

5.3: Conditional Probability Distributions

In this section, we consider the probability distribution of one random variable *given* information about the value of another random variable. As we will see in the formal definition, this kind of conditional distribution will involve the joint distribution of the two random variables under consideration, which we introduced in the previous two sections. We begin with discrete random variables, and then consider the continuous case.

Conditional Distributions of Discrete Random Variables

Recall the definition of conditional probability for *events* ([Definition 2.2.1](#)): the conditional probability of A given B is equal to

$$P(A | B) = \frac{P(A \cap B)}{P(B)}.$$

We use this same concept for events to define conditional probabilities for *random variables*.

Definition 5.3.1

If X and Y are discrete random variables with joint pmf given by $p(x, y)$, then the **conditional probability mass function of X , given that $Y = y$** , is denoted $p_{X|Y}(x|y)$ and given by

$$p_{X|Y}(x|y) = \frac{P(\{X = x\} \cap \{Y = y\})}{P(Y = y)} = \frac{p(x, y)}{p_Y(y)}, \quad \text{provided that } p_Y(y) > 0.$$

Note that if $p_Y(y) = 0$, then for that value of Y the conditional pmf of X does not exist.

Similarly, the **conditional probability mass function of Y , given that $X = x$** , is denoted $p_{Y|X}(y|x)$ and given by

$$p_{Y|X}(y|x) = \frac{P(\{Y = y\} \cap \{X = x\})}{P(X = x)} = \frac{p(x, y)}{p_X(x)}, \quad \text{provided that } p_X(x) > 0.$$

[Link to Video: Overview of Definition 5.3.1](#)

Example 5.3.1

For an example of conditional distributions for discrete random variables, we return to the context of [Example 5.1.1](#), where the underlying probability experiment was to flip a fair coin three times, and the random variable X denoted the number of heads obtained and the random variable Y denoted the winnings when betting on the placement of the first heads obtained in the three flips. We found the joint pmf for X and Y in [Table 1](#) of Section 5.1, and the marginal pmf's are given in [Table 2](#). We now find the conditional distributions of X and Y .

First, to find the conditional distribution of X given a value of Y , we can think of fixing a row in [Table 1](#) and dividing the values of the joint pmf in that row by the marginal pmf of Y for the corresponding value. For example, to find $p_{X|Y}(x|1)$, we divide each entry in the $Y = 1$ row by $p_Y(1) = 1/2$. Doing this for each row in [Table 1](#), results in the conditional distributions of X given a value of Y , which we represent in the following table.

Conditional pmf's of X given Y

x	$p_{X Y}(x -1)$	$p_{X Y}(x 1)$	$p_{X Y}(x 2)$	$p_{X Y}(x 3)$
0	1	0	0	0
1	0	1/4	1/2	1
2	0	1/2	1/2	0

x	$p_{X Y}(x -1)$	$p_{X Y}(x 1)$	$p_{X Y}(x 2)$	$p_{X Y}(x 3)$
3	0	1/4	0	0

Note that every column in the above table sums to 1. This is because each row is a different probability mass function for X given a value of Y . Specifically, if we look at the column for the conditional distribution of X given that $Y = 1$, $p_{X|Y}(x|1)$, this is the distribution of probability for the number of heads obtained, knowing that the winnings of the game are \$1. Recall that we win \$1 if the first heads is on the first toss. In that case, we know the outcome of the probability experiment must be one of $\{hhh, hht, hth, htt\}$. Working with this reduced sample space, we can see how the corresponding probabilities for the values of X arise.

Similar to the process for the conditional pmf's of X given Y , we can find the conditional pmf's of Y given X . This time, though, we take a column in Table 1, giving the joint pmf for a fixed value of X , and divide by the marginal pmf of X for the corresponding value. The following table gives the results.

Conditional pmf's of Y given X

y	$p_{Y X}(y 0)$	$p_{Y X}(y 1)$	$p_{Y X}(y 2)$	$p_{Y X}(y 3)$
-1	1	0	0	0
1	0	1/3	2/3	1
2	0	1/3	1/3	0
3	0	1/3	0	0

Note again that every column in the above table sums to 1. A conditional pmf is a pmf, just found in a specific way.

[Link to Video: Walkthrough of Example 5.3.1](#)

Informally, we can think of a conditional probability distribution as a probability distribution for a *sub-population*. In other words, a conditional probability distribution describes the probability that a randomly selected person from a sub-population has a given characteristic of interest. In this context, the joint probability distribution is the probability that a randomly selected person from the *entire* population has *both* characteristics of interest. The following example demonstrates these interpretations in a specific context.

