UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: Probability and Random Processes

Homework 03
Fall 2023

1. **Expected Norm**

   Pick two points $X = (X_1, X_2)$ and $Y = (Y_1, Y_2)$ independently and uniformly in $[0,1]^2$. Calculate $\mathbb{E}(\|X - Y\|_2^2)$.

   Solution: We observe that $X_1, X_2, Y_1, Y_2 \sim \text{Uniform}([0,1])$, and

   $$\mathbb{E}(\|X - Y\|_2^2) = \mathbb{E}((X_1 - Y_1)^2) + \mathbb{E}((X_2 - Y_2)^2).$$

   We can calculate $\mathbb{E}((X_1 - Y_1)^2) = \mathbb{E}(X_1^2) - 2\,\mathbb{E}(X_1)\,\mathbb{E}(Y_1) + \mathbb{E}(Y_1^2) = 2 \cdot \frac{1}{3} - \frac{1}{2} = \frac{1}{6}$. By symmetry, we get that $\mathbb{E}(\|X - Y\|_2^2) = \frac{1}{3}$.

2. **Joint Density for Exponential Distribution**

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

   b. If $X_1, \ldots, X_n$ are independent and Exponentially distributed with parameters $\lambda_1, \ldots, \lambda_n$, show that $\min_{1 \le k \le n} X_k \sim \text{Exponential}(\sum_{j=1}^n \lambda_j)$.

   c. Deduce that
$$\mathbb{P}\left(X_i = \min_{1 \le k \le n} X_k\right) = \frac{\lambda_i}{\sum_{j=1}^n \lambda_j}.$$

Solution:

a. By the law of total probability,
$$\mathbb{P}(X < Y) = \int_0^\infty \mathbb{P}(X < y \mid Y = y) \cdot f_Y(y)\,\mathrm{d}y.$$
Since X and Y are independent, $\mathbb{P}(X < y \mid Y = y) = \mathbb{P}(X < y)$. Plugging in the known $\mathbb{P}(X < y) = 1 - e^{-\lambda y}$ and $f_Y(y) = \mu e^{-\mu y}$, we get
$$\mathbb{P}(X < Y) = \frac{\lambda}{\lambda + \mu}.$$

b. The ccdf of $X := \min_{1 \le k \le n} X_k$ is precisely the ccdf of an $\text{Exponential}(\sum_{j=1}^n \lambda_j)$:
$$\mathbb{P}(X \ge x) = \mathbb{P}(X_1 \ge x, \ldots, X_n \ge x) = \prod_{k=1}^n \mathbb{P}(X_k \ge x) = \prod_{k=1}^n e^{-\lambda_k x} = e^{-x \sum_{k=1}^n \lambda_k}.$$

c. Now, we observe that
$$\mathbb{P}\left(X_i = \min_{1 \le k \le n} X_k\right) = \mathbb{P}\left(X_i \le \min_{k \ne i} X_k\right).$$
By part b, $\min_{k \ne i} X_k \sim \text{Exponential}(\sum_{j \ne i} \lambda_j)$. Then, by part a, the claim follows.

3. **Change of Variables**

Let X be a continuous random variable with cdf F_X and pdf $f_X > 0$ everywhere, and let $Y = g(X)$, where g is a differentiable function.

a. Suppose that g is also invertible. Find the pdf of Y, f_Y, in terms of g and f_X.
b. Let $U \sim \text{Uniform}([0,1])$. Using the conclusion from part a, show that $F_X^{-1}(U)$ has the same distribution as X. (This allows us to generate a given random variable given only a uniform random number generator.)
c. Now suppose that $g(x) = x^2$. Find the pdf of Y in terms of the pdf of X. Also find the pdf of Y when X is a standard normal random variable in particular.
(Note that this g is not invertible, unlike in part a.)

