

# Section 5: Joint distributions

STA 35C – Statistical Data Science III

**Instructor:** Akira Horiguchi

Fall Quarter 2025 (Sep 24 – Dec 12)  
MWF, 12:10 PM – 1:00 PM, Olson 158  
University of California, Davis

Based on Chapter 5 of textbook: <https://www.probabilitycourse.com/>

- Contains problems with solutions, and problems without solutions.

**1** Introduction

**2** For discrete random variables

**3** For continuous random variables

**4** Covariance and correlation

# Introduction

We are often interested in several random variables that are related to each other.

- A person's height and weight are typically related.
- A person's age and SAT score are typically related.

First we will study two random variables, but easy to extend to  $\geq 2$  random variables.

**For discrete random variables**

### Definition: Joint probability mass function (PMF)

The **joint PMF** of two discrete random variables  $X$  and  $Y$  is the function  $f_{XY}$  defined as

$$f_{XY}(a, b) = P(X = a, Y = b).$$

Can be expressed as “the probability that  $X = a$  **and**  $Y = b$ .”

- $P(X = a, Y = b)$  can be very different from  $P(X = a) P(Y = b)$ .

Consider drawing an animal from a population of orange giraffes and purple fish.

The **joint range** for  $X$  and  $Y$  is defined as

$$R_{XY} = \{(a, b) \mid f_{XY}(a, b) > 0\}.$$

- $R_{XY}$  is always a subset of  $R_X \times R_Y$ , but  $R_X \times R_Y$  might have pairs that  $R_{XY}$  does not.

## Marginal PMF

The joint PMF  $f_{XY}$  contains all the information regarding the distributions of  $X$  and  $Y$ . We can obtain the *marginal PMF* of  $X$  by

$$f_X(a) = P(X = a) = \sum_{b \in R_Y} P(X = a, Y = b) = \sum_{b \in R_Y} f_{XY}(a, b).$$

Similarly, we can obtain the *marginal PMF* of  $Y$  by

$$f_Y(b) = \sum_{a \in R_X} f_{XY}(a, b).$$

E.g. suppose we draw a student from a high school and consider their year and the math class that the student is currently taking.

## Example

Consider two random variables  $X$  and  $Y$  with joint PMF

**Table 5.1 Joint PMF of  $X$  and  $Y$  in Example 5.1**

	$Y = 0$	$Y = 1$	$Y = 2$
$X = 0$	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{8}$
$X = 1$	$\frac{1}{8}$	$\frac{1}{6}$	$\frac{1}{6}$

1. Find  $P(X = 0, Y \leq 1)$ .
2. Find the marginal PMFs of  $X$  and  $Y$ .
3. Find  $P(Y = 1 \mid X = 0)$ .
4. Are  $X$  and  $Y$  independent?



Often, we observe the value of a random variable  $X$ , and we want to update the PMF of another random variable  $Y$  whose value has not yet been observed.

E.g., regression of response  $Y$  on covariate  $X$ .

### Definition: Conditional probability mass function (PMF)

For two discrete random variables  $X$  and  $Y$  with respective marginal PMFs  $f_X$  and  $f_Y$ , the **conditional PMF of  $Y$  given  $X$**  is the function  $f_{Y|X}$  defined as

$$f_{Y|X}(b|a) = \frac{f_{XY}(a, b)}{f_X(a)}, \quad \text{for any } a \in R_X \text{ and } b \in R_Y.$$

Similarly (by symmetry), the **conditional PMF of  $X$  given  $Y$**  is the function  $f_{X|Y}$  defined as

$$f_{X|Y}(a|b) = \frac{f_{XY}(a, b)}{f_Y(b)}, \quad \text{for any } a \in R_X \text{ and } b \in R_Y.$$

# Conditional Expectation

Often want to find the mean of a conditional PMF.

Similar to finding a “regular” expected value, but replace PMF with the conditional PMF.

## Definition: Conditional expectation

For two discrete random variables  $X$  and  $Y$ , the *conditional expectation of  $Y$  given  $X = a$*  is

$$E[Y \mid X = a] = \sum_{b \in R_Y} b \cdot f_{Y|X}(b|a).$$

Similarly (by symmetry), the *conditional expectation of  $X$  given  $Y = b$*  is

$$E[X \mid Y = b] = \sum_{a \in R_X} a \cdot f_{X|Y}(a|b).$$

**For continuous random variables**

## Definition: Joint probability density function (PDF)

Two random variables  $X$  and  $Y$  are *jointly continuous* if there exists a nonnegative function  $f_{XY}: \mathbb{R}^2 \rightarrow \mathbb{R}$ , such that, for any set  $A \subset \mathbb{R}^2$ , we have

$$P((X, Y) \in A) = \iint_A f_{XY}(a, b) \, da \, db.$$

The function  $f_{XY}$  is called the *joint PDF* of  $X$  and  $Y$ .

