Multivariate Probability Distributions

March 23, 2020

As in the last section, we divide this section into two parts: discrete and continuous. We begin with the discrete case. We begin with the *marginal distribution* of a random variable. The *marginal pmf* of X_1 is simply the pmf X_1, we give it a longer name to distinguish it from the joint pmf. In the context of just one random variable this was unnecessary, as we only had one pmf to discuss.

1 Marginal Probability Distributions- Discrete

Our first theorem tells us how to compute marginal pmfs, from joint pmfs:

THEOREM 1.1. *Let Y_1 and Y_2 be discrete random variables with joint pmf $p_{Y_1,Y_2}(y_1,y_2)$. The marginal pmf of Y_1 is:*

$$p_{Y_1}(y_1)=\sum_{y_2} p_{Y_1,Y_2}(y_1,y_2)$$

and similarly, the marginal pmf of Y_2 is:

$$p_{Y_2}(y_2)=\sum_{y_1} p_{Y_1,Y_2}(y_1,y_2)$$

The right side is adding up the probability Y_2 takes on all possible values when $Y_1=y_1$. In other words, we compute the marginal pmfs by "summing out" the other random variable.

Let's consider the examples from the last section.

EXAMPLE 1.2. *Roll a fair 4-sided dice twice. Let Y_1 be the result of the first roll and Y_2 be the result of the second. Compute the marginal pmf of Y_1.*

Solution: Recall

$$p_{Y_1,Y_2}(y_1,y_2)=\frac{1}{16} \text{ for } y_1=1,2,3,4;\ y_2=1,2,3,4 \text{ and 0 otherwise}$$

so

$$p_{Y_1}(y_1)=\sum_{y_2=1}^{4}\frac{1}{16}=1/4 \text{ for } y_1=1,2,3,4 \text{ and 0 otherwise}$$

EXAMPLE 1.3. *There are 20 balls in a box, labeled 1,2,...,20. You draw 2 balls out of the box without replacement. Let X_1 be the number of the first ball you draw and X_2 be the number on the second ball. What is the marginal pmf of X_2?*

Solution: Recall

$$p_{X_1,X_2}(x_1,x_2) = \frac{1}{20\times 19} \text{ for } x_1=1,2,...,20;\ x_2=1,2,...,20 \text{ with } x_1 \neq x_2 \text{ and 0 otherwise .}$$

so

$$p_{X_2}(x_2) = \sum_{x_1=1,...20\text{but } \neq x_2} \frac{1}{20\times 19} = 1/20 \text{ for } x_2=1,2,...,20 \text{ and 0 otherwise .}$$

These last two examples could have been computed directly, but later we'll have more complicated joint pmf's and you'll really need to use this formula to compute marginal pmfs.

EXAMPLE 1.4. *If*

$$p_{X_1,X_2}(1,2)=1/3,\ p_{X_1,X_2}(5,-2)=1/2,\ p_{X_1,X_2}(6,4)=1/6 \text{ and 0 otherwise ,}$$

compute the marginal pmf of X_1*.*

Solution:

$$p_{X_1}(1)=1/3\ p_{X_1}(5)=1/2\ p_{X_1}(6)=1/6 \text{ and 0 otherwise}$$

Note that there are no sums because for each value the random variable X_1 there is only one possible value for X_2 to take.

2 Conditional Probability Distributions- Discrete

Just like we generalized the idea of computing probabilities of events to distributions of random variables, we can now generalize the idea of computing conditional probabilities of events to conditional distributions of random variables.

Sometimes we have a joint distribution of random variables but are interested in just one of the random variables, but we don't want to completely forget about the other random variable. We can use the conditional pmf to do this.

DEFINITION 2.1. *Let* Y_1 *and* Y_2 *be discrete random variables with joint pmf* $p_{Y_1,Y_2}(y_1,y_2)$*. The conditional probability mass function (conditional pmf) of* Y_1 *given* $Y_2=y_2$ *is*

$$p_{Y_1|Y_2}(y_1|y_2) = \mathbb{P}(\{Y_1=y_1\}|\{Y_2=y_2\})$$

provided $\mathbb{P}(\{Y_2=y_2\})>0$*.*

The right side of the above definition is just a conditional probability that we learned in Chapter 2, that is it satisfies the formula:

$$\mathbb{P}(\{Y_1=y_1\}|\{Y_2=y_2\}) = \frac{\mathbb{P}(\{Y_1=y_1\}\cap\{Y_2=y_2\})}{\mathbb{P}(\{Y_2=y_2\})} = \frac{p_{Y_1,Y_2}(y_1,y_2)}{p_{Y_2}(y_2)}.$$

The $\mathbb{P}(\{Y_2=y_2\})>0$, just means we don't want to condition on events that don't happen.

