Homework 2 solutions: expectation, conditional distributions, functions of rv

EE 325 (DD): Probability and Random Processes, Autumn 2016

Instructor: Animesh Kumar, EE, IIT Bombay

Set-A

1. Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with an exponential distribution with parameter λ. Let $X_{(n)}$ and $X_{(1)}$ be the largest and the smallest random variable in the sequence $X_1, X_2, \ldots, X_n$. Find the marginal (cumulative) distribution functions of $X_{(n)}$ and $X_{(1)}$.

   An exponential random variable Y with parameter $\lambda > 0$ has the pdf

   $$f_Y(x) = \lambda e^{-\lambda x}, \quad x \geq 0. \tag{1}$$

   Solution: The random variables $X_1, X_2, \ldots, X_n$ have the same pdf

   $$f_X(x) = \lambda e^{-\lambda x}, x \geq 0.$$

   First note that

   $$\mathbb{P}(X_i \leq x) = \int_0^x f_{X_i}(\nu) \mathrm{d}\nu = 1 - e^{-\lambda x}$$

   for $x \geq 0$. Similarly, $\mathbb{P}(X_i > x) = e^{-\lambda x}$ for $x \geq 0$. To evaluate the cdf of $X_{(1)}$ and $X_{(n)}$, the i.i.d. property of the random variables will be used. The complement of cdf of $X_{(1)}$ is given by

   $$\begin{aligned}
   \mathbb{P}(X_{(1)} > x) &= \mathbb{P}(\min(X_1, X_2, \ldots, X_n) > x) \\
   &= \mathbb{P}(X_1 > x, X_2 > x, \ldots, X_n > x) \\
   &= \mathbb{P}(X_1 > x)\mathbb{P}(X_2 > x) \ldots \mathbb{P}(X_n > x) \text{ since } X_1, X_2, \ldots, X_n \text{ are i.i.d.,} \\
   &= \prod_{i=1}^n e^{-\lambda x} \\
   &= e^{-n\lambda x}.
   \end{aligned}$$

   Therefore, $\mathbb{P}(X_{(1)} \leq x) = 1 - e^{-nx}$ for $x \geq 0$. For $x < 0$, the cdf will be zero since all the random variables $X_1, \ldots, X_n$ are positive.

   For the cdf of $X_{(n)}$ a similar approach will be used. The cdf of $X_{(n)}$ is given by

   $$\begin{aligned}
   \mathbb{P}(X_{(n)} \leq x) &= \mathbb{P}(\max(X_1, X_2, \ldots, X_n) < x) \\
   &= \mathbb{P}(X_1 < x, X_2 < x, \ldots, X_n < x) \\
   &= \mathbb{P}(X_1 \leq x)\mathbb{P}(X_2 \leq x) \ldots \mathbb{P}(X_n \leq x) \text{ since } X_1, X_2, \ldots, X_n \text{ are i.i.d.,} \\
   &= \prod_{i=1}^n (1 - e^{-\lambda x}) \\
   &= (1 - e^{-\lambda x})^n.
   \end{aligned}$$

   This completes the solution.

2. Let $X \sim \text{Poisson}(\lambda)$ and $Y \sim \text{Poisson}(\mu)$, with $\lambda, \mu > 0$. Assume that X and Y are independent, and n is a non-negative integer.

   (a) Find the pmf of $Z = X + Y$.

(b) Find the contional distribution of Y conditioned on $Z = n$, i.e., the pmf $p_{Y|Z}(y|n)$.
(c) What is the conditional expectation $\mathbb{E}(Y|Z)$?

Solution:

(a) If $X \sim \text{Poisson}(\lambda)$ then X is a discrete random variable with the pmf

$$p_X(k) = \frac{e^{-\lambda}\lambda^k}{k!}.$$

Using discrete convolution, we can find the pmf of Z.

$$\begin{aligned} p_Z(n) = \sum_k p_X(k)p_Y(n-k) &= \sum_{k=0}^{n} \left(\frac{e^{-\lambda}\lambda^k}{k!}\right)\left(\frac{e^{-\mu}\mu^{n-k}}{(n-k)!}\right) \\ &= \frac{e^{-(\lambda+\mu)}}{n!}\sum_{k=0}^{n}\frac{n!\lambda^k\mu^{n-k}}{k!(n-k)!} \\ &= \frac{e^{-(\lambda+\mu)}(\lambda+\mu)^n}{n!}. \end{aligned}$$

Thus, Z is also a discrete random variable with a Poisson distribution and parameter $(\lambda + \mu)$.

