Exercise Walkthrough: Conditional Expectation with Joint PDFs

Justin Lanfermann

25. June of 2025

1 Problem Statement

Let $(\Omega, \mathcal{A}, P)$ be a probability space, and let $X, Y : \Omega \to \mathbb{R}$ be two real-valued random variables (RVs). Their joint probability density function (PDF) [1] is given by:

$$p_{X,Y}(x, y) = \exp(-y) \cdot \chi_A(x, y)$$

where the set A is defined as $A = \{(x, y) \in \mathbb{R}^2 \mid 0 \le x \le y\}$. This can also be written as:

$$p_{X,Y}(x, y) = \begin{cases} e^{-y} & \text{if } 0 \le x \le y \\ 0 & \text{otherwise} \end{cases}$$

We are tasked with the following:

(i) Verify that $p_{X,Y}(x, y)$ is a valid PDF.

(ii) Compute the conditional expectation $\mathbb{E}[X \mid Y]$.

(iii) Compute the conditional expectation $\mathbb{E}[Y \mid X]$.

---

2 Step-by-Step Solution

2.1 (i) Verifying the PDF

Overview To verify that $p_{X,Y}(x, y)$ is a valid joint PDF, we need to check two conditions based on **Definition 1.39 (probability density function)** from the script:

1. **Non-negativity:** The function must be non-negative everywhere, i.e., $p_{X,Y}(x, y) \ge 0$ for all $(x, y) \in \mathbb{R}^2$.
2. **Integration to one:** The integral of the function over the entire plane must be equal to 1, i.e., $\iint_{\mathbb{R}^2} p_{X,Y}(x, y)\, dx\, dy = 1$.

Step 1: Checking Non-negativity The exponential function e^{-y} is always positive for any real y. The indicator function [2] $\chi_A(x, y)$ is either 1 (if $(x, y) \in A$) or 0 (otherwise). Therefore, their product $e^{-y}\chi_A(x, y)$ is always greater than or equal to 0. The first condition is met.

Step 2: Integration to One Now we need to compute the double integral over $\mathbb{R}^2$. The indicator function $\chi_A(x,y)$ makes the integrand non-zero only over the region A, so we can restrict our integration bounds to this region. The region A is defined by $0 \le x \le y$. This implies that for any y, x ranges from 0 to y. It also means y must be non-negative (since $x \ge 0$). We can set up the integral in two ways. Let's follow the order used in the provided solution to see how it works. The condition $0 \le x \le y$ is equivalent to $y \ge x$ for $x \ge 0$. So we integrate with respect to y first, from x to ∞, and then with respect to x from 0 to ∞.

$$\begin{aligned}
\iint_{\mathbb{R}^2} p_{X,Y}(x,y)\,dy\,dx &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-y}\chi_A(x,y)\,dy\,dx \\
&= \int_0^{\infty}\left(\int_x^{\infty} e^{-y}\,dy\right)dx \quad (\text{Bounds from } 0 \le x \le y) \\
&= \int_0^{\infty}\left[-e^{-y}\right]_{y=x}^{y=\infty} dx \quad (\text{Inner integral w.r.t. } y) \\
&= \int_0^{\infty}\left(-\lim_{y\to\infty} e^{-y} - (-e^{-x})\right) dx \\
&= \int_0^{\infty}(0 + e^{-x})\,dx \\
&= \int_0^{\infty} e^{-x}\,dx \quad (\text{Outer integral w.r.t. } x) \\
&= \left[-e^{-x}\right]_{x=0}^{x=\infty} \\
&= (-\lim_{x\to\infty} e^{-x}) - (-e^{-0}) \\
&= (0) - (-1) = 1.
\end{aligned}$$

Since both conditions are met, $p_{X,Y}(x,y)$ is a valid PDF.

2.2 (ii) Computing E[X | Y]

Overview To find the conditional expectation $\mathbb{E}[X \mid Y]$, we first need to find the conditional PDF [3] $p_{X|Y=y}(x)$. This is given by the formula:

$$p_{X|Y=y}(x) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$$

where $p_Y(y)$ is the marginal PDF [4] of Y. Once we have the conditional PDF, the conditional expectation is computed as:

$$\mathbb{E}[X \mid Y = y] = \int_{-\infty}^{\infty} x \cdot p_{X|Y=y}(x)\,dx$$

This entire process is laid out in **Definition 2.28 (conditional expectation)**.