Solution:

a. g is a continuous invertible function from $\mathbb{R}$ to $\mathbb{R}$, so g must be monotonic, i.e. strictly increasing or strictly decreasing. Let us first find the cdf of Y:

$$F_Y(y) = \mathbb{P}(g(X) \leq y) = \begin{cases} \mathbb{P}(X \leq g^{-1}(y)) = F_X(g^{-1}(y)) & \text{if } g \text{ is increasing} \\ \mathbb{P}(X \geq g^{-1}(y)) = 1 - F_X(g^{-1}(y)) & \text{if } g \text{ is decreasing.} \end{cases}$$

Then, by the chain rule of differentiation, we find the pdf of Y as

$$f_Y(y) = \frac{\mathrm{d}}{\mathrm{d}y} F_Y(y) = \left| \frac{\mathrm{d}}{\mathrm{d}y} F_X(g^{-1}(y)) \right| = f_X(g^{-1}(y)) \cdot \left| \frac{\mathrm{d}}{\mathrm{d}y} g^{-1}(y) \right|.$$

Using the inverse function rule, we can further simplify to

$$f_Y(y) = f_X(g^{-1}(y)) \cdot \frac{1}{|g'(g^{-1}(y))|}.$$

b. Let $Y = F_X^{-1}(U)$. F_X is differentiable because X is a continuous random variable, and strictly increasing because $f_X > 0$ everywhere, so its inverse F_X^{-1} is also differentiable and monotonically increasing. Using the conclusion of part a with $g = F_X^{-1}$,

$$F_Y(y) = F_U(g^{-1}(y)) = F_U(F_X(y)) = F_X(y),$$

which shows that Y has the same distribution as X. Note that $F_U(u) = \mathbb{P}(U \leq u) = u$ for $U \sim \text{Uniform}([0,1])$.

c. The cdf of $Y = X^2$ is

$$\mathbb{P}(X^2 \leq y) = \mathbb{P}(-\sqrt{y} \leq X \leq \sqrt{y}) = \int_{-\sqrt{y}}^{\sqrt{y}} f_X(x)\,\mathrm{d}x.$$

By the fundamental theorem of calculus, the pdf of Y is

$$f_Y(y) = \frac{1}{2\sqrt{y}}(f_X(-\sqrt{y}) + f_X(\sqrt{y})).$$

For $X \sim \mathcal{N}(0,1)$, the pdf of X^2 evaluates to

$$f_Y(y) = \frac{1}{2\sqrt{y}} \left(\frac{1}{\sqrt{2\pi}} e^{-y/2} + \frac{1}{\sqrt{2\pi}} e^{-y/2} \right) = \frac{1}{\sqrt{2\pi y}} e^{-y/2}.$$

(This is known as the **chi-squared** distribution with 1 degree of freedom.)

4. **Really Random Binomial**

Consider the random variables $U \sim \text{Uniform}([0,1])$ and $X|U \sim \text{Binomial}(n, U)$, where X is a binomial random variable with a random success probability. Given that $X = k$, we wish to find the conditional distribution of U, $f_{U|X}(u \mid k)$ using the steps below.

   a. Write $f_{U|X}(u \mid k)$ in terms of the distributions of X, U, and $X \mid U$ using Bayes' Rule. Plug in any distribution given in the setup.

   b. You may realize that the denominator $\mathbb{P}(X = k)$ of your expression above is hard to evaluate. It requires integrating over values of U and iterative integration by parts. Instead, we resort to an approach based on moment generating functions. Write the mgf of X as a summation in terms of $\mathbb{P}(X = k)$. Then, write $\mathbb{P}(X = k)$ as an integral over values of U and exchange the summation and integration. Use the binomial theorem to absorb the summation so we are left with an integral.

   c. Carry out the evaluation of the integral. Use the identity $\frac{1-s^{n+1}}{1-s} = \sum_{i=0}^{n} s^i$ to leave your answer as a summation. Does this expression look like the mgf of some discrete random variable, and which one?

   d. Conclude the distribution of X is the distribution of the discrete random variable you found above. Use this to find $\mathbb{P}(X = k)$, then find $f_{U|X}(u \mid k)$.