(b) The conditional pmf is given by,

$$\begin{aligned} p_{Y|Z}(y|n) = \mathbb{P}(Y = y|Z = n) &= \frac{\mathbb{P}(Y = y, X = n-y)}{\mathbb{P}(Z = n)} \\ &= \frac{e^{-\mu}\mu^y}{y!}\frac{e^{-\lambda}\lambda^{(n-y)}}{(n-y)!}\frac{n!}{e^{-(\lambda+\mu)}(\lambda+\mu)^n} \\ &= \frac{n!}{(n-y)!y!}\left(\frac{\mu}{\lambda+\mu}\right)^y\left(\frac{\lambda}{\lambda+\mu}\right)^{n-y} \end{aligned}$$

Thus, $Y|(Z = n) \sim \text{Bin}(n, p)$ with $p = \mu/(\lambda + \mu)$.

(c) It has been showed that $Y|(Z = n) \sim \text{Bin}(n, p)$ with $p = \frac{\mu}{\mu+\lambda}$. The expectation of a $\text{Bin}(n, p)$ random variable is np. Thus, $\mathbb{E}(Y|Z = n) = n\mu/(\lambda + \mu)$.

3. Assume that X is a continuous random variable with,

$$f_X(x) = \frac{c}{1+|x|^4}, x \in \mathbb{R}.$$

The constant c is selected such that $\int_{\mathbb{R}} f_X(x)\mathrm{d}x = 1$. Find the values of $\mathbb{E}(X)$ and $\mathbb{E}(X^3)$.

Solution: Whenever expectation is well defined, for a continuous random variable, it will be given by $\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x)dx$. First note that $\mathbb{E}(X^+)$ is finite since

$$\mathbb{E}(X^+) = \int_0^\infty \frac{cx}{1+x^4}\mathrm{d}x < \infty.$$

Similarly, $\mathbb{E}(X^-) < \infty$. Therefore, $\mathbb{E}(X)$ exists. Finally,

$$\begin{aligned} \mathbb{E}[X] &= \int_{-\infty}^{\infty}\frac{cx}{1+|x|^4}dx \\ &= \int_{-\infty}^{0}\frac{cx}{1+|x|^4}dx + \int_0^\infty \frac{cx}{1+|x|^4}dx \\ &\overset{(a)}{=} -\int_0^\infty \frac{cx}{1+|x|^4}dx + \int_0^\infty \frac{cx}{1+|x|^4}dx \\ &= 0, \end{aligned}$$

where (a) follows by doing the substitution $x = -x$ in the first integral term (the integral from $-\infty$ to 0). Here, $\mathbb{E}[X]$ is 0, because $\mathbb{E}[X^+]$ converges, and is equal to $\mathbb{E}[X^-]$.

Let $Y = X^3$. Then,

$$\begin{aligned}
\mathbb{E}[Y^+] &= \int_0^\infty \frac{cx^3}{1+x^4}dx \\
&\overset{(b)}{=} \int_1^\infty \frac{cdt}{4t} \\
&\overset{(c)}{=} \frac{c}{4}\log(t)|_1^\infty \\
&= \infty
\end{aligned}$$

where (b) follows by substituting $t = 1 + x^4$, and (c) follows because integral of $\frac{1}{x}$ is $\log(x)$. Similarly, $\mathbb{E}[Y^-] = \infty$, and hence, $\mathbb{E}[X^3] = \mathbb{E}[Y]$ is not defined.

Set-B

1. Assume that $\mathbb{E}(X^2) < \infty$. Show that $\alpha = \mathbb{E}(X)$ is the unique value of α that minimizes $\mathbb{E}((X-\alpha)^2)$.

   Solution: The cost function $\mathbb{E}((X-\alpha)^2)$ can be expanded as a quadratic expression in α. Observe that $\mathbb{E}((X-\alpha)^2) = \mathbb{E}(X^2 + \alpha^2 - 2\alpha X) = \mathbb{E}(X^2) + \alpha^2 - 2\alpha\mathbb{E}(X)$. The quadratic expression can be minimized by taking derivatives.

   $$\frac{\mathrm{d}}{\mathrm{d}\alpha}[\mathbb{E}((X-\alpha)^2)] = 2\alpha - 2\mathbb{E}(X) \quad \text{and} \quad \frac{\mathrm{d}^2}{\mathrm{d}\alpha^2}[\mathbb{E}((X-\alpha)^2)] = 1 > 0$$

   Equating the first derivative to 0 gives the point of (unique) minima as $\alpha = \mathbb{E}(X)$.

