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

EE 325: Probability and Random Processes, Autumn 2019
Instructor: Animesh Kumar, EE, IIT Bombay

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

$$f_X(x) = \frac{c}{1+|x|^6}, x \in \mathbb{R}.$$

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 $\mathbb{E}(X^5)$.

Solution: For a continuous rv X, its pdf is given by

$$f(x) = \frac{c}{1+|x|^6} \quad \text{for } x \in R$$

and it is known that $E[X] = E[X^+] - E[X^-]$. In this case,

$$\begin{aligned} E[X^+] &= \int_0^\infty x f(x)\mathrm{d}x \\ &= \int_0^\infty \frac{cx}{1+|x|^6}\mathrm{d}x \\ &= \int_0^\infty \frac{cx}{1+x^6}\mathrm{d}x \\ &\leq \int_0^1 \frac{cx}{1+x^6}\mathrm{d}x + \int_1^\infty \frac{cx}{x^6/2}\mathrm{d}x < \infty \end{aligned}$$

that is, $\mathbb{E}(X^+)$ exists. Similarly,

$$\begin{aligned} E[X^-] &= \int_0^\infty x f(-x) dx \\ &= \int_0^\infty \frac{cx}{1+|-x|^6} dx \\ &= E[X^+] \end{aligned}$$

and exists as well. This means, $E[X] = E[X^+] - E[X^-] = 0$. On the other hand, for $Y = X^5$,

$$\begin{aligned} E[X^{5+}] &= \int_0^\infty x^5 f(x) dx \\ &= \int_0^\infty \frac{cx^5}{1+|x|^6} dx \\ &= \int_0^\infty \frac{cx^5}{1+x^6} dx. \end{aligned}$$

With $1 + x^6 = t$. i.e. $6x^5 dx = dt$ we get,

$$\begin{aligned} E[X^{5+}] &= \int_1^\infty \frac{cdt}{6t} \\ &= \frac{c}{6}\Bigg[\log(t)\Bigg]_1^\infty \\ &= \infty \end{aligned}$$

Similarly, $E[X^{5-}] = \infty$. So, $E[X^5]$ is undefined.

2. 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 $\alpha = \mathbb{E}(X)$.

3. If the random variables (X, Y) are independent, then show that (X^2+X, Y^2+2Y) are also independent.

Solution: Let $U = X^2 + X$ and $V = Y^2 + 2Y$. Then note that $U \le u$ and $V \le v$ imply that

$$-\sqrt{u+\frac{1}{4}}-\frac{1}{2} \le X \le \sqrt{u+\frac{1}{4}}-\frac{1}{2} \text{ and } -\sqrt{v+1}-1 \le Y \le \sqrt{v+1}-1. \tag{1}$$

Since X and Y are independent, so the probability of the above joint event simplifies to (why!)

$$\begin{aligned}
\mathbb{P}(U \le u, V \le v) &= \mathbb{P}\left(-\sqrt{u+\frac{1}{4}}-\frac{1}{2} \le X \le \sqrt{u+\frac{1}{4}}-\frac{1}{2}, -\sqrt{v+1}-1 \le Y \le \sqrt{v+1}-1\right) && (2)\\
&= \mathbb{P}\left(-\sqrt{u+\frac{1}{4}}-\frac{1}{2} \le X \le \sqrt{u+\frac{1}{4}}-\frac{1}{2}\right)\mathbb{P}\left(-\sqrt{v+1}-1 \le Y \le \sqrt{v+1}-1\right) && (3)\\
&= \mathbb{P}(U \le u)\mathbb{P}(V \le v) && (4)
\end{aligned}$$

for any $v \ge -1$ and $u \ge -1/4$. For $u < -1/4$ or $v < -1$, the joint and marginal CDFs (both) are zero. So, the random variables U and V are independent.

4. 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 conditional distribution of Y conditioned on $Z = n$, i.e., the pmf $p_{Y|Z}(y|n)$.

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 $p = \mu/(\lambda+\mu)$.

5. 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,

$$\mathbb{P}(N = n) = \mathbb{P}(X_1 \leq X_2, X_1 \leq X_3, \ldots, X_1 \leq X_{n-1}, X_1 > X_n). \tag{5}$$

   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.

$$\mathbb{P}(N > n) = \mathbb{P}(X_1 \leq X_2, X_1 \leq X_3, \ldots, X_1 \leq X_{n-1}, X_1 \leq X_n) = \frac{(n-1)!}{(n)!} = 1/n \tag{6}$$

   (d) Using (5), 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.

6. 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 $Y = y$. Interpret the result obtained. What will you expect the conditional distribution of X_2 given $Y = y$ 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, $X_1|(Y=y) \sim \mathcal{N}(y/2, \sigma^2/2)$.

7. 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.