Solution:

   a. By Bayes' rule,

$$f_{U|X}(u \mid k) = \frac{p_{X|U}(k \mid u) \cdot f_U(u)}{\mathbb{P}(X = k)}.$$

   We know that $p_{X|U}(k \mid u) = \binom{n}{k} u^k (1-u)^{n-k}$ and $f_U(u) = 1$ by the distributions given.

   b.

$$\begin{aligned}
M_X(s) = \mathbb{E}(e^{sX}) &= \sum_{k=0}^{n} e^{sk}\, \mathbb{P}(X = k) \\
&= \sum_{k=0}^{n} e^{sk} \int_0^1 p_{X|U}(k \mid u) \cdot f_U(u)\, du \\
&= \sum_{k=0}^{n} e^{sk} \int_0^1 \binom{n}{k} u^k (1-u)^{n-k} \cdot 1\, du \\
&= \int_0^1 \left(\sum_{k=0}^{n} \binom{n}{k} (u \cdot e^s)^k (1-u)^{n-k} \right) du \\
&= \int_0^1 (e^s u + (1-u))^n\, du.
\end{aligned}$$

   c. Now we can substitute $v = (e^s u + (1-u))$, which gives $dv = -(1 - e^s)\, du$. Changing the limits appropriately,

$$\begin{aligned}
M_X(s) &= -\frac{1}{1-e^s} \int_1^s v^n\, dv \\
&= \frac{1 - e^{s(n+1)}}{(n+1)(1-e^s)}
\end{aligned}$$

$$= \frac{1}{n+1} \sum_{k=0}^{n} e^{sk}.$$

We observe that this MGF corresponds to a discrete random variable taking values in $0, \ldots, n$, each with probability $\frac{1}{n+1}$. In other words, X is a discrete uniform distribution over the specified range since mgf uniquely characterizes the distribution.

d. Thus, our final posterior on U is

$$f_{U|X}(u \mid k) = (n+1) \cdot \binom{n}{k} u^k (1-u)^{n-k}.$$

Remark: Here is another way of understanding why X is uniform. Recall that a binomial distribution can be written as the sum of i.i.d. Bernoulli random variables. Next, recognize that if V is uniform on $[0, 1]$, the random variable $\mathbb{1}_{V \leq u}$ is Bernoulli with parameter u. The difference is, the number u is itself another uniform random variable, call it U_{n+1}. This means we can write $X = \sum_{i=1}^{n} \mathbb{1}_{U_i \leq U_{n+1}}$, and $X = k$ implies that U_{n+1} occurs in the $(k+1)$-th position overall, which occurs with uniform probability $\frac{1}{n+1}$ since we are dealing with $n + 1$ i.i.d. random variables.

5. **Poisson Practice**

Suppose X is a Poisson random variable with parameter λ. Find the following:

a. $\mathbb{E}(X^2)$.

b. $\mathbb{P}(X \text{ is even})$. (*Hint*: Use the Taylor series expansion of e^x.)

Solution:

a. First compute

$$\begin{aligned}\mathbb{E}[X(X-1)] &= \sum_{x=2}^{\infty} x(x-1)\frac{\lambda^x \varepsilon^{-\lambda}}{x!} = \sum_{x=2}^{\infty} \frac{\lambda^x \varepsilon^{-\lambda}}{(x-2)!} = \lambda^2 \sum_{x=2}^{\infty} \frac{\lambda^{x-2}\varepsilon^{-\lambda}}{(x-2)!} \\ &= \lambda^2.\end{aligned}$$

Hence, $\mathbb{E}[X^2] = \mathbb{E}[X(X-1)] + \mathbb{E}[X] = \lambda^2 + \lambda$.