2. Assume that $g(x)$ and $h(y)$ are measurable functions on the set of real numbers. If the random variables (X, Y) are independent, then show that $(g(X), h(Y))$ are also independent.

   Solution:

   It is given that $F_{X,Y}(x,y) = F_X(x)F_Y(y)$. We will break down the solution into three parts. First, we will show that $\mathbb{P}(X \in I_x, Y \in I_y) = \mathbb{P}(X \in I_x)\mathbb{P}(Y \in I_y)$, i.e., the events $\{X \in I_x\}$ and $\{Y \in I_y\}$ are independent. Second, we will show that for a union of disjoint intervals $I_x := I_{x_1}, I_{x_2,x_3}, \ldots, I_{x_n}$ and $I_y := I_{y_1}, I_{y_2,y_3}, \ldots, I_{y_m}$, the events $\{X \in I_x\}$ and $\{Y \in I_y\}$ are independent. Finally, since $g(x)$ and $h(y)$ are measurable functions, therefore, $\{g(X) \le u\}$ and $\{h(Y) \le v\}$ will correspond to independence of X and Y on the union of disjoint intervals.

   Step 1: First, assume that $x_1 < x_2$ and $y_1 < y_2$. Also assume that these variables are finite. Then,

   $$\begin{aligned}
   \mathbb{P}(X \in (x_1, x_2], Y \in (y_1, y_2]) &= F_{X,Y}(x_2,y_2) - F_{X,Y}(x_2,y_1) - F_{X,Y}(x_1,y_2) + F_{X,Y}(x_1,y_1) \\
   &= F_X(x_2)F_Y(y_2) - F_X(x_2)F_Y(y_1) - F_X(x_1)F_Y(y_2) + F_X(x_1)F_Y(y_1) \\
   &= (F_X(x_2) - F_X(x_1))(F_Y(y_2) - F_Y(y_1)) \\
   &= \mathbb{P}(X \in (x_1, x_2])\mathbb{P}(Y \in (y_1, y_2]).
   \end{aligned}$$

   It is left as an exercise for you to prove that the above equality is true for closed and open intervals on either side. Further, the equality also holds if either x_1 or x_2 are infinite. Similarly, either y_1 or y_2 can be infinte.

   Step 2: Denote $I_{x_1} = (-\infty, x_1]$ and $I_{x_n} = [x_n, \infty)$. Now assume that $I_x = I_{x_1} \cup I_{x_2,x_3}, \ldots, I_{x_n}$ and similarly $I_y = I_{y_1}, I_{y_2,y_3}, \ldots, I_{y_m}$. Then,

   $$\begin{aligned}
   \mathbb{P}(X \in I_x, Y \in I_y) &= \sum_{i,j} \mathbb{P}(X \in [x_i, x_{i+1}], Y \in [y_j, y_{j+1}]) \\
   &= \sum_{i,j} \mathbb{P}(X \in [x_i, x_{i+1}])\mathbb{P}(Y \in [y_j, y_{j+1}]) \\
   &= \mathbb{P}(X \in I_x)\mathbb{P}(Y \in I_y).
   \end{aligned}$$

Step 3: Finally, since $g(x)$ and $h(y)$ are measurable functions, therefore, $\{g(X) \leq u\}$ and $\{h(Y) \leq v\}$ will be equivalent to $X \in I_x, Y \in I_y$, where x_1^n and y_1^m will be functions of u and v respectively. Thus,

$$\begin{aligned}
\mathbb{P}(g(X) \leq u, h(Y) \leq v) &= \mathbb{P}(X \in I_{x(u)}, Y \in I_{y(v)}) \\
&= \mathbb{P}(X \in I_{x(u)})\mathbb{P}(Y \in I_{y(v)}) \\
&= \mathbb{P}(g(X) \leq u)\mathbb{P}(h(Y) \leq v)
\end{aligned}$$

Therefore, $g(X)$ and $h(Y)$ are independent. *(Note: this proof is not rigorous but it contains all the essential ideas.)*

3. Let $\{X_i\}_{i=1}^n$ be a sequence of i.i.d. continuous random variables with probability density function $f(x)$.

   (a) Find $\mathbb{P}(X_1 \leq X_2)$.

