Probability

7. Couples of Random Variables

Ewen Gallic

Bachelor in Economics
International Program in Economics and Management
Academic Year 2026

amU FEG
Aix Marseille Université

Foreword

- In this chapter, we extend the study of probability distributions to **random vectors in dimension two**.
- We introduce:
    - joint distributions (discrete and continuous),
    - marginal and conditional distributions,
    - independence in the bivariate setting,
    - expectations of functions and covariance,
    - the distribution of sums of random variables.
- Particular attention is given to the role of the **joint density** and to change-of-variable techniques in dimension two (using notions defined in Pierre Michel's course from last semester).
- These slides are adapted from lecture notes by Mathieu Faure, kindly shared.
- Supplementary reading: Chapter 4 of Hansen (2022).

1. Introduction: Random Vectors in Dimension Two

Random Vectors

- A **random vector** is a collection of random variables considered jointly.
- Depending on the dimension, we distinguish:
  - **Univariate** random variables: a single random variable X (as in previous chapters),
  - **Bivariate** random vectors: two random variables (X, Y),
  - **Multivariate** random vectors: $d \geq 3$ random variables $(X_1, \ldots, X_d)$.
- In this chapter, we focus exclusively on the **bivariate case**.
- The general multivariate case will not be covered in this course.

Bivariate Random Variables

Definition 1. Bivariate Random Vector

*A **bivariate random vector** consists of two random variables*

$$(X, Y),$$

*defined on the same probability space and endowed with a **joint distribution**.*

- We usually write a bivariate random vector using two uppercase letters, such as (X, Y), or (X_1, X_2).
- Specific realizations are denoted by (x, y) or (x_1, x_2).
- The joint distribution captures how the two random variables behave **together**, not only individually.

Bivariate Random Vectors as Mappings to $\mathbb{R}^2$

A **bivariate random vector** is a mapping

$$(X, Y) : \Omega \to \mathbb{R}^2, \qquad \omega \mapsto (X(\omega), Y(\omega)).$$

Example: two fair coin flips

Let $\Omega = \{HH, HT, TH, TT\}$, where each outcome is equally likely. We code each flip as $0 =$ Head and $1 =$ Tail.

Then (X, Y) maps each outcome to a point in $\mathbb{R}^2$:

$$HH \mapsto (0, 0), \quad HT \mapsto (0, 1), \quad TH \mapsto (1, 0), \quad TT \mapsto (1, 1).$$

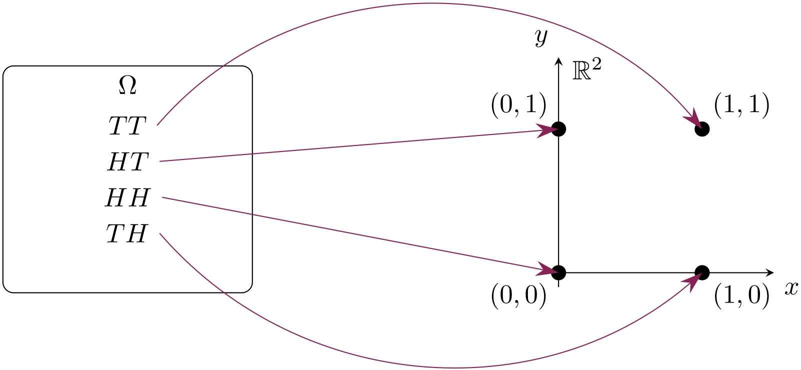


Figure 1: Sample Space and Mapping in $\mathbb{R}^2$ of Two Fair Coin Flips.

2. Discrete Random Vectors

2.1. Joint Distribution (PMF)

Joint Distribution of a Discrete Pair

Definition 2. Joint Distribution

*Let (X, Y) be a **discrete bivariate random vector**, taking values in*

$$D_X \times D_Y \subset \mathbb{R}^2.$$

*The **joint distribution** of (X, Y) is given by*

$$p_{ij} := \mathbb{P}(X = i, Y = j), \qquad i \in D_X,\ j \in D_Y.$$

- The collection $(p_{ij})_{i \in D_X, j \in D_Y}$ completely characterizes the law of (X, Y).
- The sets D_X and D_Y are the supports of X and Y.

Properties of a Joint Distribution

Properties of a Joint Distribution

The joint distribution (p_{ij}) satisfies the following properties:

- ***Non-negativity:***

$$p_{ij} \geq 0 \quad \textit{for all } i \in D_X,\ j \in D_Y.$$

- ***Normalization:***

$$\sum_{i \in D_X} \sum_{j \in D_Y} p_{ij} = 1.$$

- ***Joint probabilities:*** *for any subset $A \subset D_X \times D_Y$,*

$$\mathbb{P}\big((X, Y) \in A\big) = \sum_{(i,j) \in A} p_{ij}.$$

Support of a Discrete Pair

- The **support** of (X, Y) is the set

$$\mathcal{S}_{X,Y} := \{(i, j) \in D_X \times D_Y : p_{ij} > 0\}.$$

- Outside the support, the joint probability is zero:

$$(i, j) \notin \mathcal{S}_{X,Y} \quad \Longrightarrow \quad \mathbb{P}(X = i, Y = j) = 0.$$

- In practice, the support determines:
  - which values (X, Y) can take,
  - which sums or functions of (X, Y) are possible.

Example: Joint Distribution of Two Coin Flips

Example

Let (X, Y) denote the outcomes of two fair coin flips, with

$$0 = \textit{Head}, \quad 1 = \textit{Tail}.$$

- The support is
$$D_X = D_Y = \{0, 1\}.$$
- Since all outcomes are equally likely,
$$p_{ij} = \mathbb{P}(X = i, Y = j) = \frac{1}{4} \quad \text{for all } i, j \in \{0, 1\}.$$
- The joint distribution assigns equal mass to each point of the grid.

What the Joint Distribution Captures

- The joint distribution describes how X and Y **behave together**.
- It allows us to compute:
  - probabilities involving both variables,
  - marginal distributions,
  - conditional distributions,
  - dependence and independence.

2.2. Marginal Distributions

Marginal Distributions: Central Ideas

- A **marginal distribution** describes the behaviour of **one variable alone**.
- It is obtained by **summing out** the other variable from the joint distribution.
- The marginal distribution of X ignores the value taken by Y, and vice versa.

Marginal Distributions

Definition 3. Marginal Distribution

Let (X, Y) be a discrete bivariate random vector with joint distribution

$$p_{ij} = \mathbb{P}(X = i,\, Y = j).$$

*The **marginal distribution of X** is defined by*

$$p_i := \mathbb{P}(X = i) = \sum_{j \in D_Y} p_{ij}.$$

*Similarly, the **marginal distribution of Y** is*

$$q_j := \mathbb{P}(Y = j) = \sum_{i \in D_X} p_{ij}.$$

Basic Properties

Properties of Marginal Distributions

Let $(p_i)_{i \in D_X}$ and $(q_j)_{j \in D_Y}$ denote the marginal distributions of X and Y.

- ***Non-negativity:***

$$p_i \geq 0, \quad q_j \geq 0.$$

- ***Normalization:***

$$\sum_{i \in D_X} p_i = 1, \qquad \sum_{j \in D_Y} q_j = 1.$$

Example: Marginal Distributions for a Two Coin Flips Experiment

Let (X, Y) represent two fair coin flips, with $0 =$ Head and $1 =$ Tail.

- The joint distribution is

$$p_{ij} = \frac{1}{4} \quad \text{for all } i, j \in \{0, 1\}.$$

- Marginal distribution of X:

$$\mathbb{P}(X = 0) = \frac{1}{2}, \qquad \mathbb{P}(X = 1) = \frac{1}{2}.$$

- Marginal distribution of Y:

$$\mathbb{P}(Y = 0) = \frac{1}{2}, \qquad \mathbb{P}(Y = 1) = \frac{1}{2}.$$

From Joint to Marginal

Warning!

Marginal distributions only describe the behaviour of each variable separately. They do not capture how the variables interact with each other.

- The joint distribution $(p_{ij})_{i\in\mathcal{D}_{\mathcal{X}},j\in\mathcal{D}_Y}$ uniquely determines the marginal distributions $(p_i)_{i\in\mathcal{D}_X}$ and $(q_j)_{j\in\mathcal{D}_Y}$.
- The reverse question, whether marginal distributions determine the joint law, **is usually not true**, and will be addressed later.

2.3. Joint Distribution Table

Joint Distribution Table

Definition 4. Joint Distribution Table

Let (X, Y) be a discrete bivariate random vector with joint distribution

$$p_{ij} = \mathbb{P}(X = i, Y = j).$$

A ***joint distribution table*** *is a two–way table displaying:*

- *the joint probabilities p_{ij},*
- *the marginal distributions of X and Y.*

- Joint distribution tables provide a **complete and explicit** description of the law of a discrete pair.
- They make it easy to compute marginal distributions, compute conditional distributions, and to visualize dependence between variables.

Structure of a Joint Distribution Table

$Y \backslash X$	i_1	$\cdots$	i_k	Marginal of Y
j_1	$p_{i_1 j_1}$	$\cdots$	$p_{i_k j_1}$	q_{j_1}
$\vdots$	$\vdots$		$\vdots$	$\vdots$
j_ℓ	$p_{i_1 j_\ell}$	$\cdots$	$p_{i_k j_\ell}$	q_{j_ℓ}
Marginal of X	p_{i_1}	$\cdots$	p_{i_k}	1

- Rows correspond to values of Y.
- Columns correspond to values of X.
- Marginals are obtained by summing rows or columns.

Properties of a Joint Distribution Table

Properties of a Joint Distribution Table

Let $(p_{ij})_{i\in\mathcal{D}_X, j\in\mathcal{D}_Y}$ *be the joint distribution of* (X,Y)*, with marginals* $(p_i)_{i\in\mathcal{D}_X}$ *and* $(q_j)_{j\in\mathcal{D}_Y}$.