The conditional probability mass is still a probability mass function, it sums to 1 and is non-negative. This is similar to how conditional probabilities are still probabilities.

Just like in Chapter 2, it will sometimes happen that we're given a conditional probability $p_{Y_1|Y_2}(y_1|y_2)$ and actually want a joint probability $p_{Y_1,Y_2}(y_1,y_2)$. Rearranging the above relationship we also have:

$$p_{Y_1,Y_2}(y_1,y_2)=p_{Y_1|Y_2}(y_1|y_2)p_{Y_2}(y_2).$$

As before we're multiplying the thing we condition on by the probability that it happens.

EXAMPLE 2.2. *There are 20 balls in a box, labeled* 1,2,...,20. *You draw 2 balls out of the box without replacement. Let* X_1 *be the number of the first ball you draw and* X_2 *be the number on the second ball. What is the conditional pmf of* X_1 *given* X_2*?*

Solution: We start with the definition above, and then substitute the probabilities, which we computed earlier.

$$p_{X_1|X_2}(x_1|x_2)=\frac{p_{X_1,X_2}(x_1,x_2)}{p_{X_2}(x_2)}=\frac{\frac{1}{20*19}}{\frac{1}{20}}=\frac{1}{19}$$

for $x_1=1,...,20$, but $x_1\neq x_2$ and 0 otherwise.

This could have also been deduced directly, but hopefully this gives some intuition on the formula. In this problem the answer doesn't depend strongly on x_2 because the balls are chosen uniformly at random, in other examples the probability you compute can certainly involve the value you condition on.

Refer to example 5.7 in the book for another example of computing a conditional pmf from a joint pmf. Let's now do an example where we compute the joint and then marginal pmfs from a conditional pmf. This should look familiar from the homework, we now have introduced language to discuss this problem more in depth.

EXAMPLE 2.3. *We have a biased coin (probability of heads is equal to* $1/4$*). Consider the following 2 step process:*
In the first step we flip the coin until we get a heads. Let X *denote the trial on which the first heads occurs. In the second step we flip the coin* X *more times. Let* Y *be the number of heads in the second step. Compute joint pmf of* X *and* Y*. What is the probability* $X=5$ *and* $Y=2$*?*

Solution: The random variable X is geometric with parameter $p=1/4$:

$$p_X(x)=(3/4)^{x-1}1/4 \text{ for } x=1,2,3,...$$

the conditional distribution of Y given X is binomial with $n=X$ and $p=1/4$)

$$p_{Y|X}(y|x)=\binom{x}{y}(3/4)^{x-y}(1/4)^y \text{ for } y=0,1,2,...,x.$$

Therefore the joint pmf is

$$p_{X,Y}(x,y)=p_{Y|X}(y|x)p_X(x)=\binom{x}{y}(3/4)^{x-y}(1/4)^y\times(3/4)^{x-1}1/4 \text{ for } x=1,2,3,...,y=0,1,2,...,x.$$

To determine the probability $X=5$ and $Y=2$, we simply substitute $x=5$ and $y=2$ into the above formula:

$$p_{X,Y}(5,2)=\binom{5}{2}(3/4)^3(1/4)^2\times(3/4)^4 1/4.$$

3 Marginal Distribution of Continuous Random Variables

As we've done before we can repeat the above analysis in continuous case. We'll replace sums with integral and pmf's with pdf's.

THEOREM 3.1. *Let X_1 and X_2 be continuous random variables with joint pdf $p_{X_1,X_2}(x_1,x_2)$. The marginal pdf of X_1 is:*

$$f_{X_1}(x_1) = \int f_{X_1,X_2}(x_1,x_2)dx_2$$

and similarly, the marginal pdf of X_2 is:

$$f_{X_2}(x_2) = \int f_{X_1,X_2}(x_1,x_2)dx_1$$

The right side is integrating the the pdf of X_2 takes on all possible values when $X_1 = x_1$. In other words, we compute the marginal pdfs by "integrating out" the other random variable.

Let's consider the examples from the last section.

