.

(d) Consider the noisy version of the above channel

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{Z},$$

where $\mathbf{Z} = (Z_1, Z_2)^t$ represents Gaussian noise with Z_1 and Z_2 independent, $N(0, 1)$. We assume $\mathbf{Z}$ is independent of $\mathbf{X}$. We also assume X_1 and X_2 are jointly Gaussian.

For $\rho = 0$ and arbitrary $\mathbf{H}, \sigma_1^2, \sigma_2^2$, find the pdf of Y_1. Explain your answer.

4. You are at the first floor of Rogers Hall, trying to get to your classroom on the fifth floor. There are three elevators, and elevator i arrives first with probability $p_i, i = 1, 2, 3$. Once you get on elevator i, it takes T_i minutes to get to the fifth floor. Here $T_i, i = 1, 2, 3$ represent random variables that are not necessarily independent.

You take the first elevator that arrives to get to your class. Suppose T denotes the time it takes for you to get to the fifth floor.

(a) Find an expression for $E(T)$ in terms of p_i and $T_i, i = 1, 2, 3$.

(b) For $p_1 = 0.2, p_2 = 0.3$, $T_i \sim \mathrm{Uniform}(0, i + 1)$ calculate $E(T)$.

5. Consider a random variable θ that is uniformly distributed in the interval $(0, 1]$. Suppose for a given θ, $X_1, \ldots, X_n$ are iid according to Geometric distribution with parameter θ. That is:

$$P(X_k = i|\theta) = (1-\theta)^{i-1}\theta, i = 1, 2, \ldots,$$

for all $k = 1, \ldots, n$.

   (a) Find the pmf of X_1, that is $P(X_1 = i), i = 1, 2, \ldots$.

   (b) Find the joint pmf of (X_1, X_2). Are X_1 and X_2 independent?

   (c) i. Find $P(X_1, \ldots, X_n|\theta)$.

      ii. Find $\theta \in (0, 1]$ that maximizes $P(X_1, \ldots, X_n|\theta)$. This is called *the Maximum Likelihood (ML)* estimate of θ.

   (d) The *Maximum-a-Posteriori (MAP)* estimate of θ is the one that maximizes $P(\theta|X_1, \ldots, X_n)$. For this question would the MAP estimate and the ML estimate be the same or different? Explain your answer. *You do not need to compute the MAP estimate to answer this question.*

   Hint: You can use $\int_0^1 u^k(1-u)^{n-k}du = \frac{k!(n-k)!}{(n+1)!}$.

6. The random variables X and Y have the joint probability density function

$$f_{X,Y}(x, y) = \begin{cases} 2(y + x) & 0 \leq x \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

   (a) What is $f_{X|Y}(x|y)$, the conditional PDF of X given $Y = y$?

   (b) What is $\hat{x}_M(y)$, the MMSE estimate of X given $Y = y$?

   (c) What is $f_{Y|X}(y|x)$, the conditional PDF of Y given $X = x$?

   (d) What is $\hat{y}_M(x)$, the MMSE estimate of Y given $X = x$?