- ***Non-negativity:***

$$p_{ij} \geq 0.$$

- ***Normalization:***

$$\sum_i \sum_j p_{ij} = 1.$$

- ***Marginals from the table:***

$$p_i = \sum_j p_{ij}, \qquad q_j = \sum_i p_{ij}.$$

Example: Joint Distribution Table

Example

Let (X, Y) *take values in* $\{-1, 0, 1\} \times \{0, 1, 2\}$*, with joint probabilities:*

$$p_{-1,0} = p_{-1,1} = p_{0,0} = \frac{1}{6}, \quad p_{0,1} = p_{0,2} = \frac{1}{8}, \quad p_{1,1} = \frac{1}{4},$$

and $p_{ij} = 0$ *otherwise.*

$Y \backslash X$	-1	0	1	q_j
0	$\frac{1}{6}$	$\frac{1}{6}$	0	$\frac{1}{3}$
1	$\frac{1}{6}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{13}{24}$
2	0	$\frac{1}{8}$	0	$\frac{1}{8}$
p_i	$\frac{1}{3}$	$\frac{5}{12}$	$\frac{1}{4}$	1

Example: Supports and Marginal Distributions

- The support of X is $\mathcal{D}_X = \{-1, 0, 1\}$, and the support of Y is $\mathcal{D}_Y = \{0, 1, 2\}$.
- The marginal distribution of X:

$$\mathbb{P}(X = -1) = \tfrac{1}{3}, \quad \mathbb{P}(X = 0) = \tfrac{5}{12}, \quad \mathbb{P}(X = 1) = \tfrac{1}{4}.$$

- The marginal distribution of Y:

$$\mathbb{P}(Y = 0) = \tfrac{1}{3}, \quad \mathbb{P}(Y = 1) = \tfrac{13}{24}, \quad \mathbb{P}(Y = 2) = \tfrac{1}{8}.$$

2.4. Conditional Distributions

Conditional Distribution (Discrete Case)

The conditional distribution of X given $Y = j$ describes the behaviour of X **when the value of Y is known**.

Definition 5. Conditional Distribution

Let (X, Y) be a discrete bivariate random vector with joint distribution

$$p_{ij} = \mathbb{P}(X = i,\, Y = j),$$

and marginal distribution of Y given by $q_j = \mathbb{P}(Y = j)$, with $q_j > 0$.
*The **conditional distribution** of X given $Y = j$ is*

$$\mathbb{P}(X = i \mid Y = j) = \frac{\mathbb{P}(X = i, Y = j)}{\mathbb{P}(Y = j)} = \frac{p_{ij}}{q_j}, \qquad i \in \mathcal{D}_X.$$

Conditional Distribution

- The conditional distribution of a discrete random variable, is obtained by:
  - fixing a row of the joint distribution table,
  - normalizing it so that probabilities sum to 1.
- Conditional distributions are defined only when $\mathbb{P}(Y = j) > 0$.

Properties of Conditional Distributions

Properties

For a fixed $j \in \mathcal{D}_Y$ such that $q_j > 0$, the conditional distribution $\mathbb{P}(X = i \mid Y = j)$ satisfies:

- ***Non-negativity:***
$$\mathbb{P}(X = i \mid Y = j) \geq 0.$$
- ***Normalization:***
$$\sum_{i \in \mathcal{D}_X} \mathbb{P}(X = i \mid Y = j) = 1.$$

Example: Conditional Distribution

Example

Consider the joint distribution table introduced earlier.

$Y \backslash X$	-1	0	1	q_j
0	$\frac{1}{6}$	$\frac{1}{6}$	0	$\frac{1}{3}$
1	$\frac{1}{6}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{13}{24}$
2	0	$\frac{1}{8}$	0	$\frac{1}{8}$
p_i	$\frac{1}{3}$	$\frac{5}{12}$	$\frac{1}{4}$	1

- *Conditional distribution of X given $Y = 1$:*

$$\mathbb{P}(X = -1 \mid Y = 1) = \frac{1/6}{13/24} = \frac{4}{13},$$

$$\mathbb{P}(X = 0 \mid Y = 1) = \frac{1/8}{13/24} = \frac{3}{13},$$

$$\mathbb{P}(X = 1 \mid Y = 1) = \frac{1/4}{13/24} = \frac{6}{13}.$$

- *These probabilities sum to 1.*

Example: Conditional Distribution (Symmetry)

Example

Consider the same joint distribution table.

$Y\backslash X$	-1	0	1	q_j
0	$\frac{1}{6}$	$\frac{1}{6}$	0	$\frac{1}{3}$
1	$\frac{1}{6}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{13}{24}$
2	0	$\frac{1}{8}$	0	$\frac{1}{8}$
p_i	$\frac{1}{3}$	$\frac{5}{12}$	$\frac{1}{4}$	1

- *Conditional distribution of Y given $X = -1$:*

$$\mathbb{P}(Y = 0 \mid X = -1) = \frac{1/6}{1/3} = \frac{1}{2},$$

$$\mathbb{P}(Y = 1 \mid X = -1) = \frac{1/6}{1/3} = \frac{1}{2},$$

$$\mathbb{P}(Y = 2 \mid X = -1) = \frac{0}{1/3} = 0.$$

- *These probabilities sum to 1.*

2.5. Independence

Independence: Definition

In Chapter 5, we defined the independence of random variables. Recall this definition and apply it to two discrete random variables.

Definition 6. Independence

Two discrete random variables X and Y are said to be ***independent*** *if, for all $i \in \mathcal{D}_X$ and $j \in \mathcal{D}_Y$,*

$$\mathbb{P}(X = i, Y = j) = \mathbb{P}(X = i)\,\mathbb{P}(Y = j),$$

that is,

$$p_{ij} = p_i\, q_j.$$

Independence and Joint Distribution Tables

- **If X and Y are independent**, the joint distribution table is completely determined by the marginals.
- Each entry is the product of its row and column marginals:

$$p_{ij} = p_i q_j.$$

- Independence imposes a very strong structure on the joint law.

Example: Independent Random Variables (1/2)

Example

Let X and Y be two independent fair coin flips, with $0 =$ Head and $1 =$ Tail.
We know the marginal distributions:

$$\mathbb{P}(X = 1) = \mathbb{P}(X = 0) = \frac{1}{2} \quad \text{and} \quad \mathbb{P}(Y = 1) = \mathbb{P}(Y = 0) = \frac{1}{2}.$$

We can write:

$Y \backslash X$	0	1	q_j
0			$\frac{1}{2}$
1			$\frac{1}{2}$
p_i	$\frac{1}{2}$	$\frac{1}{2}$	1

Example: Independent Random Variables (2/2)

Since X and Y are independent, the joint distribution table can be completely determined by the marginals:

$$p_{ij} = p_i q_j \quad \text{for all} \quad i, j$$

- For example: $p_{10} = p_1 \cdot q_0 = \frac{1}{2} \cdot \frac{1}{2}$

$Y \backslash X$	0	1	q_j
0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
1	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
p_i	$\frac{1}{2}$	$\frac{1}{2}$	1

Example: Dependent Random Variables

Example

*Let (X, Y) have the same marginals as before, **but joint distribution**:*

$Y \backslash X$	0	1	q_j
0	$\frac{1}{2}$	0	$\frac{1}{2}$
1	0	$\frac{1}{2}$	$\frac{1}{2}$
	$\frac{1}{2}$	$\frac{1}{2}$	1

- Marginal distributions are unchanged (each variable looks like a fair coin).
- However, **X and Y are not** independent. This might be the case if we consider, for example an experiment where we toss a coin and define:
  - X = "outcome of the coin toss",
  - Y = "outcome of the same coin toss".

Independence and Conditional Distributions

- Recall that the **distribution of X conditional on Y**, for all $i \in \mathcal{D}_X$ and $j \in \mathcal{D}_Y$ such that $q_j > 0$ writes

$$\mathbb{P}(X = i \mid Y = j) = \frac{\mathbb{P}(X = i, Y = j)}{\mathbb{P}(Y = j)} = \frac{p_{ij}}{q_j}, \qquad i \in \mathcal{D}_X.$$

- If **X and Y are independent**, then it simplifies to

$$\mathbb{P}(X = i \mid Y = j) = \frac{\mathbb{P}(X = i)\mathbb{P}(Y = j)}{\mathbb{P}(Y = j)} = \mathbb{P}(X = i) = p_i$$

- Similarly, for all i such that $p_i > 0$,

$$\mathbb{P}(Y = j \mid X = i) = \mathbb{P}(Y = j).$$

Example: Independence and Conditional Distributions (1/3)

Example

Let X and Y be two independent random variables such that:

$$X \in \{-1, 0, 1\}, \qquad Y \in \{0, 1, 2\},$$

with marginal distributions

$$\mathbb{P}(X=-1) = \tfrac{1}{3}, \quad \mathbb{P}(X=0) = \tfrac{5}{12}, \quad \mathbb{P}(X=1) = \tfrac{1}{4},$$

$$\mathbb{P}(Y=0) = \tfrac{1}{3}, \quad \mathbb{P}(Y=1) = \tfrac{13}{24}, \quad \mathbb{P}(Y=2) = \tfrac{1}{8}.$$

Since X and Y are independent,

$$p_{ij} = p_i q_j.$$

$Y \backslash X$	-1	0	1	q_j
0	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{12}$	$\frac{1}{3}$
1	$\frac{13}{72}$	$\frac{65}{288}$	$\frac{13}{96}$	$\frac{13}{24}$
2	$\frac{1}{24}$	$\frac{5}{96}$	$\frac{1}{32}$	$\frac{1}{8}$
p_i	$\frac{1}{3}$	$\frac{5}{12}$	$\frac{1}{4}$	1

Example: Independence and Conditional Distributions (2/3)

Example

(X and Y are independent.) Fix any row (e.g. $Y = 1$). Then:

$$\mathbb{P}(X = -1 \mid Y = 1) = \frac{13/72}{13/24} = \frac{1}{3},$$

$$\mathbb{P}(X = 0 \mid Y = 1) = \frac{65/288}{13/24} = \frac{5}{12},$$

$$\mathbb{P}(X = 1 \mid Y = 1) = \frac{13/96}{13/24} = \frac{1}{4}.$$

$Y \backslash X$	-1	0	1	q_j
0	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{12}$	$\frac{1}{3}$
1	$\frac{13}{72}$	$\frac{65}{288}$	$\frac{13}{96}$	$\frac{13}{24}$
2	$\frac{1}{24}$	$\frac{5}{96}$	$\frac{1}{32}$	$\frac{1}{8}$
p_i	$\frac{1}{3}$	$\frac{5}{12}$	$\frac{1}{4}$	1

which is exactly the ***marginal distribution of X***.