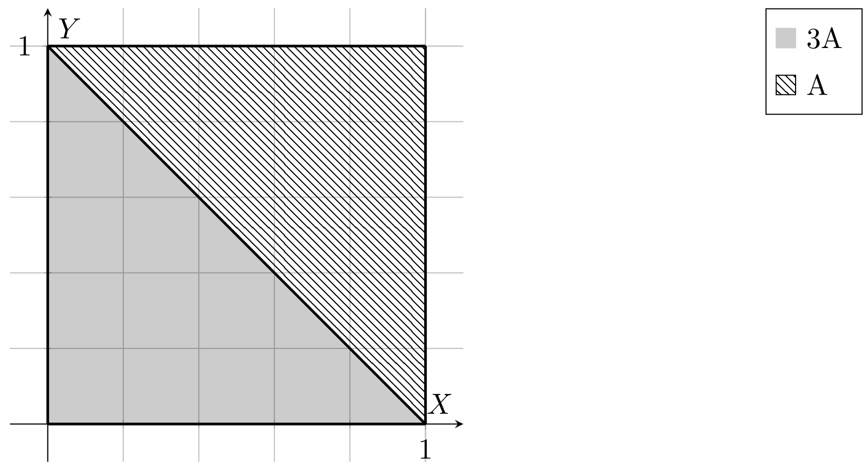
b. Note that

$$\begin{aligned}\Pr(X \text{ is even}) &= \sum_{k=0}^{\infty} \Pr(X = 2k) = \sum_{k=0}^{\infty} \frac{\lambda^{2k}\varepsilon^{-\lambda}}{(2k)!} \\ &= \frac{\varepsilon^{-\lambda}}{2}\left(\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} + \sum_{k=0}^{\infty} \frac{(-\lambda)^k}{k!}\right) = \frac{\varepsilon^{-\lambda}}{2}(\varepsilon^{\lambda} + \varepsilon^{-\lambda}) \\ &= \frac{1+\varepsilon^{-2\lambda}}{2}.\end{aligned}$$

To explain the second line, note that the odd terms cancel out and the even terms are counted twice.

6. **Graphical Density**

The following figure depicts the joint density $f_{X,Y}$ of X and Y.



a. Are X and Y independent? Remember to justify your answer.
b. What is the value of A?
c. Compute $f_X(x)$.
d. Compute $\mathbb{E}(Y \mid X = x)$. You may leave your answer as a fraction of terms containing x, but you may not have an integral.
e. What is $\mathbb{E}(X - Y \mid X + Y)$?

Solution:

a. X and Y are not independent. For example, when $X = 0$, the expected value of Y is $\frac{1}{2}$, but when $X = \frac{1}{2}$, the expected value of Y is less than $\frac{1}{2}$, since there is more probability mass in the bottom triangle.
b. The total probability density must integrate to 1, so $\frac{1}{2}(3A) + \frac{1}{2}A = 1$ implies that $A = \frac{1}{2}$.
c. The vertical line at $X = x$ breaks up into two pieces in each triangle:

$$\begin{aligned} f_X(x) = \int_0^1 f_{X,Y}(x, y)\, \mathrm{d}y &= \int_0^{1-x} \frac{3}{2}\, \mathrm{d}y + \int_{1-x}^1 \frac{1}{2}\, \mathrm{d}y \\ &= \frac{3}{2}(1 - x) + \frac{1}{2}x = \frac{3}{2} - x. \end{aligned}$$

d. As $f_{Y|X}(y \mid x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$, we can use the definition of conditional expectation directly:

$$\begin{aligned} \mathbb{E}(Y \mid X = x) &= \int_0^{1-x} y \frac{3A}{\frac{3}{2} - x}\, \mathrm{d}y + \int_{1-x}^1 y \frac{A}{\frac{3}{2} - x}\, \mathrm{d}y \\ &= \frac{3A(1 - x)^2}{3 - 2x} + \frac{A(1 - (1 - x)^2)}{3 - 2x} \\ &= \frac{3 - 4x + 2x^2}{2(3 - 2x)}. \end{aligned}$$

Alternate solution. We can split this expectation into the cases where Y falls in the $3A$ region (when it falls below $1-x$) and where Y falls in the A region. In the first case, its expectation will be $\frac{1-x}{2}$; in the second case, its expectation will be $\frac{1+1-x}{2} = \frac{2-x}{2}$. It remains to figure out the probability that Y falls below $1-x$. Let B be the (constant) density $f_{Y|X}(y \mid x)$ for y above $1-x$, so that $3B$ is the density of Y below $1-x$. In order to integrate to 1, we must have

$$3B(1-x) + Bx = 1,$$

which implies that $B = \frac{1}{3-2x}$. Then we have

$$\mathbb{E}(Y \mid X = x) = \frac{1-x}{2} \cdot \frac{3(1-x)}{3-2x} + \frac{2-x}{2} \cdot \frac{x}{3-2x},$$

which yields the same result after simplifying.

e. We see that given $X + Y = c$, which is a line parallel to the diagonal line in the graph, the values of $X - Y$ (which is perpendicular to $X + Y$) are uniformly distributed, centered around 0. So $\mathbb{E}(X - Y \mid X + Y) = 0$. Another way to see this is that $\mathbb{E}(X \mid X + Y) = \mathbb{E}(Y \mid X + Y) = \frac{X+Y}{2}$, so by linearity of expectation, $\mathbb{E}(X - Y \mid X + Y) = 0$.