Step 1: Compute the Marginal PDF of Y, $p_Y(y)$ We find $p_Y(y)$ by "integrating out" the variable x from the joint PDF, as described in **Theorem 1.63 (iii)**.

$$\begin{aligned}
p_Y(y) &= \int_{-\infty}^{\infty} p_{X,Y}(x,y)\,dx \\
&= \int_{-\infty}^{\infty} e^{-y}\chi_A(x,y)\,dx \\
&= e^{-y}\int_0^y 1\,dx \quad (\text{for } y \ge 0,\ \text{otherwise } 0) \\
&= e^{-y}\,[x]_{x=0}^{x=y} \\
&= ye^{-y} \quad \text{for } y \ge 0.
\end{aligned}$$

So, the marginal PDF is $p_Y(y) = ye^{-y}\chi_{\mathbb{R}_{\geq 0}}(y)$. This is a Gamma distribution $\Gamma(2, 1)$.

Step 2: Compute the Conditional PDF, $p_{X|Y=y}(x)$ Now we can find the conditional PDF.

$$p_{X|Y=y}(x) = \frac{p_{X,Y}(x,y)}{p_Y(y)} = \frac{e^{-y}\chi_{\{0\leq x\leq y\}}}{ye^{-y}\chi_{\{y\geq 0\}}} = \frac{1}{y}\chi_{[0,y]}(x) \quad \text{for } y > 0.$$

This is the PDF of a uniform distribution on the interval $[0, y]$, i.e., $X \mid Y = y \sim \text{Unif}(0, y)$ (see **Example 1.56 (i)**).

Step 3: Compute the Conditional Expectation, $\mathbb{E}[X \mid Y = y]$ The expectation of a uniform distribution $\text{Unif}(a, b)$ is simply $(a + b)/2$. For $X \mid Y = y \sim \text{Unif}(0, y)$, the expectation is $(0 + y)/2 = y/2$. We can also compute this explicitly:

$$\begin{aligned}
\mathbb{E}[X \mid Y = y] &= \int_{-\infty}^{\infty} x \cdot p_{X|Y=y}(x)\,dx \\
&= \int_0^y x \cdot \frac{1}{y}\,dx \\
&= \frac{1}{y}\left[\frac{x^2}{2}\right]_{x=0}^{x=y} \\
&= \frac{1}{y}\left(\frac{y^2}{2} - 0\right) = \frac{y}{2}.
\end{aligned}$$

The conditional expectation is a function of the value of the conditioning variable. So we write the final answer as a random variable: $\mathbb{E}[X \mid Y] = \frac{Y}{2}$.

2.3 (iii) Computing E[Y | X]

Overview The procedure is symmetric to part (ii). We first find the marginal PDF of X, $p_X(x)$, then the conditional PDF $p_{Y|X=x}(y)$, and finally compute the expectation $\mathbb{E}[Y \mid X = x]$.

Step 1: Compute the Marginal PDF of X, $p_X(x)$ We integrate out y from the joint PDF. The integration is over $y \geq x$ for a fixed $x \geq 0$.

$$\begin{aligned}
p_X(x) &= \int_{-\infty}^{\infty} p_{X,Y}(x,y)\,dy \\
&= \int_x^{\infty} e^{-y}\,dy \quad (\text{for } x \geq 0,\ \text{otherwise } 0) \\
&= \left[-e^{-y}\right]_{y=x}^{y=\infty} \\
&= (0) - (-e^{-x}) = e^{-x} \quad \text{for } x \geq 0.
\end{aligned}$$

So, $p_X(x) = e^{-x}\chi_{\mathbb{R}_{\geq 0}}(x)$. This is the exponential distribution $\text{Exp}(1)$ (see **Example 1.56 (iv)**).

Step 2: Compute the Conditional PDF, $p_{Y|X=x}(y)$ For $x \geq 0$:

$$p_{Y|X=x}(y) = \frac{p_{X,Y}(x,y)}{p_X(x)} = \frac{e^{-y}\chi_{\{y\geq x\}}}{e^{-x}\chi_{\{x\geq 0\}}} = e^{x-y}\chi_{[x,\infty)}(y).$$

This is a "shifted" exponential distribution.