EXAMPLE 3.2. *Let X_1 and X_2 be continuous random variables with jpdf:*

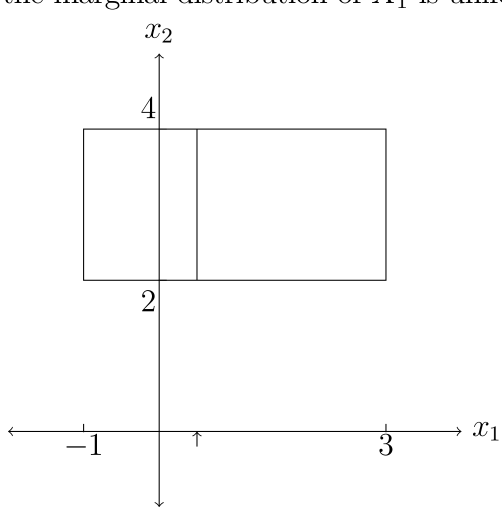
$$f_{X_1,X_2}(x_1,x_2) = \begin{cases} 1/8, & \text{if } -1 \le x_1 \le 3 \text{ and } 2 \le x_2 \le 4 \\ 0, & \text{otherwise.} \end{cases}$$

What is the marginal pdf of X_1?

Solution: When x_1 is not between -1 and 3 then the integral $\int_{-\infty}^{\infty} f_{X_1,X_2}(x_1,x_2)dx_2$ is just integrating the 0 function, and hence 0. When $-1 \le x_1 \le 3$ we have:

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1,x_2)dx_2 = \int_2^4 \frac{1}{8}dx_2 = \frac{1}{4}.$$

So the marginal distribution of X_1 is uniform on the interval $[-1,3]$.



EXAMPLE 3.3. *Let Y_1 and Y_2 be continuous random variables with jpdf:*

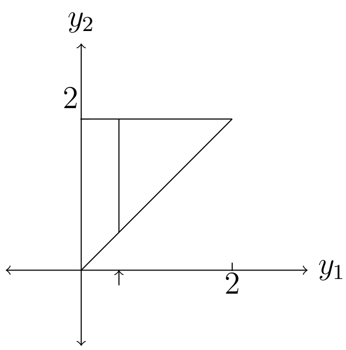
$$f_{Y_1,Y_2}(y_1,y_2)=\begin{cases} \frac{3}{4}y_1^2, & \textit{if } 0\leq y_1\leq y_2\leq 2 \\ 0, & \textit{otherwise.} \end{cases}$$

(a) Compute the marginal pdf of Y_1.

(b) Compute the marginal pdf of Y_2.

Solution:

(a) As in the last example if y_1 is not between 0 and 2 then the marginal pdf is 0. To compute the the pdf for $0\leq y_1\leq 2$ we begin by picking arbitrary point in this range (marked by the arrow) and compute the pdf at this point:



Now we compute the integral over this line.

$$f_{Y_1}(y_1)=\int_{y_1}^{2}\frac{3}{4}y_1^2dy_2 =\frac{3}{4}y_1^2(2-y_1)$$

So the final answer is

$$f_{Y_1}(y_1)=\begin{cases} \frac{3}{4}y_1^2(2-y_1), & \text{if } 0\leq y_1\leq 2 \\ 0, & \text{otherwise.} \end{cases}$$

Note that marginal distribution is a function of y_1 only, the y_2 terms have been integrated out, and we are left with just a function of y_1, as we expect.

(b) Exercise for you. The answer is:

$$f_{Y_2}(y_2)=\begin{cases} \frac{1}{4}y_2^3, & \text{if } 0\leq y_2\leq 2 \\ 0, & \text{otherwise.} \end{cases}$$

4 Conditional Distribution of Continuous Random Variables

Just as in the discrete case, we can define a conditional pdf

DEFINITION 4.1. *Let X_1 and X_2 be continuous random variables with joint pdf $f_{X_1,X_2}(x_1,x_2)$. The conditional probability density function (conditional pdf) of X_1 given $X_2=x_2$ is*

$$f_{X_1|X_2}(x_1|x_2)=\frac{f_{X_1,X_2}(x_1,x_2)}{f_{X_2}(x_2)}.$$

Remember the the actual value of the probability density of a random variable is not such a meaningful quantity. What is actually meaningful is the value of its integral over sets, these integrals give the probability a random variable lives in that set. The same is true here with conditional densities, and the connections to the conditional CDFs, which gives actual probabilities is explained in the book after Def. 5.6. Perhaps the best way to think about the conditional pdf is as just replacing the quantities in the conditional pmf with the respective pdf quantities.