Rows are scaled versions of the ***marginal of X***.

Example: Independence and Conditional Distributions (3/3)

Example

(X and Y are independent.) Fix any column (e.g. $X = 1$). Then:

$$\mathbb{P}(Y = 0 \mid X = -1) = \frac{1/9}{1/3} = \frac{1}{3}$$

$$\mathbb{P}(Y = 1 \mid X = -1) = \frac{13/72}{1/3} = \frac{13}{24}$$

$$\mathbb{P}(Y = 2 \mid X = -1) = \frac{1/24}{1/3} = \frac{1}{8}$$

which is exactly the ***marginal distribution of Y***.

$Y \backslash X$	-1	0	1	q_j
0	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{12}$	$\frac{1}{3}$
1	$\frac{13}{72}$	$\frac{65}{288}$	$\frac{13}{96}$	$\frac{13}{24}$
2	$\frac{1}{24}$	$\frac{5}{96}$	$\frac{1}{32}$	$\frac{1}{8}$
p_i	$\frac{1}{3}$	$\frac{5}{12}$	$\frac{1}{4}$	1

Columns are scaled versions of the ***marginal of Y***.

2.6. Conditional Moments and Regression Function

Conditional Expectation (Discrete Case)

Definition 7. Conditional Expectation

Let (X, Y) be a discrete bivariate random vector. For any $i \in \mathcal{D}_X$ such that $\mathbb{P}(X = i) > 0$, the ***conditional expectation of Y given $X = i$*** *is*

$$\mathbb{E}(Y \mid X = i) = \sum_{j \in \mathcal{D}_Y} j\, \mathbb{P}(Y = j \mid X = i).$$

- It is the expectation of Y under the conditional distribution.
- It is a *number*, not a random variable.

Conditional Moments

- More generally, for any function g, the conditional moment of $g(Y)$ given $X = i$ is

$$\mathbb{E}(g(Y) \mid X = i) = \sum_{j \in \mathcal{D}_Y} g(j)\, \mathbb{P}(Y = j \mid X = i).$$

- Important special cases:
  - Conditional second moment:
  $$\mathbb{E}(Y^2 \mid X = i),$$
  - Conditional variance:
  $$\mathrm{Var}(Y \mid X = i) = \mathbb{E}(Y^2 \mid X = i) - \left[\mathbb{E}(Y \mid X = i)\right]^2.$$

Example: Conditional Expectation of Y Given X

Example

Using the joint distribution table introduced earlier, compute $\mathbb{E}(Y \mid X = -1)$.

$Y \setminus X$	-1	0	1	q_j
0	$\frac{1}{6}$	$\frac{1}{6}$	0	$\frac{1}{3}$
1	$\frac{1}{6}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{13}{24}$
2	0	$\frac{1}{8}$	0	$\frac{1}{8}$
p_i	$\frac{1}{3}$	$\frac{5}{12}$	$\frac{1}{4}$	1

- *Marginal distribution of X:*

$$\mathbb{P}(X = -1) = \frac{1}{3}.$$

- *Conditional distribution of Y given $X = -1$:*

$$\mathbb{P}(Y = 0 \mid X = -1) = \frac{\mathbb{P}(Y = 0, X = -1)}{\mathbb{P}(X = -1)} = \frac{1/6}{1/3} = \frac{1}{2},$$

$$\mathbb{P}(Y = 1 \mid X = -1) = \frac{\mathbb{P}(Y = 1, X = -1)}{\mathbb{P}(X = -1)} = \frac{1}{2},$$

$$\mathbb{P}(Y = 2 \mid X = -1) = \frac{\mathbb{P}(Y = 2, X = -1)}{\mathbb{P}(X = -1)} = 0.$$

- *Therefore,*

$$\mathbb{E}(Y \mid X = -1) = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} + 2 \cdot 0 = \frac{1}{2}.$$

Regression Function of Y on X

Regression Function

Let (X, Y) be a discrete bivariate random vector. The ***regression function of*** Y ***on*** X *is defined by*

$$r_Y(i) := \mathbb{E}(Y \mid X = i), \qquad i \in \mathcal{D}_X.$$

- It assigns to each value of X the conditional mean of Y.
- It describes how the average value of Y depends on X.
- For a fixed value $X = i$, $r_Y(i)$ is the best prediction (in mean) of Y given $X = i$.

Example: Computing the Regression Function

Using the joint distrib. table introduced earlier, compute $r_Y(i) = \mathbb{E}(Y \mid X = i)$ *for* $i = -1, 0, 1$.

$Y \backslash X$	-1	0	1	q_j
0	$\frac{1}{6}$	$\frac{1}{6}$	0	$\frac{1}{3}$
1	$\frac{1}{6}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{13}{24}$
2	0	$\frac{1}{8}$	0	$\frac{1}{8}$
p_i	$\frac{1}{3}$	$\frac{5}{12}$	$\frac{1}{4}$	1

- *For* $X = -1$*:* $r_Y(-1) = \frac{1}{2}$.
- *For* $X = 0$*:*

  $\mathbb{P}(Y = 0 \mid X = 0) = \frac{2}{5}$, $\mathbb{P}(Y = 1 \mid X = 0) = \frac{3}{10}$, $\mathbb{P}(Y = 2 \mid X = 0) = \frac{3}{10}$,

  $$r_Y(0) = 0 \cdot \frac{2}{5} + 1 \cdot \frac{3}{10} + 2 \cdot \frac{3}{10} = \frac{9}{10}.$$
- *For* $X = 1$*:*

  $$\mathbb{P}(Y = 1 \mid X = 1) = 1,$$

  $$r_Y(1) = 1$$

2.7. Moments of Functions of (X, Y) (Covariance)

Expectation of a Function of (X, Y)

Definition 8. Expectation of a Function

Let (X, Y) be a discrete bivariate random vector and let $h : \mathbb{R}^2 \to \mathbb{R}$. The expectation of $h(X, Y)$ is defined by

$$\mathbb{E}[h(X,Y)] = \sum_{i \in \mathcal{D}_X} \sum_{j \in \mathcal{D}_Y} h(i,j)\, \mathbb{P}(X = i, Y = j).$$

- This generalizes the expectation of a single random variable.
- All moments of (X, Y) are obtained as special cases.

Special Cases: Expectation

Using this definition, we can revisit some results as special cases:

- Mean of a component: **with** $h(x, y) = x$,

$$\begin{aligned}
\mathbb{E}[h(X,Y)] &= \sum_{i\in\mathcal{D}_X}\sum_{j\in\mathcal{D}_Y} h(ij)\,\mathbb{P}(X=i,Y=j)\\
&= \sum_{i\in\mathcal{D}_X} i \sum_{j\in\mathcal{D}_Y} \mathbb{P}(X=i,Y=j)\\
&= \sum_{i\in\mathcal{D}_X} i\mathbb{P}(X=i)\\
&= \mathbb{E}(X).
\end{aligned}$$

Special Cases: Covariance

- Covariance: **with** $h(x,y) = (x - \mathbb{E}(X))(y - \mathbb{E}(Y))$,

$$\begin{aligned}
\mathbb{E}[h(X,Y)] &= \sum_{i \in \mathcal{D}_X} \sum_{j \in \mathcal{D}_Y} h(ij)\, \mathbb{P}(X = i, Y = j) \\
&= \sum_i \sum_j (i - \mathbb{E}(X))(j - \mathbb{E}(Y))\, \mathbb{P}(X = i, Y = j) \\
&= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) \\
&= \mathrm{Cov}(X,Y).
\end{aligned}$$

Remark: Independence Implies Zero Covariance

If X and Y are independent, $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$. In such a case, as seen in Chapter 5, we have:

$$\mathrm{Cov}(X,Y) = 0$$

***Warning**: the converse is false!*

Counterexample: Zero Covariance Does Not Imply Independence

Let X be a random variable taking values in $\{-1, 0, 1\}$ with

$$\mathbb{P}(X = -1) = \mathbb{P}(X = 0) = \mathbb{P}(X = 1) = \tfrac{1}{3},$$

and define $Y = X^2$.

$Y \backslash X$	-1	0	1	q_j
0	0	$\frac{1}{3}$	0	$\frac{1}{3}$
1	$\frac{1}{3}$	0	$\frac{1}{3}$	$\frac{2}{3}$
p_i	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	1

- *Marginal expectations:*

$$\mathbb{E}(X) = 0, \qquad \mathbb{E}(Y) = \mathbb{E}(X^2) = \tfrac{2}{3}.$$

- $\mathbb{E}(XY) = \mathbb{E}(X^3) = 0$
- *Hence,*

$$\mathrm{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = 0.$$

- *However, joint probabilities do not factorize:*

$$\mathbb{P}(X = -1, Y = 1) = \frac{1}{3} \neq \frac{1}{3}\frac{2}{3} = \mathbb{P}(X = -1)\mathbb{P}(Y = 1)$$

- *Zero covariance does not imply independence.*

2.8. Distribution of a Sum $S = X + Y$ (Discrete Case)

Sum of Two Discrete Random Variables

Definition 9. Sum of Two Discrete Random Variables

Let (X, Y) be a discrete bivariate random vector. We define a new random variable

$$S = X + Y.$$

- S is a function of the pair (X, Y).
- Its distribution is fully determined by the joint distribution of (X, Y).
- The support of S is

$$\mathcal{D}_S = \{\, i + j \;:\; i \in \mathcal{D}_X,\ j \in \mathcal{D}_Y \,\}.$$

- Different pairs (i, j) may lead to the same value of S.
- Therefore, probabilities must be obtained by summing over all pairs (i, j) such that $i + j = s$.

