UC Berkeley
Department of Electrical Engineering and Computer Sciences

EE126: Probability and Random Processes

Homework III
Spring 2020

Problem 1. **Graphical Density**
Figure **??** shows the joint density $f_{X,Y}$ of the random variables X and Y.

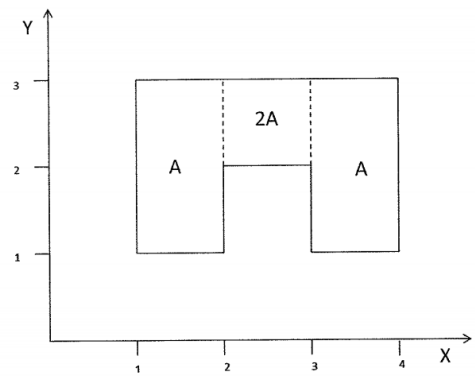

Figure 1: Joint density of X and Y.

(a) Find A and sketch f_X, f_Y, and $f_{X|X+Y\leq 3}$.

(b) Find $\mathbb{E}[X \mid Y = y]$ for $1 \leq y \leq 3$ and $\mathbb{E}[Y \mid X = x]$ for $1 \leq x \leq 4$.

(c) Find $\text{cov}(X, Y)$.

Problem 2. **Records** Let n be a positive integer and $X_1, X_2, \ldots, X_n$ be a sequence of i.i.d. continuous random variable with common probability density f_X. For any integer $2 \leq k \leq n$, define X_k as a record-to-date of the sequence if $X_k > X_i$ for all $i = 1, \ldots, k-1$. (X_1 is automatically a record-to-date.)

(a) Find the probability that X_2 is a record-to-date.

Hint: You should be able to do it without rigorous computation.

(b) Find the probability that X_n is a record-to-date.

(c) Find the expected number of records-to-date that occur over the first n trials (Hint: Use indicator functions.) Compute this when $n \to \infty$.

Problem 3. **Joint Density for Exponential Distribution**

(a) If $X \sim Exp(\lambda)$ and $Y \sim Exp(\mu)$, X and Y independent, compute $\mathbb{P}(X < Y)$.

(b) If X_k, $1 \leq k \leq n$ are exponentially distributed with parameters $\lambda_1, \ldots, \lambda_n$, show that,

$$\mathbb{P}(X_i = \min_{1 \leq k \leq n} X_k) = \frac{\lambda_i}{\sum_{j=1}^{n} \lambda_j}$$

Problem 4. **Change of Variables**
Let X be a R.V with PDF $f_X(x)$ and CDF $F_X(x)$. Let $g(X)$ be an invertible function. The change of variables problem asks for the density of $Y = g(X)$ which is a new R.V. To find this distribution, we use the definition of the CDF

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \leq g^{-1}(y)) = F_X(g^{-1}(y))$$

How would you do this if the function g was not invertible? For example $g(x) = x^2 \forall x \in \mathbb{R}$ has two values in the domain mapping to each value in its range. In these cases we have to include all parts of the domain that contribute to the probability. In the case of a discrete distribution using the above g, this would look like

$$P(Y = y) = P(g(X) = y) = P(X \in \{-\sqrt{y}, \sqrt{y}\})$$

1. Suppose that X has the **standard normal distribution**, that is, X is a continuous random variable with density function

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{x^2}{2}\Big).$$

   What is the density function of $\exp X$? (The answer is called the **lognormal distribution**.)

2. Suppose that X is a continuous random variable with density f. What is the density of X^2?

3. What is the answer to the previous question when X has the standard normal distribution? (This is known as the **chi-squared distribution**.)

Problem 5. **Unit Circle on a Grid**
A circle of unit radius is thrown on an infinite sheet of grid paper that is has square grids of unit length on each side. Find the expected number of vertex points of the grid that fall inside the circle.

Problem 6. **Matrix Sketching**
Matrix sketching is an important technique in randomized linear algebra to do large computations efficiently. For example, to compute the multiplication $\mathbf{A}^T \times \mathbf{B}$ of two large matrices $\mathbf{A}$ and $\mathbf{B}$, we can use a random sketch matrix $\mathbf{S}$ to compute a "sketch" $\mathbf{SA}$ of $\mathbf{A}$ and a "sketch" $\mathbf{SB}$ of $\mathbf{B}$. Such a sketching matrix has the property that $\mathbf{S}^T\mathbf{S} \approx \mathbf{I}$ so that the approximate multiplication $\mathbf{A}^T\mathbf{S}^T\mathbf{SB}$ is close to $\mathbf{A}^T\mathbf{B}$.
In this problem, we will discuss two popular sketching schemes and understand how they help in approximate computation. Let $\hat{\mathbf{I}} = \mathbf{S}^T\mathbf{S}$ and the dimension of sketch matrix $\mathbf{S}$ be $d \times n$ (typically $d \ll n$).

1. (**Gaussian-sketch**) Define

$$\mathbf{S} = \frac{1}{\sqrt{d}} \begin{bmatrix} S_{11} & \dots & \dots & S_{1n} \\ \vdots & \ddots & & \vdots \\ S_{d1} & \dots & \dots & S_{dn} \end{bmatrix}$$

such that S_{ij}'s are chosen i.i.d. from $\mathcal{N}(0,1)$ for all $i \in [1,d]$ and $j \in [1,n]$. Find the element-wise mean and variance (as a function of d) of the matrix $\hat{\mathbf{I}} = \mathbf{S}^T\mathbf{S}$, that is, find $\mathbb{E}[\hat{I}_{ij}]$ and $\mathrm{Var}[\hat{I}_{ij}]$ for all $i \in [1,n]$ and $j \in [1,n]$.

2. (**Count-sketch**) For each column $j \in [1,n]$ of $\mathbf{S}$, choose a row i uniformly randomly from $[1,d]$ such that

$$S_{ij} = \begin{cases} 1, & \text{with probability } 0.5 \\ -1, & \text{with probability } 0.5 \end{cases}$$

and assign $S_{kj} = 0$ for all $k \neq i$. An example of a 3×8 count-sketch is

$$\mathbf{S} = \begin{bmatrix} 0 & -1 & 1 & 0 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \end{bmatrix}$$

Again, find the element-wise mean and variance (as a function of d) of the matrix $\hat{\mathbf{I}} = \mathbf{S}^T\mathbf{S}$.

Note that for sufficiently large d, the matrix $\hat{\mathbf{I}}$ is close to the identity matrix for both cases. We will use this fact in the lab to do an approximate matrix multiplication. **Note:** You can use the fact that the fourth moment of a standard Gaussian is 3 without proof.