Note that the conditional probability density is still a probability density, it integrates to 1 and is non-negative.

Let's see this works in an example.

EXAMPLE 4.2. *Let Y_1 and Y_2 be continuous random variables with jpdf:*

$$f_{Y_1,Y_2}(y_1,y_2) = \begin{cases} \frac{3}{4}y_1^2, & \text{if } 0 \le y_1 \le y_2 \le 2 \\ 0, & \text{otherwise.} \end{cases}$$

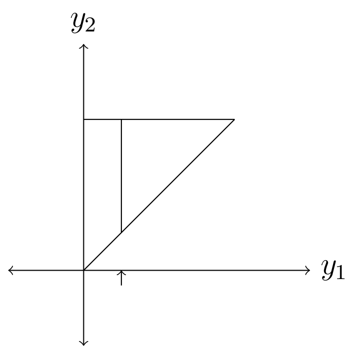
(a) Compute the conditional pdf of Y_2 given $Y_1 = y_1$.

(b) Given that $Y_1 = 1/2$, what is the probability Y_2 is less than 1?

(c) Compute the conditional pdf of Y_1 given $Y_2 = y_2$.

Solution:

(a) The idea here is that we want to restrict the the density a segment where y_1 is constant. Like



in this picture:
but we can't just take the value of the joint pdf along this line, because we want it to be a function that integrates to 1, so we need to normalize it. Since we already computed the marginal pdfs, we just need to plug them into the formula and remember the bounds.

$$f_{Y_2|Y_1}(y_2|y_1) = \frac{f_{Y_1,Y_2}(y_1,y_2)}{f_{Y_1}(y_1)} = \frac{\frac{3}{4}y_1^2}{\frac{3}{4}y_1^2(2-y_1)} = \frac{1}{2-y_1}.$$

for $y_1 \le y_2 \le 2$, and zero otherwise. You can get the bounds on y_2 from looking at the picture. (Remember we're thinking of y_1 as a fixed number and the conditional pdf is then a function of y_2)

(b) To get this conditional probability, we integrate the conditional probability from $-\infty$ to 1.

$$\mathbb{P}(Y_2 \le 1 | Y_1 = 1/2) = \int_{1/2}^{1} \frac{1}{2-1/2} dy_2 = \frac{1/2}{3/2} = 1/3$$

Note the the lower bound is $y_1 = 1/2$ because that is where the density starts being positive in the above picture. Since we are told $Y_1 = 1/2$ we replace the y_1 with $1/2$ everywhere in the equation.

(c) This is an exercise for you. The answer is:

$$f_{Y_1|Y_2}(y_1|y_2) = \frac{f_{Y_1,Y_2}(y_1,y_2)}{f_{Y_2}(y_2)} = \frac{\frac{3}{4}y_1^2}{\frac{1}{4}y_2^3} = \frac{3y_1^2}{y_2^3}$$

for $0 \le y_1 \le y_2$. (DON'T FORGET THE BOUNDS!)

We can also consider problems where we use a conditional probability and the marginal distribution of the random variable we condition on, to get the joint distribution of the two random variables.

EXAMPLE 4.3. *Let X be a uniform random variable on the interval $[0,1]$. Conditioned on the event $X = x$, let Y be an exponential random variable with mean $1/x$.*

(a) Compute the joint distribution of X and Y.

(b) Compute the marginal distribution of Y.

Solution: We are given the pdf of X is

$$f_X(x) = \begin{cases} 1, & \text{if } 0 \le x \le 1 \\ 0, & \text{otherwise.} \end{cases}$$

and that the conditional pdf of Y is

$$f_Y|X(y|x) = \begin{cases} xe^{-xy}, & \text{if } 0 \le y \\ 0, & \text{otherwise.} \end{cases}$$

So the joint pdf is

$$f_{X,Y}(x,y) = f_Y|X(y|x) f_X(x) = \begin{cases} xe^{-xy} * 1, & \text{if } 0 \le y,\ 0 \le x \le 1 \\ 0, & \text{otherwise.} \end{cases}$$

The marginal pdf of Y is then

$$f_Y(y) = \int f_{X,Y}(x,y)dx = \int_0^1 xe^{-xy}dx$$

After a bit of integration by parts and simplification one arrives at:

$$f_Y(y) = \begin{cases} \frac{1-e^{-y}(1+y)}{y^2} & \text{if } y \ge 0 \\ 0, & \text{otherwise.} \end{cases}$$