Distribution of the Sum of Two Discrete Random Variables

Distribution of $S = X + Y$

For any $s \in \mathcal{D}_S$,

$$\mathbb{P}(S = s) = \sum_{\substack{i \in \mathcal{D}_X,\, j \in \mathcal{D}_Y \\ i+j=s}} \mathbb{P}(X = i, Y = j).$$

- This formula holds whether or not X and Y are independent.

Example

Recall a previous example

$Y \backslash X$	-1	0	1	q_j
0	$\frac{1}{6}$	$\frac{1}{6}$	0	$\frac{1}{3}$
1	$\frac{1}{6}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{13}{24}$
2	0	$\frac{1}{8}$	0	$\frac{1}{8}$
p_i	$\frac{1}{3}$	$\frac{5}{12}$	$\frac{1}{4}$	1

The support is $\mathcal{D}_S = \{i + j : i \in \{-1, 0, 1\},\ j \in \{0, 1, 2\}\} = \{-1, 0, 1, 2, 3\}$.

- *Example for* $s = \mathbf{0}$*:* $(-1, 1)$ *or* $(0, 0)$

$$\mathbb{P}(S = 0) = \mathbb{P}(X = -1, Y = 1) + \mathbb{P}(X = 0, Y = 0)$$
$$= \frac{1}{6} + \frac{1}{6} = \frac{1}{3}.$$

Therefore:

S	*Outcomes* (X, Y) *such that* $X + Y = S$	$\mathbb{P}(S = s)$
-1	$(-1, 0)$	$\frac{1}{6}$
0	$(-1, 1),\ (0, 0)$	$\frac{1}{6} + \frac{1}{6} = \frac{1}{3}$
1	$(0, 1)$	$\frac{1}{8}$
2	$(0, 2),\ (1, 1)$	$\frac{1}{8} + \frac{1}{4} = \frac{3}{8}$
3	$(1, 2)$	0

3. Exercises

3.1. Exercise 1

Exercise 1: Joint Distribution Table

Let (X, Y) be a pair of discrete random variables with joint distribution:

$Y \backslash X$	-1	0	1
-1	0	$\frac{1}{8}$	0
0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
1	0	$\frac{1}{8}$	$\frac{1}{8}$
2	0	0	$\frac{1}{4}$

1. *Determine the marginal distributions of X and Y.*
2. *Let $D = |X - Y|$. Determine the distribution of D and compute $\mathbb{E}(D)$.*
3. *Compute $\mathrm{Cov}(X, Y)$.*
4. *Compute $\mathbb{E}(Y \mid X = 1)$.*

Solutions: Marginals

- The marginal distributions are obtained using: $\begin{cases} p_i = \sum_j p_{ij} & \text{(marginal of } X\text{)} \\ q_j = \sum_i p_{ij} & \text{(marginal of } Y\text{)} \end{cases}$
- The marginal distribution of X:

$$\mathbb{P}(X=-1)=\tfrac{1}{8}, \qquad \mathbb{P}(X=0)=\tfrac{3}{8}, \qquad \mathbb{P}(X=1)=\tfrac{1}{2}.$$

- The marginal distribution of Y:

$$\mathbb{P}(Y=-1)=\tfrac{1}{8}, \qquad \mathbb{P}(Y=0)=\tfrac{3}{8}, \qquad \mathbb{P}(Y=1)=\tfrac{1}{4}, \qquad \mathbb{P}(Y=2)=\tfrac{1}{4}.$$

The marginal distributions can be added to the joint distribution table:

$Y \backslash X$	-1	0	1	q_j
-1	0	$\frac{1}{8}$	0	$\frac{1}{8}$
0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{3}{8}$
1	0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$
2	0	0	$\frac{1}{4}$	$\frac{1}{4}$
p_j	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{2}$	1

Solution: Distribution of $D = |X - Y|$

- Nonzero outcomes yield only $D \in \{0, 1\}$.

- $D = 0$ occurs when $X = Y$, i.e. $(0, 0)$ or $(1, 1)$:

$$\mathbb{P}(D = 0) = \mathbb{P}(X = 0, Y = 0) + \mathbb{P}(X = 1, Y = 1) = \frac{1}{8} + \frac{1}{8} = \frac{1}{4}.$$

- Hence

$$\mathbb{P}(D = 1) = 1 - \mathbb{P}(D = 0) = \frac{3}{4}.$$

- Therefore

$$\mathbb{E}(D) = 0 \cdot \frac{1}{4} + 1 \cdot \frac{3}{4} = \frac{3}{4}.$$

Solution: $\text{Cov}(X, Y)$

- Firs, compute $\mathbb{E}(X)$ and $\mathbb{E}(Y)$:

$$\mathbb{E}(X) = (-1)\tfrac{1}{8} + 0 \cdot \tfrac{3}{8} + 1 \cdot \tfrac{1}{2} = \tfrac{3}{8}, \qquad \mathbb{E}(Y) = (-1)\tfrac{1}{8} + 0 \cdot \tfrac{3}{8} + 1 \cdot \tfrac{1}{4} + 2 \cdot \tfrac{1}{4} = \tfrac{5}{8}.$$

- Compute

$$\mathbb{E}(XY) = \sum_{i \in \mathcal{D}_X} \sum_{j \in \mathcal{D}_Y} ij \, \mathbb{P}(X = i, Y = j)$$

Only $(X, Y) = (1, 1)$ and $(1, 2)$ contribute:

$$\mathbb{E}(XY) = 1 \cdot \mathbb{P}(1, 1) + 2 \cdot \mathbb{P}(1, 2) = 1 \cdot \tfrac{1}{8} + 2 \cdot \tfrac{1}{4} = \tfrac{5}{8}.$$

$Y \backslash X$	-1	0	1	q_j
-1	0	$\frac{1}{8}$	0	$\frac{1}{8}$
0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{3}{8}$
1	0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$
2	0	0	$\frac{1}{4}$	$\frac{1}{4}$
p_j	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{2}$	1

- Hence,

$$\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \tfrac{5}{8} - \tfrac{3}{8} \cdot \tfrac{5}{8} = \tfrac{25}{64}.$$

Solution $\mathbb{E}(Y \mid X = 1)$

$Y \backslash X$	-1	0	1	q_j
-1	0	$\frac{1}{8}$	0	$\frac{1}{8}$
0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{3}{8}$
1	0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$
2	0	0	$\frac{1}{4}$	$\frac{1}{4}$
p_j	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{2}$	1

Compute:

$$\mathbb{E}(Y \mid X = 1) = \sum_{j \in \mathcal{D}_Y} j \, \mathbb{P}(Y = j \mid X = 1).$$

- First note $\mathbb{P}(X = 1) = \frac{1}{2}$.
- Conditional probabilities (from the **column** $X = 1$):

$$\mathbb{P}(Y = -1 \mid X = 1) = \frac{\mathbb{P}(Y = -1, X = 1)}{\mathbb{P}(X = 1)} = \frac{0}{1/2} = 0,$$

$$\mathbb{P}(Y = 0 \mid X = 1) = \frac{\mathbb{P}(Y = 0, X = 1)}{\mathbb{P}(X = 1)} = \frac{1/8}{1/2} = \frac{1}{4},$$

$$\mathbb{P}(Y = 1 \mid X = 1) = \frac{\mathbb{P}(Y = 1, X = 1)}{\mathbb{P}(X = 1)} = \frac{1/8}{1/2} = \frac{1}{4},$$

$$\mathbb{P}(Y = 2 \mid X = 1) = \frac{\mathbb{P}(Y = 2, X = 1)}{\mathbb{P}(X = 1)} = \frac{1/4}{1/2} = \frac{1}{2}.$$

- Therefore

$$\mathbb{E}(Y \mid X = 1) = -1 \cdot 0 + 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} + 2 \cdot \frac{1}{2} = \frac{5}{4}.$$

3.2. Exercise 2

Exercise 2: Linear Transformations

Let X and Y be two independent Bernoulli random variables with parameter $p \in (0, 1)$.
Define

$$U = X + Y, \qquad V = X - Y.$$

1. *Determine the joint distribution of the pair (U, V).*
2. *Compute $\mathrm{Cov}(U, V)$.*
3. *Are the random variables U and V independent?*

Solution: Law of (U, V)

The setup gives X, Y independent $\mathcal{B}(p)$ (hence $\mathcal{D}_X = \mathcal{D}_Y = \{0, 1\}$, and

$$U = X + Y, \qquad V = X - Y.$$

Since X and Y are independent Bernoulli(p),

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y).$$

Mapping $(X, Y) \mapsto (U, V)$

(X, Y)	(U, V)	$\mathbb{P}$
$(0, 0)$	$(0, 0)$	$(1-p)^2$
$(0, 1)$	$(1, -1)$	$p(1-p)$
$(1, 0)$	$(1, 1)$	$p(1-p)$
$(1, 1)$	$(2, 0)$	p^2

Solution: $\mathrm{Cov}(U, V)$

- Compute UV:

$$UV = (X+Y)(X-Y) = X^2 - Y^2.$$

- Since $X^2 = X$ and $Y^2 = Y$ (Bernoulli),

$$\mathbb{E}(UV) = \mathbb{E}(X^2) - \mathbb{E}(Y^2) = \mathbb{E}(X) - \mathbb{E}(Y) = p - p = 0.$$

- Also $\mathbb{E}(V) = \mathbb{E}(X-Y) = p - p = 0$, hence

$$\mathrm{Cov}(U, V) = \mathbb{E}(UV) - \mathbb{E}(U)\mathbb{E}(V) = 0.$$

Solution: Are U and V independent?

