Exercise Walkthrough: Conditional PDF and Probability

Justin Lanfermann

25. June 2025

Abstract

This document provides a step-by-step walkthrough for an exercise on continuous random variables. We are given a joint probability density function (PDF) for two random variables, X and Y. Our goal is to first derive the conditional PDF of X given $Y = y$, and then to compute the conditional probability $P(X < 1/2 | Y = y)$. We will meticulously follow the formal definitions and theorems as laid out in the "Discrete Probability Theory" script by Niki Kilbertus.

1 Problem Statement

Let X and Y be two jointly continuous random variables with the joint PDF [1] given by:

$$p_{X,Y}(x,y) = \begin{cases} \frac{x^2}{4} + \frac{y^2}{4} + \frac{xy}{6} & \text{if } x \in [0,1], y \in [0,2] \\ 0 & \text{otherwise.} \end{cases}$$

For a given $y \in [0,2]$, we need to find:

(i) The conditional PDF of X given $Y = y$, which is denoted as $p_{X|Y}(x|y)$.

(ii) The conditional probability $P\left(X < \frac{1}{2} \,\middle|\, Y = y\right)$.

2 Solution Walkthrough

We will solve this problem in two parts, following the structure of the question.

2.1 Part (i): Finding the Conditional PDF

Step 1: Recall the formula for conditional PDF. The foundation for this part is the definition of a conditional PDF. According to **Definition 1.73 (ii)** from the script, the conditional PDF of a random variable X given $Y = y$ is defined as the joint PDF divided by the marginal PDF of the conditioning variable.

$$p_{X|Y}(x|y) = \frac{p_{X,Y}(x,y)}{p_Y(y)} \tag{1}$$

This formula is valid for values of y where the marginal PDF $p_Y(y)$ is greater than zero. Our first task, therefore, is to compute this marginal PDF [2].

Step 2: Calculate the marginal PDF of Y. To find the marginal PDF $p_Y(y)$, we need to "integrate out" the other variable, X, from the joint PDF. This process is described in **Theorem 1.63 (iii)** of the script. The integration is performed over the entire support of X, which is $[0, 1]$ in this case.

$$p_Y(y) = \int_{-\infty}^{\infty} p_{X,Y}(x, y)\, dx$$

For a specific $y \in [0, 2]$, we have:

$$p_Y(y) = \int_0^1 \left(\frac{x^2}{4} + \frac{y^2}{4} + \frac{xy}{6} \right) dx$$

We compute the integral with respect to x, treating y as a constant:

$$= \left[\frac{x^3}{12} + \frac{y^2}{4}x + \frac{x^2 y}{12} \right]_{x=0}^{x=1}$$

Now, we evaluate the antiderivative at the limits of integration:

$$= \left(\frac{1^3}{12} + \frac{y^2}{4}(1) + \frac{1^2 y}{12} \right) - \left(\frac{0^3}{12} + \frac{y^2}{4}(0) + \frac{0^2 y}{12} \right)$$

$$= \frac{1}{12} + \frac{y^2}{4} + \frac{y}{12}$$

To simplify, we find a common denominator (12):

$$= \frac{1}{12} + \frac{3y^2}{12} + \frac{y}{12} = \frac{3y^2 + y + 1}{12}$$

So, the marginal PDF for Y is $p_Y(y) = \frac{3y^2+y+1}{12}$ for $y \in [0, 2]$.

Step 3: Assemble the conditional PDF. Now we have both components needed for Equation 1: the joint PDF $p_{X,Y}(x, y)$ and the marginal PDF $p_Y(y)$. We can substitute them into the formula. For $x \in [0, 1]$ and $y \in [0, 2]$:

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

$$= \frac{\frac{x^2}{4} + \frac{y^2}{4} + \frac{xy}{6}}{\frac{3y^2+y+1}{12}}$$

To simplify this fraction, we can multiply the numerator and the denominator by 12:

$$= \frac{12 \left(\frac{x^2}{4} + \frac{y^2}{4} + \frac{xy}{6} \right)}{12 \left(\frac{3y^2+y+1}{12} \right)}$$

$$= \frac{3x^2 + 3y^2 + 2xy}{3y^2 + y + 1}$$

Thus, the complete conditional PDF [3] of X given $Y = y$ is:

$$p_{X|Y}(x|y) = \begin{cases} \frac{3x^2+3y^2+2xy}{3y^2+y+1} & \text{if } x \in [0, 1] \\ 0 & \text{otherwise.} \end{cases}$$

2.2 Part (ii): Finding the Conditional Probability

Step 1: Set up the integral. To find the probability of an event for a continuous random variable, we integrate its PDF over the interval corresponding to that event. Since we are asked for a *conditional* probability, we must use the *conditional* PDF that we just derived in Part (i). This is the standard method for finding a probability from a PDF [4], applied to the conditional case.