The *joint range* for  $X$  and  $Y$  is defined as

$$R_{XY} = \{(a, b) \mid f_{XY}(a, b) > 0\}.$$

The joint PDF  $f_{XY}$  contains all the information regarding the distributions of  $X$  and  $Y$ . We can obtain the *marginal PDF* of  $X$  by

$$f_X(a) = \int_{-\infty}^{\infty} f_{XY}(a, b) \, db, \quad \text{for all } a.$$

Similarly, we can obtain the *marginal PDF* of  $Y$  by

$$f_Y(b) = \int_{-\infty}^{\infty} f_{XY}(a, b) \, da, \quad \text{for all } b.$$

### Definition: Conditional probability density function (PDF)

For two jointly continuous random variables  $X$  and  $Y$  with respective marginal PDFs  $f_X$  and  $f_Y$ , the **conditional PDF of  $Y$  given  $X$**  is the function  $f_{Y|X}$  defined as

$$f_{Y|X}(b|a) = \frac{f_{XY}(a, b)}{f_X(a)} \quad \text{for any } a \in R_X \text{ and } b \in R_Y.$$

Similarly (by symmetry), the **conditional PDF of  $X$  given  $Y$**  is the function  $f_{X|Y}$  defined as

$$f_{X|Y}(a|b) = \frac{f_{XY}(a, b)}{f_Y(b)} \quad \text{for any } a \in R_X \text{ and } b \in R_Y.$$

# Conditional expectation and variance

For these definitions, suppose  $X$  and  $Y$  are two jointly continuous random variables.

## Definition: Conditional expectation

The *conditional expectation of  $Y$  given  $X = a$*  is

$$E[Y | X = a] = \int_{-\infty}^{\infty} b \cdot f_{Y|X}(b|a) db.$$

## Definition: Conditional variance

The *conditional variance of  $Y$  given  $X = a$*  is

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

# Covariance and correlation



A joint PMF/PDF characterizes the relationship between two random variables  $X$  and  $Y$ .

- Simpler characterizations are the *covariance* and *correlation* of  $X$  and  $Y$ .

## Definition 4: Covariance $\text{Cov}(\cdot, \cdot)$

Let  $X, Y$  be random variables with  $E(X^2), E(Y^2) < \infty$ . Then the **covariance** between  $X$  and  $Y$  is defined as

$$\text{Cov}(X, Y) := E[(X - EX)(Y - EY)]. \quad (1)$$

- The covariance between  $X$  and  $Y$  can also be calculated as

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y). \quad (2)$$

- We say  $X$  and  $Y$  are *uncorrelated* if  $\text{Cov}(X, Y) = 0$ . Then  $X$  and  $Y$  have no linear relationship, and  $E(XY) = E(X)E(Y)$ .
- $\text{Cov}(X, Y) > 0$  indicate a positive linear relationship between  $X$  and  $Y$ .
- $\text{Cov}(X, Y) < 0$  indicate a negative linear relationship between  $X$  and  $Y$ .
- Covariance is symmetric:  $\text{Cov}(X, Y) = \text{Cov}(Y, X)$ .

From either (1) or (2), we can deduce that  $\text{Cov}(X, X) = \text{Var}(X)$ .

- In general, we can use covariance to prove/deduce many results for variance.

## Definition 5: Correlation coefficient $\rho(\cdot, \cdot)$

Let  $X, Y$  be random variables with  $E(X^2), E(Y^2) < \infty$ . Then, the **correlation coefficient** between  $X$  and  $Y$  is defined as, provided  $\text{Var}(X) > 0$  and  $\text{Var}(Y) > 0$ ,

$$\rho(X, Y) := \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}} \in [-1, 1]. \quad (3)$$

- $\rho(X, Y) = 0 \Rightarrow$  between  $X$  and  $Y$  is no linear relationship.
- $\rho(X, Y) = -1$  ( $1$ )  $\Rightarrow$  all values of  $X$  and  $Y$  lie on a line with negative (positive) slope.
- If  $\rho(X, Y)$  is close to  $-1$  ( $1$ ), there is a strong negative (positive) linear relationship between  $X$  and  $Y$ .