Recitation/Discussion 7

Joint Distributions

accompaniment to MATH 170E: Introduction to Probability and Statistics 1: Probability at the University of California, Los Angeles

Allen Lin

In this recitation, we will introduce and solve problems with joint distributions—both discrete and continuous.

Exit Survey

Please fill out the exit survey.

This is so I can get feedback and know what to focus on next time!

Joint Distributions of the Discrete Type

Until now, we have only considered situations involving one random variable. In many practical settings, however, will involve more than one random variable. For now, we consider two random variables.

Definition 1. Let X and Y be two random variables defined on discrete sample spaces S_X and S_Y, respectively. Let $S = S_X \times S_y$. The function $f(x,y) = P(X = x \text{ and } Y = y)$ is the *joint probability mass function* (joint pmf) and satisfies

- $0 \leq f(x,y) \leq 1$ for all $(x,y) \in S$,
- $\sum_{(x,y)\in S} f(x,y) = 1$, and
- $P(A) = \sum_{(x,y)\in A} f(x,y)$ for all $A \subseteq S$.

Note this follows the same definition as a regular pmf of a discrete random variable. Can you generalize this to arbitrary random variables?

Exercise 1. Roll a pair of fair 6-sided die. Let X denote the smaller of the two rolls, and let Y denote the larger of the two rolls. Find the pmf of the joint distribution of X and Y.

Solution. Consider the event $(X,Y) = (x,y)$ in the cases where $x = y$ and $x \neq y$. Here $S_X = \{1,\ldots,6\} = S_Y$. For a fixed $x \in S_X$ and $y \in S_Y$, when $x = y$ the event $(X,Y) = (x,y)$ occurs in only one way: when both rolls result in x and y. This happens with probability $1/36$. When $x \neq y$ the event $(X,Y) = (x,y)$ occurs in two ways: the first roll is x and the second roll is y, or the first roll is y and the second roll is x. This happens with probability $2/36$. Thus the joint pmf is

$$f(x,y) = \begin{cases} 1/36 & \text{if } 1 \leq x = y \leq 6, \\ 2/36 & \text{if } 1 \leq x < y \leq 6, \\ 0 & \text{otherwise.} \end{cases}$$

If we restrict a joint distribution to one random variable, we have a marginal distribution.

Definition 2. Let X and Y with spaces S_X and S_Y, respectively, have joint probability mass function $f(x,y)$ in total space $S = S_X \times S_Y$. The *marginal probability mass function of* X is defined by

$$f_X(x) = \sum_{y \in S_Y} f(x,y) = P(X = x)$$

Essentially fix x and sum over all y.

for all $x \in S_X$. Similarly, the *marginal probability mass function of* Y is defined by

$$f_Y(y) = \sum_{x \in S_X} f(x,y) = P(Y = y)$$

Essentially fix y and sum over all x.

for all $y \in S_Y$.

Definition 3. Let X and Y with spaces S_X and S_Y, respectively, have joint probability mass function $f(x,y)$. Two random variables X and Y are *independent* if and only if

$$f(x,y) = f_X(x) f_Y(y)$$

for all $x \in S_X$ and $y \in S_Y$, or equivalently

$$P(X = x \text{ and } Y = y) = P(X = x)P(Y = y).$$

Exercise 2. In the previous exercise, calculate the marginal distributions and determine if X and Y are independent. Justify your answer.

Solution. Using the first exercise we have the following joint pmf.

Y \ X	1	2	3	4	5	6	$f_Y(y)$
1	1/36	0	0	0	0	0	1/36
2	2/36	1/36	0	0	0	0	3/36
3	2/36	2/36	1/36	0	0	0	5/36
4	2/36	2/36	2/36	1/36	0	0	7/36
5	2/36	2/36	2/36	2/36	1/36	0	9/36
6	2/36	2/36	2/36	2/36	2/36	1/36	11/36
$f_X(x)$	11/36	9/36	7/36	5/36	3/36	1/36	