Step 3: Compute the Conditional Expectation, $\mathbb{E}[Y \mid X = x]$ For $x \geq 0$, we compute:

$$\mathbb{E}[Y \mid X = x] = \int_{-\infty}^{\infty} y \cdot p_{Y|X=x}(y)\, dy = \int_{x}^{\infty} y \cdot e^{x-y}\, dy = e^{x} \int_{x}^{\infty} y e^{-y}\, dy.$$

This integral requires integration by parts [5] ($\int u\, dv = uv - \int v\, du$). Let $u = y$ and $dv = e^{-y}\, dy$. Then $du = dy$ and $v = -e^{-y}$.

$$\begin{aligned}
\int_{x}^{\infty} y e^{-y}\, dy &= \left[y(-e^{-y})\right]_{x}^{\infty} - \int_{x}^{\infty} (-e^{-y})\, dy \\
&= \left(\lim_{y \to \infty} -y e^{-y} - (-x e^{-x})\right) + \int_{x}^{\infty} e^{-y}\, dy \\
&= (0 + x e^{-x}) + \left[-e^{-y}\right]_{x}^{\infty} \\
&= x e^{-x} + (0 - (-e^{-x})) \\
&= x e^{-x} + e^{-x} = (x+1)e^{-x}.
\end{aligned}$$

Now, we substitute this back into our expression for the expectation:

$$\mathbb{E}[Y \mid X = x] = e^{x} \cdot \left((x+1)e^{-x}\right) = x + 1.$$

This holds for $x \geq 0$. As a random variable, the result is $\mathbb{E}[Y \mid X] = X + 1$.

3 Summary of Takeaways

- **PDF Validation:** Always check non-negativity and that the integral over the entire domain equals 1. The support of the PDF (where it's non-zero) is crucial for setting up the correct integration limits.
- **Marginalization:** To find the distribution of a single variable from a joint distribution, you "integrate out" the other variable(s). This is a fundamental operation.
- **Conditioning:** The conditional PDF $p_{Y|X}$ tells you the distribution of Y given that X has taken a specific value. It is found by dividing the joint PDF by the marginal PDF of the conditioning variable.
- **Conditional Expectation:** $\mathbb{E}[Y \mid X]$ is the expected value of Y computed using the conditional distribution $p_{Y|X}$. It is not a constant but a random variable itself, as its value depends on the value of X.

---

4 In-depth Explanations

[1] Probability Density Function (PDF) A function $p : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ is a PDF if it's non-negative and its integral over the entire space $\mathbb{R}^n$ is 1. For a continuous random variable X, the probability of X falling into a set A is given by integrating the PDF over that set: $P(X \in A) = \int_A p(x)dx$. Crucially, for any single point c, $P(X = c) = 0$, and the value $p(c)$ is not a probability itself, but a measure of probability *density*. *(Script Reference: Definition 1.39)*

[2] Indicator Function (χ_A) The indicator function of a set A, denoted $\chi_A(x)$, is a simple but powerful tool. It is defined as:

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

In probability, it's used to define PDFs that are non-zero only on a specific region (their support), simplifying the notation for integration bounds. *(Script Reference: Used in Proposition 1.42)*

[3] Conditional PDF and Expectation Given two continuous RVs X and Y with joint PDF $p_{X,Y}(x, y)$, the conditional PDF of Y given $X = x$ is defined as $p_{Y|X=x}(y) = \frac{p_{X,Y}(x,y)}{p_X(x)}$, provided $p_X(x) > 0$. It describes the probability distribution of Y when we know the value of X. The conditional expectation $\mathbb{E}[Y \mid X = x]$ is the mean of this conditional distribution. *(Script Reference: Definitions 1.73, 2.28)*

[4] Marginal PDF The marginal PDF of a single variable is obtained from a joint PDF by integrating over all possible values of the other variables. For a joint PDF $p_{X,Y}(x, y)$, the marginal PDF of X is $p_X(x) = \int_{-\infty}^{\infty} p_{X,Y}(x, y)\, dy$. It represents the distribution of X irrespective of the value of Y. *(Script Reference: Theorem 1.63)*

[5] Integration by Parts This is a fundamental calculus technique used to integrate the product of two functions. The formula stems from the product rule for differentiation and is given by:

$$\int u\,dv = uv - \int v\,du$$

Choosing the functions for u (which will be differentiated) and dv (which will be integrated) is the key to successfully applying this method. A common mnemonic for choosing u is LIATE (Logarithmic, Inverse trig, Algebraic, Trigonometric, Exponential).