Example 5.3.2

Suppose we are interested in the relationship between an individual's hair and eye color. Based on a random sample of Saint Mary's students, we have the following joint pmf, with marginal pmf's given in the margins:

$p(x, y)$	Hair Color (X)				
Eye Color (Y)	blonde (1)	red (2)	brown (3)	black (4)	$p_Y(y)$
blue (1)	0.12	0.05	0.12	0.01	0.30
green (2)	0.12	0.07	0.09	0	0.28
brown (3)	0.16	0.07	0.16	0.03	0.42
$p_X(x)$	0.40	0.19	0.37	0.04	1.00

The probabilities in the last row and column (orange cells) give the marginal pmf's for X and Y , while the probabilities in the interior (white and grey cells) give the joint pmf for pairs (X, Y) . For instance, $p(3, 2) = 0.09$ indicates the joint probability that a randomly selected SMC student has brown hair ($X = 3$) and green eyes ($Y = 2$) is 9%, $p_X(3) = 0.37$ indicates the marginal probability that a randomly selected SMC student has brown hair is 37%, and $p_Y(2) = 0.28$ indicates the marginal probability that a randomly selected SMC student has green eyes is 28%.

Given this table of probabilities, we can calculate conditional pmf values:

- First, let's find $p_{X|Y}(2|1)$:

$$p_{X|Y}(2|1) = \frac{p(2,1)}{p_Y(1)} = \frac{0.05}{0.3} = \frac{1}{6} \approx 0.167$$

Note that we can write $p_{X|Y}(2|1) = P(X=2 | Y=1) = P(\text{red hair} | \text{blue eyes})$. Here we are finding the probability that an individual in the sub-population of individuals with blue eyes has red hair. Specifically, we found that 1/6 (or approximately 16.7%) of SMC students with blue eyes have red hair.

- Now, let's reverse the order of X and Y , and find $p_{Y|X}(2|1)$:

$$p_{Y|X}(2|1) = \frac{p(1,2)}{p_X(1)} = \frac{0.12}{0.40} = 0.3$$

Now the sub-population is individuals with blonde hair, and we find the probability that an individual in this sub-population has green eyes. Specifically, we found that 30% of SMC students with blonde hair have green eyes.

[Link to Video: Interpretation of Joint & Conditional Probabilities \(Example 5.3.2\)](#)

Properties of Conditional PMF's

1. Conditional pmf's are valid pmf's. In other words, the conditional pmf for X , given $Y = y$, for a fixed y , is a valid pmf satisfying the following:

$$0 \leq p_{X|Y}(x|y) \leq 1 \quad \text{and} \quad \sum_x p_{X|Y}(x|y) = 1$$

Similarly, for a fixed x , we also have the following for the conditional pmf of Y , given $X = x$:

$$0 \leq p_{Y|X}(y|x) \leq 1 \quad \text{and} \quad \sum_y p_{Y|X}(y|x) = 1$$

2. In general, the conditional distribution of X given Y does not equal the conditional distribution of Y given X , i.e.,

$$p_{X|Y}(x|y) \neq p_{Y|X}(y|x).$$

3. If X and Y are independent, discrete random variables, then the following are true:

$$p_{X|Y}(x|y) = p_X(x)$$

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

In other words, if X and Y are independent, then knowing the value of one random variable does not affect the probability of the other one.

Now that we have defined conditional distributions, we define conditional expectation.

Definition 5.3.2

For discrete random variables X and Y , the **conditional expected value of X , given $Y = y$** , is given by

$$\mu_{X|Y=y} = E[X|Y = y] = \sum_x x p_{X|Y}(x|y),$$

and the **conditional expected value of Y , given $X = x$** , is given by

$$\mu_{Y|X=x} = E[Y|X = x] = \sum_y y p_{Y|X}(y|x).$$

Similarly, we can define conditional variances. The **conditional variance of X , given $Y = y$** , is given by

$$\begin{aligned} \sigma_{X|Y=y}^2 &= \text{Var}(X|Y = y) = E[(X - \mu_{X|Y=y})^2 | Y = y] = \sum_x (x - \mu_{X|Y=y})^2 p_{X|Y}(x|y) \\ &= E[X^2 | Y = y] - \mu_{X|Y=y}^2 = \left(\sum_x x^2 p_{X|Y}(x|y) \right) - \mu_{X|Y=y}^2 \end{aligned}$$

Example 5.3.3

Continuing in the context of [Example 5.3.2](#), we calculate the conditional mean and variance of hair color (X), given that a randomly selected student has brown eyes ($Y = 3$).

First, we derive the conditional pmf of X , given $Y = 3$, by taking the row for brown eyes in the joint pmf table and dividing each by the marginal pmf for Y at 3, i.e., $p_Y(3) = 0.42$. This gives the following:

x	$p_{X Y}(x 3) = p(x, 3)/p_Y(3)$
1	$0.16/0.42 = 8/21$
2	$0.07/0.42 = 1/6$
3	$0.16/0.42 = 8/21$
4	$0.03/0.42 = 1/14$

Using the conditional probabilities in the table above, we calculate the following:

$$\begin{aligned} E[X|Y = 3] &= \sum_{x=1}^4 x \cdot p_{X|Y}(x|3) = (1) \left(\frac{8}{21} \right) + (2) \left(\frac{1}{6} \right) + (3) \left(\frac{8}{21} \right) + (4) \left(\frac{1}{14} \right) = \frac{15}{7} \approx 2 \\ E[X^2|Y = 3] &= \sum_{x=1}^4 x^2 \cdot p_{X|Y}(x|3) = (1^2) \left(\frac{8}{21} \right) + (2^2) \left(\frac{1}{6} \right) + (3^2) \left(\frac{8}{21} \right) + (4^2) \left(\frac{1}{14} \right) = \frac{118}{21} \\ \Rightarrow \text{Var}(X|Y = 3) &= E[X^2|Y = 3] - (E[X|Y = 3])^2 = \frac{118}{21} - \left(\frac{15}{7} \right)^2 \approx 1 \end{aligned}$$

Thus, the expected hair color of a student with brown eyes is red.