Reading from the table we obtain the marginal distributions

$$f_Y(y) = \begin{cases} 1/36 & \text{if } y = 1, \\ 3/36 & \text{if } y = 2, \\ 5/36 & \text{if } y = 3, \\ 7/36 & \text{if } y = 4, \\ 9/36 & \text{if } y = 5, \\ 11/36 & \text{if } y = 6, \end{cases} \qquad f_X(x) = \begin{cases} 11/36 & \text{if } x = 1, \\ 9/36 & \text{if } x = 2, \\ 7/36 & \text{if } x = 3, \\ 5/36 & \text{if } x = 4, \\ 3/36 & \text{if } x = 5, \\ 1/36 & \text{if } x = 6. \end{cases}$$

From here it is clear that X and Y are not independent:

$$f(1,1) = 1/36 \neq 11/36^2 = f_X(1) f_Y(1).$$

Or any other choice of x and y.

Finally, we have an analogous definition of expected value.

Definition 4. Let X and Y be discrete random variables with joint pmf $f(x,y)$ on the space S. Let $u(X,Y)$ be a function on X and Y. Then

$$\mathbb{E}[u(X,Y)] = \sum_{(x,y)\in S} u(x,y)f(x,y).$$

The most important cases of u are

- $u(X,Y) = X$ or $u(X,Y) = Y$, so

$$\begin{aligned}\mathbb{E}[u(X,Y)] &= \sum_{(x,y)\in S} xf(x,y) = \sum_{x\in S_X} x \sum_{y\in S_Y} f(x,y)\\ &= \sum_{x\in S_X} xf_X(x)\\ &= \mathbb{E}[X];\end{aligned}$$

Can you derive a similar expression for $\mathbb{E}[Y]$?

- $u(X,Y) = XY$, so

$$\mathbb{E}[u(X,Y)] = \sum_{(x,y)\in S} xyf(x,y);$$

- $u(X,Y) = X+Y$, so

$$\begin{aligned}\mathbb{E}[u(X,Y)] &= \sum_{(x,y)\in S} (x+y)f(x,y)\\ &= \sum_{x\in S_X} xf_X(x) + \sum_{y\in S_Y} yf_Y(y)\\ &= \mathbb{E}[X] + \mathbb{E}[Y].\end{aligned}$$

Exercise 3. There are ten chips in a bowl. One is labeled $(-1,-1)$, one is labeled $(0,-2)$, one is labeled $(-3,1)$, three are labeled $(0,0)$, two are labeled $(1,0)$, two are marked $(0,1)$, and one is marked $(1,1)$. You select a chip at random and earn the sum of two times the first coordinate and three times the second coordinate in coins. What is the expected number of coins that you will earn in one round of this game?

Solution. Let X and Y be the first and second coordinate of the chosen chip, respectively. Then the joint pmf of X and Y is

$$f(x,y) = \begin{cases} 1/10 & \text{if } (x,y) = (-1,-1),\\ 1/10 & \text{if } (x,y) = (0,-2),\\ 3/10 & \text{if } (x,y) = (0,0),\\ 2/10 & \text{if } (x,y) = (1,0),\\ 2/10 & \text{if } (x,y) = (0,1),\\ 1/10 & \text{if } (x,y) = (1,1).\end{cases}$$

The required expected value is

$$\mathbb{E}[2X+3Y] = \sum_{(x,y)\in S} (2x+3y)f(x,y) = 0.2,$$

so you are expected to win 0.2 coins per round.

Joint Distributions of the Continuous Type

Joint distributions of the continuous type have a similar definition: replace pmf with pdf and replace all the sums with integrals.

Definition 5. let X and Y be two random variables defined on continuous sample spaces S_X and S_Y, respectively. Let $S = S_X \times S_Y$. The function $f(x, y)$ is the *joint probability density function* (joint pdf) of X and Y and satisfies

- $0 \leq f(x, y) \leq 1$ for all $(x, y) \in S$,
- $\int_S f(x, y)\, dx\, dy = 1$, and
- $P(A) = \int_A f(x, y)\, dx\, dy$.

If we restrict a joint distribution to one random variable, we have a marginal distribution.

