Lecture Note: Week 9

MATH 461: Probability Theory, Spring 2021
Daesung Kim

Lecture 22. Joint Distribution Functions (Sec 6.1)

In this section, we consider a collection of random variables $X_1, X_2, \cdots, X_n$ defined on a sample space S. In particular, we are interested in modeling relationships between them. For example, collections of random variables are involved in the following:

(i) X_1 is the price of a product today, X_2 the price tomorrow, and so on.

(ii) X_1 is rainfall in IL, X_2 rainfall in IN, and so on.

(iii) Statistical topics such as time series, multivariate analysis, multiple linear regression, factor models.

(iv) Probability topics such as Markov chains, stochastic processes.

First, we focus on two random variable X and Y on a sample space S. The probability of X and Y can be realized by their joint cumulative distribution function.

> **Definition: Joint cumulative distribution functions**
> For random variables X and Y, the joint cumulative distribution function $F(a,b)$ on $\mathbb{R}^2$ is defined by
> $$F(a,b) = \mathbb{P}(X \le a, Y \le b).$$

The distributions of each X and Y can be obtained from the joint distribution. Indeed, we have

$$F_X(a) = \mathbb{P}(X \le a) = \lim_{b\to\infty} \mathbb{P}(X \le a, Y \le b) = \lim_{b\to\infty} F(a,b),$$
$$F_Y(b) = \mathbb{P}(Y \le b) = \lim_{a\to\infty} \mathbb{P}(X \le a, Y \le b) = \lim_{a\to\infty} F(a,b).$$

Example 1. Describe $\mathbb{P}(X > a, Y > b)$ and $\mathbb{P}(a_1 < X \le a_2, b_1 < Y \le b_2)$ for $a, a_1, a_2, b, b_1, b_2 \in \mathbb{R}$ in terms of distribution functions.

> **Definition: Joint probability mass function**
> For discrete random variables X and Y, the joint probability mass function is defined by
> $$p(x,y) = \mathbb{P}(X = x, Y = y).$$

The probability mass functions for each X and Y are

$$p_X(x) = \mathbb{P}(X = x) = \sum_y \mathbb{P}(X = x, Y = y) = \sum_y p(x,y),$$
$$p_Y(y) = \mathbb{P}(Y = y) = \sum_x \mathbb{P}(X = x, Y = y) = \sum_x p(x,y).$$

Example 2. Two fair dice are rolled. Find the joint probability mass function of X and Y when X is the largest value obtained on any die and Y is the sum of the values.

Definition: Joint continuity
We say that X and Y are jointly continuous if there exists a nonnegative function $f(x, y)$ on $\mathbb{R}^2$ such that

$$\mathbb{P}((X, Y) \in C) = \iint_C f(x, y)\, dxdy$$

for every C in $\mathbb{R}^2$. The function $f(x, y)$ is called the joint probability density function of X and Y.

The probability mass functions for each X and Y are

$$p_X(x) = \mathbb{P}(X = x) = \sum_y \mathbb{P}(X = x, Y = y) = \sum_y p(x, y),$$
$$p_Y(y) = \mathbb{P}(Y = y) = \sum_x \mathbb{P}(X = x, Y = y) = \sum_x p(x, y).$$

Remark 3. Suppose X and Y are jointly continuous.

(i) For any sets $A, B \subseteq \mathbb{R}$,

$$\mathbb{P}(X \in A, Y \in B) = \int_B \int_A f(x, y)\, dxdy.$$

(ii) The cumulative distribution function $F(a, b)$ is

$$F(a, b) = \mathbb{P}(X \leq a, Y \leq b) = \int_{-\infty}^{b} \int_{-\infty}^{a} f(x, y)\, dxdy.$$

Differentiating with respect to a and b, we have

$$\frac{\partial^2}{\partial a \partial b} F(a, b) = f(a, b).$$

(iii) The density of X can be obtained from the joint density $f(x, y)$. For $A \subseteq \mathbb{R}$, one can see that

$$\int_A f_X(x)\, dx = \mathbb{P}(X \in A) = \mathbb{P}(X \in A, Y \in \mathbb{R}) = \int_{\mathbb{R}} \int_A f(x, y)\, dxdy = \int_A \left(\int_{\mathbb{R}} f(x, y)\, dy \right) dx$$

Let $A = (-\infty, x)$ and differentiate in x, then

$$f_X(x) = \int_{\mathbb{R}} f(x, y)\, dy.$$

Similarly, we have

$$f_Y(y) = \int_{\mathbb{R}} f(x, y)\, dx.$$

We call f_X and f_Y the marginal densities of X and Y.

Example 4. The joint density function of X and Y is given by

$$f(x, y) = \begin{cases} 2e^{-x}e^{-2y}, & 0 < x < \infty, 0 < y < \infty \\ 0, & \text{otherwise.} \end{cases}$$

Compute $\mathbb{P}(X < Y)$.

Example 5. The joint density function of X and Y is given by

$$f(x,y) = \begin{cases} c(y^2 - x^2)e^{-y}, & -y \le x \le y, 0 < y < \infty \\ 0, & \text{otherwise.} \end{cases}$$

(i) Find c.

(ii) Compute $\mathbb{P}(0 < X < 1, Y < 1)$.

(iii) Find the marginal density of Y.

(iv) Compute $\mathbb{E}[Y]$.