Conditional Distributions of Continuous Random Variables

We now turn to the continuous setting. Note that definitions and results in the discrete setting transfer to the continuous setting by simply replacing sums with integrals and pmf's with pdf's. The following definition gives the formulas for conditional distributions and expectations of continuous random variables.

Definition 5.3.3

If X and Y are continuous random variables with joint pdf given by $f(x, y)$, then the **conditional probability density function (pdf) of X , given that $Y = y$** , is denoted $f_{X|Y}(x|y)$ and given by

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)}.$$

The **conditional expected value of X , given $Y = y$** , is

$$E[X|Y = y] = \int_{\mathbb{R}} x f_{X|Y}(x|y) dx$$

and the **conditional variance of X , given $Y = y$** , is

$$\text{Var}(X|Y = y) = E[X^2 | Y = y] - (E[X|Y = y])^2.$$

Similarly, we can define the conditional pdf, expected value, and variance of Y , given $X = x$, by swapping the roles of X and Y in the above.

Properties of Conditional PDF's

1. Conditional pdf's are valid pdf's. In other words, the conditional pdf for X , given $Y = y$, for a fixed y , is a valid pdf satisfying the following:

$$0 \leq f_{X|Y}(x|y) \quad \text{and} \quad \int_{\mathbb{R}} f_{X|Y}(x|y) dx = 1$$

2. In general, the conditional distribution of X given Y does not equal the conditional distribution of Y given X , i.e.,

$$f_{X|Y}(x|y) \neq f_{Y|X}(y|x).$$

3. If X and Y are independent, discrete random variables, then the following are true:

$$\begin{aligned} f_{X|Y}(x|y) &= f_X(x) \\ f_{Y|X}(y|x) &= f_Y(y) \end{aligned}$$

In other words, if X and Y are independent, then knowing the value of one random variable does not affect the probability of the other one.

[Link to Video: Overview of Definition 5.3.3 & Properties of Conditional PDF's](#)

Example 5.3.4

We verify the third property of conditional pdf's for the radioactive particle example ([Example 5.2.1](#)). Recall that we had the following joint pdf and marginal pdf's for X and Y :

$$\begin{aligned}f(x, y) &= 1, & \text{for } 0 \leq x \leq 1 \text{ and } 0 \leq y \leq 1 \\f_X(x) &= 1, & \text{for } 0 \leq x \leq 1 \\f_Y(y) &= 1, & \text{for } 0 \leq y \leq 1\end{aligned}$$

We showed in [Example 5.2.3](#) that X and Y are independent. So, if we fix $y \in [0, 1]$, the following shows that the conditional pdf of X , given $Y = y$, is equal to the marginal pdf of X , as stated in the third property of conditional pdf's above:

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)} = \frac{1}{1} = f_X(x)$$

Example 5.3.5

Continuing in the context of [Example 5.2.2](#), where X gave the amount of gas stocked and Y gave the amount of gas sold at a given gas station in a given week, we find the conditional pdf of Y given that $X = 0.5$. In other words, we find the conditional probability distribution for the amount of gas sold in a given week, when only half of the tank was stocked.

First, we find the marginal pdf for X :

$$f_X(x) = \int_{\mathbb{R}} f(x, y) dy = \int_0^x 3x dy = 3xy \Big|_0^x = 3x^2, \quad \text{for } 0 \leq x \leq 1.$$

So, if $X = 0.5$, then $f_X(0.5) = 3(0.5)^2 = 0.75$, and the conditional pdf of Y in this case is

$$f_{Y|X}(y|0.5) = \frac{f(0.5, y)}{f_X(0.5)} = \frac{3(0.5)}{0.75} = 2, \quad \text{for } 0 \leq y \leq 0.5.$$

Note the $f_{Y|X}(y|0.5)$ is the pdf for a uniform distribution on the interval $[0, 0.5]$. Thus, the conditional distribution of the amount of gas sold in a week, given that only half of the tank is stocked, is uniformly distributed between 0 and 0.5. Recognizing this, we can easily compute the conditional expected value of Y , given that $X = 0.5$:

$$E[Y|X = 0.5] = \frac{0.5}{2} = 0.25.$$

In other words, given that 50% of tank is stocked, we expect that 25% will be sold.

[Link to Video: Walkthrough of Example 5.3.5](#)

This page titled [5.3: Conditional Probability Distributions](#) is shared under a [not declared](#) license and was authored, remixed, and/or curated by [Kristin Kuter](#).