Reminder

Recall that two random variables U and V are independent if and only if, for all $u \in \mathcal{D}_U$ and all $v \in \mathcal{D}_V$:

$$\mathbb{P}(U = u, V = v) = \mathbb{P}(U = u)\mathbb{P}(V = v).$$

- We have $\mathbb{P}(U = 2, V = 0) = \mathbb{P}(X = 1, Y = 1) = p^2$.
- But

$$\mathbb{P}(U = 2) = p^2, \qquad \mathbb{P}(V = 0) = \mathbb{P}(X = Y) = \underbrace{p^2}_{\text{both } 1} + \underbrace{(1-p)^2}_{\text{both } 0}.$$

- Therefore

$$\mathbb{P}(U = 2)\mathbb{P}(V = 0) = p^2(p^2 + (1-p)^2) \neq p^2$$

(for $p \in (0, 1)$). Hence U and V are not independent.

3.3. Exercise 3

Exercise

Let N be the number of users arriving at a post office, where

$$N \sim \mathcal{P}(\lambda), \quad \lambda > 0.$$

Each user performs exactly one operation:

- *with probability $p \in (0,1)$, they send a letter;*
- *with probability $1-p$, they perform another operation.*

Operations are independent across users.
Let X be the number of users sending a letter, and define $Y = N - X$.

1. *Determine the distribution of $X \mid N = n$.*
2. *Determine the joint distribution of (X, N).*
3. *Determine the distribution of X, and deduce the distribution of Y.*
4. *Show that X and Y are independent.*
5. *Compute $\mathrm{Cov}(X, N)$.*

Solution: Law of $X \mid N = n$

Conditionally on $N = n$, X counts the number of successes among n independent trials:

$$X \mid (N = n) \;\sim\; \mathrm{Bin}(n, p).$$

So, for $0 \leq k \leq n$,

$$\mathbb{P}(X = x \mid N = n) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Given that $N = n$ people arrived at the post office, X counts the number of people among those $X = n$ that will send a letter.

Solution: Joint law of (X, N)

Since $N \sim \mathcal{P}(\lambda)$, for $n \in \mathbb{N}$,

$$\mathbb{P}(N = n) = e^{-\lambda} \frac{\lambda^n}{n!}.$$

For $0 \leq k \leq n$,

$$\mathbb{P}(X = k, N = n) = \mathbb{P}(N = n)\, \mathbb{P}(X = k \mid N = n).$$

Thus,

$$\mathbb{P}(X = k, N = n) = e^{-\lambda} \frac{\lambda^n}{n!} \binom{n}{k} p^k (1-p)^{n-k} = e^{-\lambda} \frac{\lambda^n}{k!(n-k)!} p^k (1-p)^{n-k}.$$

Solution: Marginal law of X

For $k \in \mathbb{N}$,

$$\begin{aligned}
\mathbb{P}(X = k) &= \sum_{n \geq k} \mathbb{P}(X = k, N = n) = e^{-\lambda} \frac{p^k}{k!} \sum_{n \geq k} \frac{\lambda^n}{(n-k)!} (1-p)^{n-k} \\
&= e^{-\lambda} \frac{(\lambda s p)^k}{k!} \sum_{n \geq k} \frac{\lambda^{n-k}}{(n-k)!} (1-p)^{n-k} \\
&= e^{-\lambda} \frac{(\lambda p)^k}{k!} \sum_{m \geq 0} \lambda^m \frac{1}{m!} (1-p)^m \qquad (\text{with } \ = n - k) \\
&= e^{-\lambda} \frac{(\lambda p)^k}{k!} \sum_{m \geq 0} \frac{(\lambda(1-p))^m}{m!} \\
&= e^{-\lambda} \frac{(\lambda p)^k}{k!} e^{\lambda(1-p)} \qquad (\text{power series expansion: } e^x = \sum_{n=0}^{+\infty} \frac{x^n}{n!}) \\
&= e^{-\lambda p} \frac{(\lambda p)^k}{k!}.
\end{aligned}$$

Hence, $X \sim \mathcal{P}(\lambda p)$.

Solution: Marginal Law of Y

Each of the N users either:

- sends "letter" with probability $p \Rightarrow$ counted by X,
- performs "another operation" with probability $1-p \Rightarrow$ counted by $Y = N - X$.

By the same computation as for X (replace p by $1-p$), for $\ell \in \mathbb{N}$,

$$\mathbb{P}(Y = \ell) = e^{-\lambda(1-p)} \frac{(\lambda(1-p))^{\ell}}{\ell!}.$$

Hence,

$$Y \sim \mathcal{P}(\lambda(1-p)).$$

Solution: Independence of X and Y (1/2)

Recall $Y = N - X$. Hence, for $k, l \in \mathbb{N}$,

$$\{X = k,\, Y = l\} = \{X = k,\, N = k + l\}, \quad \text{so} \quad \mathbb{P}(X = k, Y = l) = \mathbb{P}(X = k, N = k + l).$$

Earlier, we found that for $0 \leq k \leq n$,

$$\mathbb{P}(X = k, N = n) = e^{-\lambda} \frac{\lambda^n}{k!(n-k)!} p^k (1-p)^{n-k}.$$

Taking $n = k + l$, we obtain

$$\mathbb{P}(X = k, Y = l) = e^{-\lambda} \frac{\lambda^{k+l}}{k!\, l!} p^k (1-p)^l.$$

Finally, since $e^{-\lambda} = e^{-\lambda p} e^{-\lambda(1-p)}$,

$$\mathbb{P}(X = k, Y = l) = \underbrace{\left(e^{-\lambda p} \frac{(\lambda p)^k}{k!}\right)}_{\mathbb{P}(X=k)} \underbrace{\left(e^{-\lambda(1-p)} \frac{(\lambda(1-p))^l}{l!}\right)}_{\mathbb{P}(Y=l)}.$$

Solution: Independence of X and Y (2/2)

Therefore,

$$\mathbb{P}(X = k, Y = l) = \mathbb{P}(X = k)\mathbb{P}(Y = l) \quad \forall k, l \in \mathbb{N},$$

hence X and Y are independent.

Solution: $\text{Cov}(X, N)$

Since $N = X + Y$,

$$\text{Cov}(X, N) = \text{Cov}(X, X + Y) = \text{Cov}(X, X) + \text{Cov}(X, Y) = \text{Var}(X) + \text{Cov}(X, Y).$$

But X and Y are independent, so $\text{Cov}(X, Y) = 0$. Therefore

$$\text{Cov}(X, N) = \text{Var}(X) = \lambda p.$$

4. Continuous Random Pairs

Continuous Random Pairs

- We now consider a random vector (X, Y) taking values in $\mathbb{R}^2$.
- In the continuous case, the distribution of the pair cannot be described by a probability table.
- Instead, it is characterized by its **joint distribution function**.

4.1. Joint Distribution Function

Definition: Joint Distribution Function

Definition 10. Joint CDF

The joint distribution function of (X, Y) is the function

$$F_{X,Y}(x, y) = \mathbb{P}(X \leq x,\ Y \leq y), \qquad (x, y) \in \mathbb{R}^2.$$

It completely characterizes the law of (X, Y).

Basic Properties of $F_{X,Y}$

Definition 11. Basic Properties of $F_{X,Y}$

For all $(x, y) \in \mathbb{R}^2$, the function $F_{X,Y}$ satisfies:

1. ***Bounds:***
$$0 \leq F_{X,Y}(x, y) \leq 1.$$
2. ***Monotonicity:*** *$F_{X,Y}$ is non-decreasing in each variable separately:*
$$x_1 \leq x_2 \Rightarrow F_{X,Y}(x_1, y) \leq F_{X,Y}(x_2, y),$$
$$y_1 \leq y_2 \Rightarrow F_{X,Y}(x, y_1) \leq F_{X,Y}(x, y_2).$$
3. ***Limits:***
$$\lim_{x \to -\infty} F_{X,Y}(x, y) = 0, \qquad \lim_{y \to -\infty} F_{X,Y}(x, y) = 0,$$
$$\lim_{x \to +\infty,\, y \to +\infty} F_{X,Y}(x, y) = 1.$$

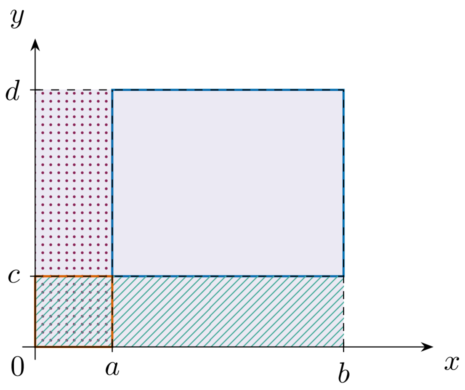
Probability of a Rectangle

For $a < b$ and $c < d$, we can compute:

$$\mathbb{P}(a < X \leq b,\ c < Y \leq d)$$

using inclusion–exclusion:

$$\mathbb{P}(a < X \leq b,\ c < Y \leq d)$$
$$= F_{X,Y}(b,d) - F_{X,Y}(a,d) - F_{X,Y}(b,c)$$
$$+ F_{X,Y}(a,c).$$



This is the two-dimensional analogue of

$$\mathbb{P}(a < X \leq b) = F_X(b) - F_X(a).$$

Independence and the Joint CDF

Definition 12. Independence and the Joint CDF

Recall that X and Y are independent if

$$\mathbb{P}(X \leq x,\ Y \leq y) = \mathbb{P}(X \leq x)\mathbb{P}(Y \leq y) \quad \forall (x, y).$$

Equivalently,

$$F_{X,Y}(x, y) = F_X(x)\, F_Y(y).$$

Independence is characterized directly at the level of the joint CDF.