More than two random Variables
For random variables $X_1, X_2, \cdots, X_n$, the joint cumulative distribution function is defined by

$$F(a_1, a_2, \cdots, a_n) = \mathbb{P}(X_1 \le a_1, X_2 \le a_2, \cdots, X_n \le a_n).$$

The random variables $X_1, X_2, \cdots, X_n$ are jointly continuous if there exists a nonnegative function $f(x_1, x_2, \cdots, x_n)$ such that

$$\mathbb{P}((X_1, X_2, \cdots, X_n) \in C) = \iint \cdots \int_C f(x_1, x_2, \cdots, x_n)\, dx_1 dx_2 \cdots dx_n$$

for all $C \subseteq \mathbb{R}^n$.

Lecture 23. Independent Random Variables (Sec 6.2)

Definition
Two random variables X and Y are independent if for any sets A and B

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B).$$

It is equivalent to the following:

(i) $F(a,b) = F_X(a)F_Y(b)$ for all $a, b \in \mathbb{R}$;

(ii) (Discrete case) $p(x,y) = p_X(x)p_Y(y)$ for all $x, y \in \mathbb{R}$;

(iii) (Discrete case) $p(x,y) = h(x)g(y)$ for all $x, y \in \mathbb{R}$, for some h and g;

(iv) (Jointly continuous case) $f(x,y) = f_X(x)f_Y(y)$ for all $x, y \in \mathbb{R}$;

(v) (Jointly continuous case) $f(x,y) = h(x)g(y)$ for all $x, y \in \mathbb{R}$, for some h and g.

Otherwise, we say that X and Y are dependent.

Remark 6. Let E, F be events on a sample space S. Recall that E, F are independent if $\mathbb{P}(E \cup F) = \mathbb{P}(E)\mathbb{P}(F)$. Define

$$X = I_E = \begin{cases} 1, & E \text{ occurs,} \\ 0, & \text{otherwise,} \end{cases} \qquad Y = I_F = \begin{cases} 1, & F \text{ occurs,} \\ 0, & \text{otherwise.} \end{cases}$$

Then, X and Y are independent if and only if E and F are independent.

Example 7. If the joint density function of X and Y is

$$f(x,y) = \begin{cases} 6e^{-2x}e^{-3y}, & 0 < x < \infty, 0 < y < \infty \\ 0, & \text{otherwise,} \end{cases}$$

are the random variables independent? Find the marginal densities f_X and f_Y.

Example 8. If the joint density function of X and Y is

$$f(x,y) = \begin{cases} 24xy, & 0 < x < 1, 0 < y < 1, 0 < x + y < 1 \\ 0, & \text{otherwise,} \end{cases}$$

are the random variables independent? Find the marginal densities f_X and f_Y.

Example 9. Two points are selected randomly on a line of length L so as to be on opposite sides of the midpoint of the line. In other words, the two points X and Y are independent random variables such that X is uniformly distributed over $(0, L/2)$ and Y is uniformly distributed over $(L/2, L)$. Find the probability that the distance between the two points is greater than $L/3$.

Definition
Random variables $X_1, X_2, \cdots, X_n$ are independent if, for any sets $A_1, A_2, \cdots, A_n$

$$\mathbb{P}(X_1 \in A_1, X_2 \in A_2, \cdots, X_n \in A_n) = \mathbb{P}(X_1 \in A_1)\mathbb{P}(X_2 \in A_2)\cdots\mathbb{P}(X_n \in A_n).$$

If the random variables are jointly continuous, it is equivalent to

$$f(x_1, x_2, \cdots, x_n) = f_{X_1}(x_1)f_{X_2}(x_2)\cdots f_{X_n}(x_n)$$

for all $x_1, x_2, \cdots, x_n$.

Example 10. If X_1, X_2, X_3 are independent random variables that are uniformly distributed over $(0, 1)$, compute the probability that the largest of the three is greater than the sum of the other two.

Remark 11. Let X_1, X_2, X_3, X_4 be independent uniform random variables on $[0, 1]$. Define $X^{(i)}$ be the i-th smallest random variable between X_1, X_2, X_3, X_4 for $i = 1, 2, 3, 4$. Let $Y = X^{(2)}$ and $Z = 1 - X^{(3)}$, then one can see that the joint density of Y and Z is

$$f(y,z) = \begin{cases} 24yz, & 0 < y < 1, 0 < z < 1, 0 < y + z < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Further examples

Example 12. Suppose that the number of people who enter a post office on a given day is a Poisson random variable with parameter λ. Each person who enters the post office is a male with probability p and a female with probability $1 - p$. Show that the number of males and females entering the post office are independent Poisson random variables with respective parameters λp and $\lambda(1 - p)$.

Example 13. Buffon's needle problem A table is ruled with equidistant parallel lines a distance D apart. A needle of length L, where $L < D$, is randomly thrown on the table. What is the probability that the needle will intersect one of the lines (the other possibility being that the needle will be completely contained in the strip between two lines)? Let us determine the position of the needle by specifying

(i) the distance X from the middle point of the needle to the nearest parallel line and

(ii) the angle θ between the needle and the projected line of length X.

References

[SR] Sheldon Ross, *A First Course in Probability*, 9th Edition, Pearson

Department of Mathematics, University of Illinois at Urbana-Champaign
E-mail address: daesungk@illinois.edu