Chapter 9: Independence and Conditioning (Joint Distributions)

Meike Niederhausen and Nicky Wakim

2024-10-21

Learning Objectives

1. Calculate probabilities for a pair of discrete random variables
2. Calculate a *joint, marginal, and conditional* probability mass function (pmf)
3. Calculate a *joint, marginal, and conditional* cumulative distribution function (CDF)

Where are we?

Basics of probability

- Outcomes and events
- Sample space
- Probability axioms
- Probability properties
- Counting
- Independence
- Conditional probability
- Bayes' Theorem
- Random Variables

Probability for discrete random variables

- Functions: pmfs/CDFs
- Important distributions
- Joint distributions
- Expected values and variance

Probability for continuous random variables

- Calculus
- Functions: pdfs/CDFs
- Important distributions
- Joint distributions
- Expected values and variance

Advanced probability

- Central limit theorem
- Functions: moment generating functions

What is a joint pmf?

Definition: joint pmf

The **joint pmf** of a pair of discrete r.v.'s X and Y is

$$p_{X,Y}(x, y) = \mathbb{P}(X = x \; and \; Y = y) = \mathbb{P}(X = x, Y = y)$$

This chapter's main example

Example 1

Let X and Y be two random draws from a box containing balls labelled 1, 2, and 3 without replacement.

1. Find $p_{X,Y}(x, y)$.
2. Find $\mathbb{P}(X + Y = 3)$.
3. Find $\mathbb{P}(Y = 1)$.
4. Find $\mathbb{P}(Y \leq 2)$.
5. Find the joint CDF $F_{X,Y}(x, y)$ for the joint pmf $p_{X,Y}(x, y)$
6. Find the marginal CDFs $F_X(x)$ and $F_Y(y)$
7. Find $p_{X|Y}(x|y)$.
8. Are X and Y independent? Why or why not?

Joint pmf

Example 1

Let X and Y be two random draws from a box containing balls labelled 1, 2, and 3 without replacement.

1. Find $p_{X,Y}(x, y)$.
2. Find $\mathbb{P}(X + Y = 3)$.

Marginal pmf's

Example 1

Let X and Y be two random draws from a box containing balls labelled 1, 2, and 3 without replacement.

3. Find $\mathbb{P}(Y = 1)$.
4. Find $\mathbb{P}(Y \leq 2)$.

Remarks on the joint pmf

Some properties of joint pmf's:

- A joint pmf $p_{X,Y}(x, y)$ must satisfy the following properties:
  - $p_{X,Y}(x, y) \geq 0$ for all x, y.
  - $\sum_{\{all\ x\}} \sum_{\{all\ y\}} p_{X,Y}(x, y) = 1.$
- Marginal pmf's:
  - $p_X(x) = \sum_{\{all\ y\}} p_{X,Y}(x, y)$
  - $p_Y(y) = \sum_{\{all\ x\}} p_{X,Y}(x, y)$

What is a joint CDF?

Definition: joint CDF

The **joint CDF** of a pair of discrete r.v.'s X and Y is

$$F_{X,Y}(x, y) = \mathbb{P}(X \leq x \; and \; Y \leq y) = \mathbb{P}(X \leq x, Y \leq y)$$

Joint CDFs

Example 1

Let X and Y be two random draws from a box containing balls labelled 1, 2, and 3 without replacement.

5. Find the joint CDF $F_{X,Y}(x, y)$ for the joint pmf $p_{X,Y}(x, y)$

Marginal CDFs

Example 1

Let X and Y be two random draws from a box containing balls labelled 1, 2, and 3 without replacement.

6. Find the marginal CDFs $F_X(x)$ and $F_Y(y)$

Remarks on the joint and marginal CDF

- $F_X(x)$: right most columns of the CDf table (where the Y values are largest)
- $F_Y(y)$: bottom row of the table (where X values are largest)
- $F_X(x) = \lim_{y \to \infty} F_{X,Y}(x, y)$
- $F_Y(y) = \lim_{x \to \infty} F_{X,Y}(x, y)$

Independence and Conditioning

Recall that for *events* A and B,

- $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$
- A and B are independent if and only if
  - $\mathbb{P}(A|B) = \mathbb{P}(A)$
  - $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$

Independence and conditioning are defined similarly for r.v.'s, since

$$p_X(x) = \mathbb{P}(X = x) \text{ and } p_{X,Y}(x, y) = \mathbb{P}(X = x, Y = y).$$

What is the conditional pmf?

Definition: conditional pmf

The **conditional pmf** of a pair of discrete r.v.'s X and Y is defined as

$$p_{X|Y}(x|y) = \mathbb{P}(X = x|Y = y) = \frac{\mathbb{P}(X = x \textit{ and } Y = y)}{\mathbb{P}(Y = y)} = \frac{p_{X,Y}(x, y)}{p_Y(y)}$$

if $p_Y(y) > 0$.

Remarks on the conditional pmf

The following properties follow from the conditional pmf definition:

- If $X \perp Y$ (independent)
  - $p_{X|Y}(x|y) = p_X(x)$ for all x and y
  - $p_{X,Y}(x,y) = p_X(x)p_Y(y)$ for all x and y
  - Which also implies ($\Rightarrow$): $F_{X,Y}(x,y) = F_X(x)F_Y(y)$ for all x and y
- If $X_1, X_2, \ldots, X_n$ are independent
  - $$p_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n) = P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = \prod_{i=1}^{n} p_{X_i}(x_i)$$
  - $$F_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n) = P(X_1 \le x_1, X_2 \le x_2, \ldots, X_n \le x_n) = \prod_{i=1}^{n} P(X_i \le x_i) = \prod_{i=1}^{n} F_{X_i}(x_i)$$

Conditional pmf's

Example 1

Let X and Y be two random draws from a box containing balls labelled 1, 2, and 3 without replacement.

7. Find $p_{X|Y}(x|y)$.

8. Are X and Y independent? Why or why not?

Remark:

- To show that X and Y are *not* independent, we just need to find one counter example
- However, to show that they *are* independent, we need to verify this for all possible pairs of x and y