4.2. Joint Density

Definition: Joint Density

Definition 13. Joint Density

We say that the pair (X, Y) is ***absolutely continuous*** *if there exists a function $f_{X,Y} : \mathbb{R}^2 \to \mathbb{R}_+$ such that, for every set $A \subset \mathbb{R}^2$,*

$$\mathbb{P}((X, Y) \in A) = \iint_A f_{X,Y}(x, y)\, dx\, dy.$$

The function $f_{X,Y}$ is called the ***joint density*** *of (X, Y).*

In particular, $f_{X,Y}$ must satisfy:

$$f_{X,Y}(x, y) \geq 0, \qquad \iint_{\mathbb{R}^2} f_{X,Y}(x, y)\, dx\, dy = 1.$$

Joint Density as a Second Derivative (when smooth)

Definition 14. When the joint CDF is sufficiently smooth

If the joint CDF $F_{X,Y}$ is twice differentiable (in a suitable sense), then

$$f_{X,Y}(x,y) = \frac{\partial^2}{\partial x\,\partial y} F_{X,Y}(x,y).$$

This is the 2D analogue of the 1D identity $f_X(x) = F'_X(x)$ (when the derivative exists).

Independence and Joint Density

Proposition 1. Product form under independence

1. *If X and Y are independent with marginal densities f_X and f_Y, then (X, Y) has joint density*
$$f_{X,Y}(x, y) = f_X(x)\, f_Y(y).$$
2. *Conversely, if $f_{X,Y}(x, y)$ factorizes as $g(x)h(y)$ with g, h densities, then X and Y are independent.*

Proof: Independence $\Rightarrow$ Product Density

Assume X and Y are independent and admit marginal densities f_X and f_Y.

For any $x, y \in \mathbb{R}$,

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

By independence,

$$\mathbb{P}(X \leq x,\ Y \leq y) = \mathbb{P}(X \leq x)\mathbb{P}(Y \leq y) = F_X(x)\, F_Y(y).$$

Assuming the CDFs are differentiable (so densities exist),

$$f_{X,Y}(x, y) = \frac{\partial^2}{\partial x\, \partial y} F_{X,Y}(x, y) = \frac{\partial^2}{\partial x\, \partial y}\big(F_X(x) F_Y(y)\big) = F'_X(x)\, F'_Y(y) = f_X(x)\, f_Y(y).$$

$\square$

Proof: Product Density $\Rightarrow$ Independence (1/2)

Assume (X, Y) has a joint density of the form

$$f_{X,Y}(x, y) = g(x)\, h(y),$$

where g and h are densities on $\mathbb{R}$.

Compute the joint CDF for any x, y:

$$F_{X,Y}(x, y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(u, v)\, dv\, du = \int_{-\infty}^{x}\int_{-\infty}^{y} g(u)h(v)\, dv\, du.$$

$$= \left(\int_{-\infty}^{x} g(u)\, du\right)\left(\int_{-\infty}^{y} h(v)\, dv\right)$$

Proof: Product Density ⇒ Independence (2/2)

The marginals from the joint density are:

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\,dy = \int_{-\infty}^{+\infty} g(x)\,h(y)\,dy = g(x)\underbrace{\int_{-\infty}^{+\infty} h(y)\,dy}_{=1},$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\,dx = \int_{-\infty}^{+\infty} g(x)\,h(y)\,dx = h(y)\underbrace{\int_{-\infty}^{+\infty} g(x)\,dx}_{=1}.$$

Let

$$F_X(x) = \int_{-\infty}^{x} f_X(u)\,du = \int_{-\infty}^{x} g(u)\,du, \qquad F_Y(y) = \int_{-\infty}^{y} f_Y(v)\,dv = \int_{-\infty}^{y} h(v)\,dv.$$

Then,

$$F_{X,Y}(x,y) = F_X(x)\,F_Y(y) \quad \forall x, y,$$

which is exactly the definition of independence. □

Example: A Joint Density with Restricted Support

Example

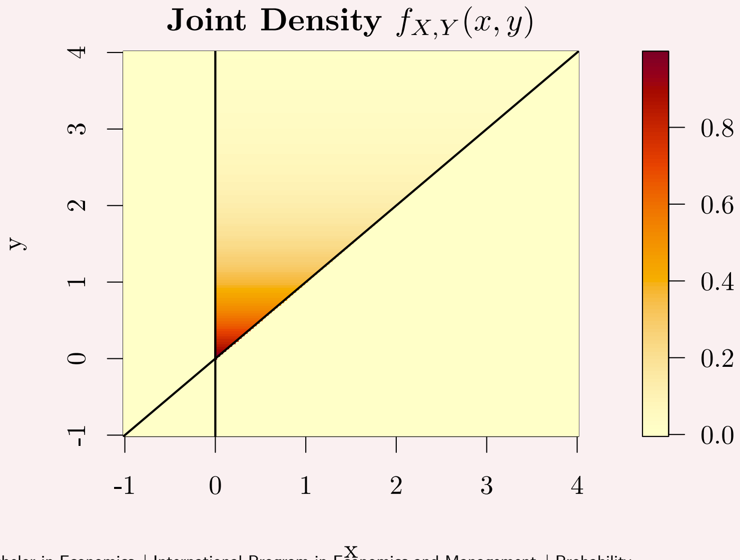
Consider the function

$$f_{X,Y}(x,y) = \begin{cases} e^{-y} & \text{if } 0 \leq x \leq y, \\ 0 & \text{otherwise.} \end{cases}$$

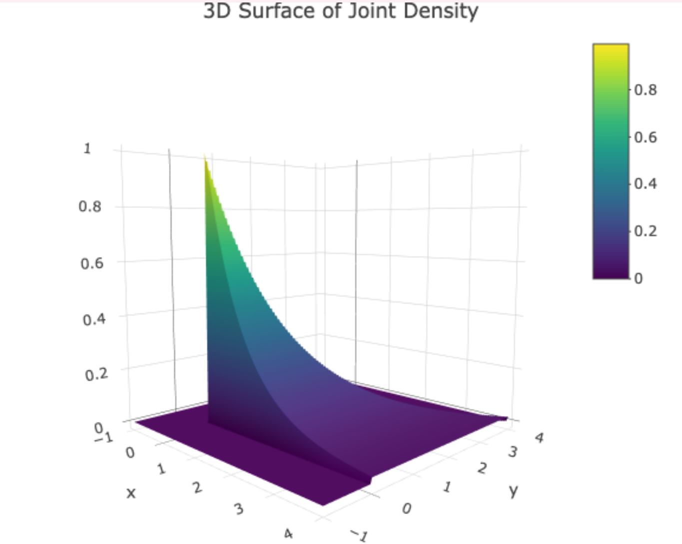
- *The support is the region* $\{(x,y) : 0 \leq x \leq y\}$ *(a wedge above the diagonal)*

1. *We can verify it is a density by checking* $\iint_{\mathbb{R}^2} f_{X,Y} = 1$.

Example: Visualization of the Joint Density



Example: Visualization of the Joint Density (3D)



Checking $\iint f_{X,Y} = 1$

$$\iint_{\mathbb{R}^2} f_{X,Y}(x,y)\,dx\,dy = \int_0^{+\infty} \int_0^y e^{-y}\,dx\,dy$$
$$= \int_0^{+\infty} e^{-y} \left(\int_0^y 1\,dx \right) dy$$
$$= \int_0^{+\infty} e^{-y} \left[x\right]_0^y dy = \int_0^{+\infty} ye^{-y}\,dy.$$

Using integration by parts:

$$u(y) = y, \quad u'(y) = 1, \quad v'(y) = e^{-y}, \quad v(y) = -e^{-y},$$

Hence,

$$\int_0^{+\infty} ye^{-y}\,dy = \Big[-ye^{-y}\Big]_0^{+\infty} + \int_0^{+\infty} e^{-y}\,dy = 0 + 1 = 1.$$

So $f_{X,Y}$ is a valid joint density.

4.3. Recovering the CDF from the Density

From Density to Joint CDF

Suppose (X, Y) admits a joint density $f_{X,Y}$.

By definition,

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

Since probabilities are computed by integration,

$$F_{X,Y}(x, y) = \iint_{(-\infty,x]\times(-\infty,y]} f_{X,Y}(u, v)\, du\, dv.$$

General Formula

Proposition 2. Recovering the joint CDF

If (X, Y) *has joint density* $f_{X,Y}$*, then for all* $(x, y) \in \mathbb{R}^2$*,*

$$F_{X,Y}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(u, v)\, dv\, du.$$

This is the 2D analogue of

$$F_X(x) = \int_{-\infty}^{x} f_X(u)\, du.$$

Example: A Joint Density with Restricted Support

Example

Consider again the function

$$f_{X,Y}(x,y) = \begin{cases} e^{-y} & \text{if } 0 \leq x \leq y, \\ 0 & \text{otherwise.} \end{cases}$$

1. *We verified that it is a density by checking* $\iint_{\mathbb{R}^2} f_{X,Y} = 1$.
2. *Now, we can compute the cumulative distribution function* $F_{X,Y}$.