   (b) Find $\mathbb{P}(X_1 \leq X_2, X_1 \leq X_3)$.

   (c) Let N be a new integer-valued random variable defined as follows. N is the index of the first random variable that is less than X_1, that is,

   $$\mathbb{P}(N = n) = \mathbb{P}(X_1 \leq X_2, X_1 \leq X_3, \ldots, X_1 \leq X_{n-1}, X_1 > X_n). \tag{2}$$

   Find $\mathbb{P}(N > n)$ as a function of n.

   (d) Show that $\mathbb{E}(N) = \infty$

Solution:

(a) Using disjoint events $X_1 > X_2$, $X_1 = X_2$, and $X_1 < X_2$, we get $\mathbb{P}(X_1 > X_2) + \mathbb{P}(X_1 = X_2) + \mathbb{P}(X_1 < X_2) = 1$. Since X_1, X_2 are continuous, therefore $\mathbb{P}(X_1 = X_2) = 0$. Further X_1 and X_2 are i.i.d. Therefore, $\mathbb{P}(X_1 < X_2) = \mathbb{P}(X_2 < X_1)$. Hence $\mathbb{P}(X_1 \leq X_2) = 1/2$.

(b) Using a similar argument as in the previous part, there are six ways in which X_1, X_2, X_3 can be ordered (or arranged). The chance that any two or all three are equal is zero. Out of those six ways, X_1 is smallest in two ways. Finally, due to i.i.d. nature of random variables, all these orderings are equiprobable. Hence $\mathbb{P}(X_1 \leq X_2, X_1 \leq X_3) = 2/6 = 1/3$.

(c) We again utilize the fact that two or more random variables are equal with zero probability, and any ordering of these random variables is equiprobable. Fixing X_N in least and X_1 in least but one position, there are $(n-2)!$ ways in which $N = n$ can happen. Hence, $\mathbb{P}(N = n) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$. For $N > n$, Fixing X_1 in the least value position, now with remaining elements we can have $(n-1)!$ combinations.

$$\mathbb{P}(N > n) = \mathbb{P}(X_1 \leq X_2, X_1 \leq X_3, \ldots, X_1 \leq X_{n-1}, X_1 \leq X_n) = \frac{(n-1)!}{(n)!} = 1/n \tag{3}$$

(d) Using (3), we note that $\mathbb{E}(N) = \sum_{n=2}^{\infty} \mathbb{P}(N \geq n) = \sum_{k=1}^{\infty} \mathbb{P}(N > k)$. Since $\mathbb{P}(N > k)$ is decreasing as $1/k$, its summation leads to infinity.

4. Let X_1 and X_2 be IID Gaussian random variables with $X_i \sim \mathcal{N}(0, \sigma^2), i = 1, 2$. Let $Y = X_1 + X_2$. Then answer the following questions.

   (a) Find the distribution of Y by using 'functions of random variable' approach. You can use the convolution of pdf formula if it is required.

   (b) Find the conditional distribution of X_1 given Y. Interpret the result obtained. What will you expect the conditional distribution of X_2 given Y to be?

Solution:

(a) The pdf of Y will be a convolution of the pdfs of X_1 and X_2. Thus,

$$f_Y(y) = f_X(y) * f_X(y) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^2 \int_{\mathbb{R}} \exp\left(-\frac{x^2 + (y-x)^2}{2\sigma^2}\right) \mathrm{d}x$$

Upon simplifying the above expression we get,

$$
\begin{aligned}
f_Y(y) &= \left(\frac{1}{(\sqrt{2\pi})\sigma\sqrt{2}}\exp\left(-y^2/(4\sigma^2)\right)\right)\left(\frac{1}{\sqrt{2\pi}(\sigma/\sqrt{2})}\int_{\mathbb{R}}\exp\left(-\frac{(x-y/2)^2}{2(\sigma/\sqrt{2})^2}\right)\mathrm{d}x\right) \\
&= \frac{1}{(\sqrt{2\pi})\sigma\sqrt{2}}\exp\left(-y^2/(4\sigma^2)\right)
\end{aligned}
$$

Thus, $Y \sim= \mathcal{N}(0, 2\sigma^2)$.