The event is $X < \frac{1}{2}$, which corresponds to the interval $[0, 1/2)$ for x.

$$P\left(X < \frac{1}{2} \,\middle|\, Y = y\right) = \int_0^{1/2} p_{X|Y}(x|y)\, dx$$

Step 2: Calculate the integral. We substitute the expression for $p_{X|Y}(x|y)$ and compute the integral. Note that for a fixed y, the denominator of the conditional PDF is a constant with respect to x.

$$P\left(X < \frac{1}{2} \,\middle|\, Y = y\right) = \int_0^{1/2} \frac{3x^2 + 3y^2 + 2xy}{3y^2 + y + 1}\, dx$$
$$= \frac{1}{3y^2 + y + 1} \int_0^{1/2} (3x^2 + 3y^2 + 2xy)\, dx$$

We find the antiderivative of the integrand:

$$= \frac{1}{3y^2 + y + 1} \left[x^3 + 3y^2 x + x^2 y \right]_{x=0}^{x=1/2}$$

Now, we evaluate at the limits:

$$= \frac{1}{3y^2 + y + 1} \left(\left(\left(\frac{1}{2}\right)^3 + 3y^2\left(\frac{1}{2}\right) + \left(\frac{1}{2}\right)^2 y \right) - (0) \right)$$
$$= \frac{1}{3y^2 + y + 1} \left(\frac{1}{8} + \frac{3y^2}{2} + \frac{y}{4} \right)$$
$$= \frac{\frac{1}{8} + \frac{y}{4} + \frac{3y^2}{2}}{3y^2 + y + 1}$$

This is our final answer. It is a function of y, which makes sense: the probability of X being less than $1/2$ depends on the specific value that Y takes.

3 Summary and Key Takeaways

This exercise demonstrates a fundamental procedure in probability theory:

1. **Marginalization:** We first computed the marginal PDF $p_Y(y)$ by integrating the joint PDF $p_{X,Y}(x, y)$ over all possible values of x.

2. **Conditioning:** We then found the conditional PDF $p_{X|Y}(x|y)$ by dividing the joint PDF by the marginal PDF. This can be intuitively understood as restricting our universe to the "slice" where $Y = y$ and renormalizing the probabilities to sum to one.

3. **Probability Calculation:** Finally, we computed the desired conditional probability by integrating the conditional PDF over the specified range for x.

This process of marginalize $\rightarrow$ condition $\rightarrow$ compute is a cornerstone for working with joint distributions.

In-depth Explanations

[1] Joint Probability Density Function (PDF)

A joint PDF, like the $p_{X,Y}(x,y)$ given in this problem, is a function used to describe the probability distribution of two or more continuous random variables. It doesn't give probabilities directly. Instead, the probability that the pair (X, Y) falls into a certain region in the xy-plane is found by integrating the PDF over that region. A valid PDF must satisfy two main properties (see **Definition 1.39/1.40**):

1. **Non-negativity:** $p_{X,Y}(x,y) \geq 0$ for all x, y.

2. **Normalization:** The total integral over the entire plane must equal 1. That is, $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_{X,Y}(x,y) \, dx \, dy = 1$.

Think of it as a "probability density" surface. The volume under this surface over a certain area gives the probability of the random variables taking values in that area.

[2] Marginal PDF

The marginal PDF of one variable (say, Y) from a joint distribution represents the probability distribution of just that variable, irrespective of the other. As outlined in **Theorem 1.63 (iii)**, you obtain it by "averaging out" the influence of the other variable, which in the continuous case means integrating it out.

$$p_Y(y) = \int_{-\infty}^{\infty} p_{X,Y}(x,y) \, dx$$

An intuitive analogy comes from discrete probability tables. If you have a joint probability table for two variables, you can find the probability of one variable taking a specific value by summing across the rows (or columns) for the other variable. These sums are often written in the "margins" of the table, hence the name "marginal probability". Integration is the continuous analog of this summation.

[3] Conditional PDF

The conditional PDF $p_{X|Y}(x|y)$, as defined in **Definition 1.73 (ii)**, describes the probability distribution of a random variable X *given that* the random variable Y has taken a specific value y. Intuitively, imagine the 3D surface of the joint PDF $p_{X,Y}(x,y)$. When we "condition" on $Y = y$, we are essentially taking a 2D "slice" of this surface at that specific y value. This slice shows how the probability density for X is distributed, knowing that $Y = y$. However, this slice doesn't usually integrate to 1 on its own. The division by the marginal PDF $p_Y(y)$ is a normalization step. It scales the slice up or down so that it becomes a valid PDF for X (i.e., it integrates to 1), representing the updated state of knowledge about X.

[4] Probability from a PDF

For any continuous random variable, whether its distribution is described by a regular PDF like $p_X(x)$ or a conditional PDF like $p_{X|Y}(x|y)$, the way to calculate the probability of it falling within an interval $[a, b]$ is always the same: you integrate the corresponding PDF over that interval. As seen in **Example 1.58**,

$$P(a \leq X \leq b) = \int_a^b p_X(x) \, dx$$

4

In our exercise, we simply applied this principle to the conditional PDF to find a conditional probability:

$$P\left(a \leq X \leq b \,|\, Y = y\right) = \int_a^b p_{X|Y}(x|y)\, dx$$