Recall

$$f_{X,Y}(x,y) = \begin{cases} e^{-y} & \text{if } 0 \leq x \leq y, \\ 0 & \text{otherwise.} \end{cases}$$

We compute

$$F_{X,Y}(x,y) = \mathbb{P}(X \leq x,\ Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(u,v)\, du\, dv.$$

Since the support is $\{0 \leq u \leq v\}$:

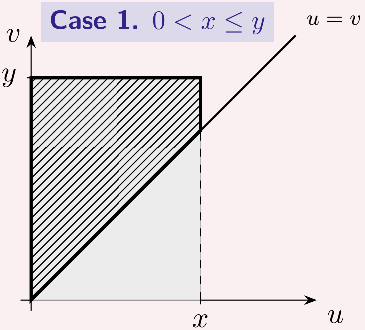
- If $x \leq 0$ or $y \leq 0$, the integration region does not intersect the support. Hence

$$F_{X,Y}(x,y) = 0.$$

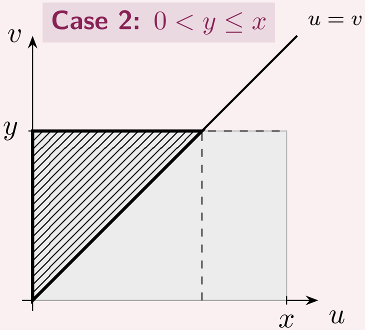
- If $x > 0$ and $y > 0$, we must distinguish whether $x \leq y$ or $y \leq x$.

- The CDF rectangle is $(-\infty, x] \times (-\infty, y]$,
- The density is non zero only on $\{0 \leq u \leq v\}$,
- So, the true integration region is:

$$(-\infty, x] \times (-\infty, y] \cap \{0 \leq u \leq v\}$$



Hatched: $0 \leq u \leq x,\ u \leq v \leq y$



Hatched: $0 \leq v \leq y,\ 0 \leq u \leq v$

Case 1: $0 < x \leq y$

Integration region:

$$0 \leq u \leq x, \qquad u \leq v \leq y.$$

$$\begin{aligned} F_{X,Y}(x,y) &= \int_0^x \int_u^y e^{-v}\, dv\, du \\ &= \int_0^x \left[-e^{-v}\right]_u^y\, du \\ &= \int_0^x \left(e^{-u} - e^{-y}\right) du \\ &= 1 - e^{-x} - xe^{-y}. \end{aligned}$$

Case 2: $0 < y \leq x$

Integration region:

$$0 \leq u \leq v, \qquad 0 \leq v \leq y.$$

$$\begin{aligned}
F_{X,Y}(x,y) &= \int_0^y \int_0^v e^{-v}\,du\,dv \\
&= \int_0^y \left[e^{-v}u\right]_0^v \,dv \\
&= \int_0^y ve^{-v}\,dv \\
&= \left[-ve^{-v}\right]_0^y + \int_0^y e^{-v}\,dv \\
&= (-ye^{-y} - 0) + \left[-e^{-v}\right]_0^y \\
&= -ye^{-y} + (-e^{-y} - (-1)) \\
&= 1 - (1+y)e^{-y}.
\end{aligned}$$

Hence,

$$F_{X,Y}(x,y) = \begin{cases} 0 & \text{if } x \leq 0 \text{ or } y \leq 0, \\ 1 - e^{-x} - xe^{-y} & \text{if } 0 < x \leq y, \\ 1 - (1+y)e^{-y} & \text{if } 0 < y \leq x. \end{cases}$$

- The formula changes because the geometry of the intersection region $(-\infty, x] \times (-\infty, y] \cap \{0 \leq u \leq v\}$ changes.
- We distinguished between two cases because of the diagonal $u = v$.

4.4. Marginal Distributions

Marginal Densities

Definition 15. Marginal Densities

Suppose (X, Y) has joint density $f_{X,Y}$.
*The **marginal densities** are obtained by integrating out the other variable:*

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x, y)\, dy, \qquad f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x, y)\, dx.$$

(Provided the integrals exist.)

Marginal CDFs

Definition 16. Marginal Cumulative Distribution Functions

*The **marginal CDFs** are*

$$F_X(x) = \mathbb{P}(X \leq x), \qquad F_Y(y) = \mathbb{P}(Y \leq y).$$

They can be recovered from the joint CDF by

$$F_X(x) = \lim_{y \to +\infty} F_{X,Y}(x, y), \qquad F_Y(y) = \lim_{x \to +\infty} F_{X,Y}(x, y).$$

If marginal densities exist, then equivalently:

$$F_X(x) = \int_{-\infty}^{x} f_X(u)\, du, \qquad F_Y(y) = \int_{-\infty}^{y} f_Y(v)\, dv.$$

Example: Marginal Density f_X

We apply this to the example:

$$f_{X,Y}(x,y) = \begin{cases} e^{-y} & \text{if } 0 \leq x \leq y, \\ 0 & \text{otherwise.} \end{cases}$$

For fixed x, the density is nonzero when $0 \leq x \leq y$, i.e.

$$y \in [x, +\infty), \quad \text{and only if } x \geq 0.$$

Thus:

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\,dy = \begin{cases} \int_x^{+\infty} e^{-y}\,dy & \text{if } x \geq 0, \\ 0 & \text{if } x < 0. \end{cases}$$

Marginal Density f_X

We compute the integral:

$$\int_x^{+\infty} e^{-y}\, dy = \Big[-e^{-y} \Big]_x^{+\infty} = 0 - (-e^{-x}) = e^{-x}.$$

Hence,

$$f_X(x) = \begin{cases} e^{-x} & \text{if } x \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, $X \sim \text{Exp}(1)$.

(Recall that for $\lambda > 0$, the density function of an exponential distribution is: $\lambda e^{-\lambda x}$ for $x \geq 0$ and 0 otherwise.)

Example: Marginal Density f_Y

For fixed y, the density is nonzero when $0 \leq x \leq y$.

This requires $y \geq 0$, and then $x \in [0, y]$.

Thus:

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x, y)\, dx = \begin{cases} \int_0^y e^{-y}\, dx & \text{if } y \geq 0, \\ 0 & \text{if } y < 0. \end{cases}$$

Since e^{-y} does not depend on x,

$$\int_0^y e^{-y}\, dx = e^{-y} \int_0^y 1\, dx = ye^{-y}.$$

Hence,

$$f_Y(y) = \begin{cases} ye^{-y} & \text{if } y \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Remark

For the record, $Y \sim \Gamma(2,1)$. Indeed, The Gamma density with shape α and rate λ is

$$f(y) = \frac{\lambda^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\lambda y}, \qquad y \geq 0.$$

Recall that for $k \in \mathbb{N}$:

$$\Gamma(k) = (k-1)!.$$

Hence,

$$\Gamma(2) = 1! = 1,$$

So, for $\alpha = 2$ and $\lambda = 1$,

$$f(y) = ye^{-y}, \qquad y \geq 0.$$

Therefore,

$$f_Y(y) = \begin{cases} ye^{-y} & \text{if } y \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

4.5. Conditional Densities

Conditional Densities

Definition 17. Conditional Density

Suppose (X, Y) *has joint density* $f_{X,Y}$ *and* $f_X(x) > 0$.
The ***conditional density of*** Y ***given*** $X = x$ *is*

$$f_{Y|X}(y \mid x) = \frac{f_{X,Y}(x, y)}{f_X(x)}.$$

Similarly, if $f_Y(y) > 0$,

$$f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}.$$

This is the continuous analogue of

$$\mathbb{P}(Y = y \mid X = x) = \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(X = x)}.$$

Interpretation

$$f_{Y|X}(y \mid x) = \frac{f_{X,Y}(x,y)}{f_X(x)}.$$

- Fix $X = x$.
  - The joint density becomes a function of y (we restrict the function to a “slice” $x =$ constant).
  - But this “slice” does not integrate to 1:

    $$\int_{\mathbb{R}} f_{X,Y}(x,y)\, dy = \int_{R} f_X(x) \neq 1$$

  - Dividing by $f_X(x)$ normalizes it so that

    $$\int_{-\infty}^{+\infty} f_{Y|X}(y \mid x)\, dy = 1.$$

- So $f_{Y|X}(\cdot \mid x)$ is a genuine density in y.

Example: Recall the Setup

Recall the previous example, where the joint density is

$$f_{X,Y}(x,y) = \begin{cases} e^{-y} & \textit{if } 0 \leq x \leq y, \\ 0 & \textit{otherwise.} \end{cases}$$

The marginal density of X is

$$f_X(x) = \begin{cases} e^{-x} & \textit{if } x \geq 0, \\ 0 & \textit{otherwise.} \end{cases}$$

The marginal density of Y is

$$f_Y(y) = \begin{cases} ye^{-y} & \textit{if } y \geq 0, \\ 0 & \textit{otherwise.} \end{cases}$$

Density of $Y \mid X = x$

For $x \geq 0$,

$$f_{Y|X}(y \mid x) = \frac{f_{X,Y}(x,y)}{f_X(x)}.$$

Since the joint density is nonzero only when $y \geq x$,

$$f_{Y|X}(y \mid x) = \begin{cases} \dfrac{e^{-y}}{e^{-x}} & \text{if } y \geq x, \\ 0 & \text{otherwise.} \end{cases}$$

Which simplifies to

$$f_{Y|X}(y \mid x) = \begin{cases} e^{-(y-x)} & \text{if } y \geq x, \\ 0 & \text{otherwise.} \end{cases}$$

Conditional Density of $Y \mid X = x$

Let
$$Z = Y - x.$$

Then $z = y - x \geq 0$, and

$$f_Z(z) = \begin{cases} e^{-z} & \text{if } z \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Hence,
$$Z \sim \text{Exp}(1).$$

Therefore,
$$Y \mid (X = x) = x + Z, \qquad Z \sim \text{Exp}(1).$$

Thus,
$$Y \mid (X = x) \sim x + \text{Exp}(1).$$

Conditional Density of $X \mid Y = y$

For $y \geq 0$,

$$f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}.$$

Since $0 \leq x \leq y$,

$$f_{X|Y}(x \mid y) = \begin{cases} \dfrac{e^{-y}}{ye^{-y}} & \text{if } 0 \leq x \leq y, \\ 0 & \text{otherwise.} \end{cases}$$