Definition 6. Let X and Y with spaces S_X and S_Y, respectively, have joint probability density function $f(x, y)$ in total space $S = S_X \times S_Y$. The *marginal probability density function of* X is defined by

$$f_X(x) = \int_{S_Y} f(x, y)\, dy$$

for all $x \in S_X$. Similarly, the *marginal probability density function of* Y is defined by

$$f_Y(y) = \int_{S_X} f(x, y)\, dx$$

for all $y \in S_Y$.

Definition 7. Let X and Y with spaces S_X and S_Y, respectively, have joint probability density function $f(x, y)$. Two random variables X and Y are *independent* if and only if

$$f(x, y) = f_X(x) f_Y(y)$$

for all $x \in S_X$ and $y \in S_Y$.

Definition 8. Let X and Y be discrete random variables with joint pdf $f(x, y)$ on the space S. Let $u(X, Y)$ be a function on X and Y. Then

$$\mathbb{E}[u(X, Y)] = \int_S u(x, y) f(x, y)\, dx\, dy.$$

The most important cases of u are

- $u(X, Y) = X$ or $u(X, Y) = Y$, so

$$\mathbb{E}[u(X, Y)] = \int_S x f(x, y)\, dx;$$

Can you derive a similar expression for $\mathbb{E}[Y]$?

- $u(X,Y) = XY$, so
$$\mathbb{E}[u(X,Y)] = \int_S xyf(x,y)\,dx\,dy;$$
- $u(X,Y) = X + Y$, so
$$\mathbb{E}[u(X,Y)] = \int_S (x+y)f(x,y)\,dx\,dy = \mathbb{E}[X] + \mathbb{E}[Y].$$

Exercise 4. Let X and Y have joint pdf $f(x,y) = cx^2y$ for $-y \le x \le 1$ and $0 \le y \le 1$.

(a) Determine the value of c.

(b) Find the marginal distributions f_X and f_Y.

(c) Are X and Y independent?

(d) Calculate the probability $P(0 \le Y \le X \le 1)$.

Solution. (a) To determine c, we have
$$\begin{aligned} 1 &= \int_0^1 \int_{-y}^1 f(x,y)\,dx\,dy \\ &= \int_0^1 \int_{-y}^1 cx^2y\,dx\,dy \\ &= \int_0^1 \frac{c}{3}\left(y + y^4\right)\,dy \\ &= \frac{7c}{30} \end{aligned}$$
so $c = 30/7$.

(b) The marginal distribution of Y is
$$f_Y(y) = \int_{-y}^1 f(x,y)\,dx = \int_{-y}^1 \frac{30}{7}x^2y\,dx = \frac{10}{7}\left(y + y^4\right).$$
There are two regions for x depending on y: when $-1 \le x \le 0$ we have $-x \le y \le 1$, and when $0 \le x \le 1$ we have $0 \le y \le 1$. Thus
$$\begin{aligned} f_X(x) &= \begin{cases} \int_{-x}^1 f(x,y)\,dx\,dy & \text{if } -1 \le x < 0, \\ \int_0^1 f(x,y)\,dx\,dy & \text{if } 0 \le x \le 1, \end{cases} \\ &= \begin{cases} \int_{-x}^1 \frac{30}{7}x^2y\,dx\,dy & \text{if } -1 \le x < 0, \\ \int_0^1 \frac{30}{7}x^2y\,dx\,dy & \text{if } 0 \le x \le 1, \end{cases} \\ &= \begin{cases} \frac{15}{7}x^2\left(1 - x^2\right) & \text{if } -1 \le x < 0, \\ \frac{15}{7}x^2 & \text{if } 0 \le x \le 1. \end{cases} \end{aligned}$$

(c) No since $f(x,y) = 30x^2y/7 \ne 10\left(y + y^4\right)/7 \cdot 15x^2/7 = f_Y(y)f_X(x)$ when $x \in [0,1]$.

(d) We calculate
$$\int_0^1 \int_0^x f(x,y)\,dy\,dx = \int_0^1 \int_0^x \frac{30}{7}x^2y\,dy\,dx = \int_0^1 \frac{15}{7}x^4\,dx = \frac{3}{7}.$$