(b) Since $f_{X_1}(x)$ and $f_Y(y)$ are non-zero everywhere, we can use the Baye's rule to find the conditional pdfs. In particular,

$$
\begin{aligned}
f_{X_1|Y}(x|y) &= \frac{f_{Y|X_1}(y|x)f_{X_1}(x)}{f_Y(y)} \\
&= \left(\frac{\exp(-x^2/(2\sigma^2)}{\sqrt{2\pi}\sigma}\right)\left(\frac{\exp(-(y-x)^2/(2\sigma^2))}{\sqrt{2\pi}\sigma}\right)\left(\frac{\sqrt{2\pi}\sqrt{2}\sigma}{\exp(-y^2/(4\sigma^2))}\right) \\
&= \left(\frac{\exp(-(x-y/2)^2/\sigma^2}{\sqrt{2\pi}\sigma/\sqrt{2}}\right).
\end{aligned}
$$

Thus, $X_1|(Y = y) \sim \mathcal{N}(y/2, \sigma^2/2)$.

5. Let X, Y be a continuous random variables having a cumulative distribution function $F(x, y)$. Let their marginal (cumulative) distributions be $G(x)$ and $H(y)$.

(a) Show that $G(X)$ is uniformly distributed in $(0, 1)$.

(b) Suppose that you have access to a random variable U uniformly distributed in $(0, 1)$ (for example, in MATLAB or C, you will have access to a uniform random variable). How would you use it to simulate a continuous random variable X having a distribution function G? Justify rigorously.

(c) Suppose now you have two IID random variables U_1 and U_2 distributed uniformly in $(0, 1)$. How would you use them to simulate random variable pair (X, Y) having a joint distribution $F(x, y)$?

Solution:

(a) First assume that $f_X(x)$ is non-zero on some support $I \subseteq \mathbb{R}$. Let $U = G(X)$. The random variable X takes values in I and $G(x)$ is strictly increasing on I. Thus $G^{-1}(x)$ exists for all $x \in I$. Thus, G is invertible on the support set I of the pdf of X.
Consider the cdf of U. For $0 \le u \le 1$, we have $F_U(u) = \mathbb{P}(U \le u) = \mathbb{P}(G(X) \le u) = \mathbb{P}(X \le G^{-1}(u)) = G(G^{-1}(u)) = u$, which is the cdf of a uniformly distributed random variable in $(0, 1)$. $G^{-1}(u)$ exists since $G(x)$ is monotonically increasing (cdf).[1] Finally, if $u < 0$, then $\mathbb{P}(U \le u) = 0$ and if $u \ge 1$, then $\mathbb{P}(U \le u) = 1$. Thus, $U \sim \text{Unif}(0, 1)$.

(b) Based on part (a) and invertibility of G, we expect $G^{-1}(U)$ to satisfy the requirements. $\mathbb{P}(G^{-1}(U) \le x) = \mathbb{P}(U \le G(x)) = G(x)$ for $x \in I$. Therefore, $X = G^{-1}(U)$ can be used to simulate a continuous random variable X with any arbitrary cdf $G(x)$ of a continuous random variable.

(c) One can use U_1 to generate X using G^{-1} and U_2 to generate Y using H^{-1}. However, since U_1 and U_2 are independent, so we will not obtain the dependence between X and Y using this procedure. To obtain dependence, we need to use the conditional cdf of Y given X. Let $G(x)$ and $F_{Y|X}(y|x)$ be the cdfs. Both of these are valid cdfs and hence they will be invertible. For each $x \in I$, let $F^{-1}_{Y|X}(y|x)$ be the inverse function. Then $X_1 = G^{-1}(U_1)$ will have the cdf $G(x)$. Further, depending on the realized value of X_1 from U_1, generate $Y_1 = F^{-1}_{Y|X}(U_2|X_1)$. Then (X_1, Y_1) have the same joint cdf as $F(x, y)$. To see this formally,

$$
\begin{aligned}
\mathbb{P}(x - \delta x < X_1 \le x, Y_1 \le y) &= (f_X(x)\delta x)\,\mathbb{P}(F^{-1}_{Y|X}(U_2|x) \le y) \\
&= (f_X(x)\delta x)\,\mathbb{P}(U_2 \le F_{Y|X}(y|x)) = (f_X(x)\delta x)F_{Y|X}(y|x).
\end{aligned}
$$

Integrating over x gives the result. We have conditioned on $X_1 = x$, and utilized the fact that U_2 is independent of X_1.

[1] One needs to show that if G^{-1} is also monotonic; it is a routine exercise in mathematical analysis. Use derivatives or see Walter Rudin's book.