Which simplifies to

$$f_{X|Y}(x \mid y) = \begin{cases} \dfrac{1}{y} & \text{if } 0 \leq x \leq y, \\ 0 & \text{otherwise.} \end{cases}$$

Thus,

$$X \mid (Y = y) \sim \text{Uniform}(0, y).$$

Exercise

Let $c \in \mathbb{R}$ and consider the function

$$f_{X,Y}(x,y) = \begin{cases} c\, y e^{-y} & \text{if } 0 \leq x \leq y, \\ 0 & \text{otherwise.} \end{cases}$$

1. *Find the value of c for which $f_{X,Y}$ is a density.*
2. *Let (X,Y) be a pair of random variables with joint density $f_{X,Y}$. Determine the marginal densities of X and Y.*
3. *Determine the conditional densities of $X \mid Y = y$ and $Y \mid X = x$.*

Solution: Finding c

We need $\iint_{\mathbb{R}^2} f_{X,Y}(x,y)\,dx\,dy = 1$. Since the support is $\{(x,y) : 0 \leq x \leq y\}$,

$$\begin{aligned}
1 &= \int_{y=0}^{\infty}\int_{x=0}^{y} c\,ye^{-y}\,dx\,dy \\
&= c\int_{y=0}^{\infty} ye^{-y}\left(\int_{x=0}^{y} 1\,dx\right)dy \\
&= c\int_{0}^{\infty} ye^{-y}\,[x]_0^y\,dx = c\int_0^\infty y^2e^{-y}\,dy \\
&= c\left(\left[-y^2e^{-y}\right]_0^\infty + \int_0^\infty 2ye^{-y}\,dy\right) \quad \text{(IBP)} \\
&= c\left(\left[-2ye^{-y}\right]_0^\infty + \int_0^\infty 2e^{-y}\,dy\right) \quad \text{(IBP)} \\
&= c(\left[-2e^{-y}\right]_0^\infty) = 2c.
\end{aligned}$$

Hence $c = 1/2$.

Solution: Marginal densities (1/2)

With $c = 1/2$,

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{2}\, ye^{-y} & \text{if } 0 \leq x \leq y, \\ 0 & \text{otherwise.} \end{cases}$$

Marginal of Y. For $y \geq 0$,

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx = \int_0^y \frac{1}{2}\, ye^{-y}\,dx = \frac{1}{2}\, ye^{-y} \cdot y = \frac{1}{2}\, y^2 e^{-y},$$

and $f_Y(y) = 0$ for $y < 0$.

Hence,

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

Note: you can check that $Y \sim \Gamma(3,1)$.

Solution: Marginal densities (2/2)

Marginal of X. For $x \geq 0$ (and recall from the support that $x \leq y$, hence $y \geq x$),

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy \\ &= \tfrac{1}{2}\int_{x}^{\infty} ye^{-y}\,dy \\ &= \tfrac{1}{2}\left(\left[-ye^{-y}\right]_x^{+\infty} + \int_x^{\infty} e^{-y}\,dy\right) \quad \text{(IBP)} \\ &= \tfrac{1}{2}\left(xe^{-x} + \left[-e^{-y}\right]_x^{\infty}\right) \\ &= \tfrac{1}{2}\left(xe^{-x} + e^{-x}\right) \\ &= \tfrac{1}{2}(x+1)e^{-x} \end{aligned}$$

hence

$$f_X(x) = \begin{cases} \tfrac{1}{2}\,(x+1)e^{-x} & \text{if} \quad x \geq 0 \\ ,0 & \text{if} \quad x < 0. \end{cases}$$

Solution: Conditional Densities

Conditional density of $X \mid Y = y$. For $y > 0$, since $f_Y(y) = \frac{1}{2}y^2e^{-y}$,

$$f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \begin{cases} \dfrac{\frac{1}{2}\,ye^{-y}}{\frac{1}{2}\,y^2e^{-y}} = \dfrac{1}{y} & \text{if } 0 \leq x \leq y, \\ 0 & \text{otherwise.} \end{cases}$$

Thus, $X \mid Y = y \sim \text{Uniform}(0, y)$.

Conditional density of $Y \mid X = x$. For $x \geq 0$, since $f_X(x) = \frac{1}{2}(x+1)e^{-x}$,

$$f_{Y|X}(y \mid x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \begin{cases} \dfrac{\frac{1}{2}\,ye^{-y}}{\frac{1}{2}\,(x+1)e^{-x}} = \dfrac{y}{x+1}\,e^{-(y-x)} & \text{if } y \geq x, \\ 0 & \text{otherwise.} \end{cases}$$

Note that if $x = 0$, then $f_{Y|X}(y \mid 0) = ye^{-y}, \qquad y \geq 0$ (density of a $\Gamma(2,1)$ distribution). Hence, $Y \mid X = 0 \sim \Gamma(2,1)$.

4.6. Distribution of a Sum $S = X + Y$

Distribution of a Sum

Definition 18. Distribution of $S = X + Y$

Let (X, Y) be a pair of continuous random variables and define

$$S = X + Y.$$

The distribution of S is given by its CDF:

$$F_S(s) = \mathbb{P}(X + Y \leq s).$$

Suppose (X, Y) has joint density $f_{X,Y}$. Then

$$\begin{aligned} F_S(s) &= \mathbb{P}(X + Y \leq s) \\ &= \iint_{\{x+y\leq s\}} f_{X,Y}(x, y)\, dx\, dy \\ &= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{s-x} f_{X,Y}(x, y)\, dy \right) dx. \end{aligned}$$

Density of a Sum

Proposition 3. Density of $S = X + Y$

Let (X, Y) be a pair of continuous random variables with joint density $f_{X,Y}$. Define

$$S = X + Y.$$

Then S admits a density given by

$$f_S(s) = \int_{-\infty}^{+\infty} f_{X,Y}(x,\, s - x)\, dx.$$

Proof (1/5): Sketch

Let $S = X + Y$.

The objective here is to prove that S admits a density of the form

$$f_S(s) = \int_{-\infty}^{+\infty} f_{X,Y}(x,\, s - x)\, dx.$$

To obtain a density, we proceed as follows:

1. Express $F_S(s) = \mathbb{P}(S \leq s)$ as a double integral,
2. Perform a 2D change of variables,
3. Rewrite $F_S(s)$ in the form

$$F_S(s) = \int_{-\infty}^{s} (\cdots)\, dz,$$

4. Identify the density.

Proof (2/5): CDF as a Double Integral

For $s \in \mathbb{R}$,

$$F_S(s) = \mathbb{P}(X + Y \leq s).$$

Since (X, Y) admits joint density $f_{X,Y}$,

$$F_S(s) = \iint_{\{(x,y):\, x+y \leq s\}} f_{X,Y}(x, y)\, dx\, dy$$

We now perform a change of variables that isolates $x + y$.

Reminder

Change of Variables in $\mathbb{R}^2$

Let $(u, v) = \varphi(x, y)$ be a C^1 bijective transformation with inverse $(x, y) = \varphi^{-1}(u, v)$. Then

$$\iint_A f(x, y)\, dx\, dy = \iint_{\varphi(A)} f\left(\varphi^{-1}(u, v)\right) \left|\det\left(\frac{\partial(x, y)}{\partial(u, v)}\right)\right| du\, dv.$$

The interpretation, in a nutshell:

- A small rectangle $du\, dv$ in the (u, v)-plane is transformed into a small parallelogram in the (x, y)-plane.
- The area of that parallelogram is

$$\left|\det\left(\frac{\partial(x, y)}{\partial(u, v)}\right)\right| du\, dv.$$

- The determinant measures the *local area distortion* caused by the transformation.

This translates, mathematically, to applying this correction factor

$$dx\, dy = \left|\det\left(\frac{\partial(x, y)}{\partial(u, v)}\right)\right| du\, dv.$$

Proof (3/5): Connecting of this Theorem to the Proof

In our proof of $S = X + Y$:

- Original variables: (x, y)
- New variables: $(u, v) = (x, z)$ with $z = x + y$
- Transformation: $\varphi(x, y) = (x,\ x + y)$
- Inverse transformation: $\varphi^{-1}(x, z) = (x,\ z - x)$
- Integration domain: $A = \{(x, y) \in \mathbb{R}^2 : x + y \leq s\}$
- Transformed domain: $\varphi(A) = \{(x, z) \in \mathbb{R}^2 : z \leq s\}$

Proof (4/5): 2D Change of Variables and Jacobian

The Jacobian of the inverse is:

$$\frac{\partial(x,y)}{\partial(x,z)} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

$$\det = 1 \cdot 1 - 0 \cdot (-1) = 1.$$

Hence the area element transforms as

$$dx\,dy = \left|\det\left(\frac{\partial(x,y)}{\partial(x,z)}\right)\right| dx\,dz = 1 \cdot dx\,dz.$$

The constraint $x + y \leq s$ becomes $z \leq s$.

Proof (5/5): Rewrite and Identify the Density

Under this transformation,

$$F_S(s) = \iint_{\{(x,z):\, z \leq s\}} f_{X,Y}(x,\, z-x) \underbrace{\left| \det\left(\frac{\partial(x,y)}{\partial(x,z)} \right) \right|}_{=1} dx\, dz$$

$$= \int_{-\infty}^{s} \left(\int_{-\infty}^{+\infty} f_{X,Y}(x,\, z-x)\, dx \right) dz.$$

Thus

$$F_S(s) = \int_{-\infty}^{s} f_S(z)\, dz,$$

with

$$f_S(z) = \int_{-\infty}^{+\infty} f_{X,Y}(x,\, z-x)\, dx.$$

Therefore,

$$f_S(s) = \int_{-\infty}^{+\infty} f_{X,Y}(x,\, s-x)\, dx.$$

Example: Sum of a Uniform and an Exponential

Let X and Y be independent random variables. Assume

$$X \sim \mathcal{U}(0,1), \qquad Y \sim \mathrm{Exp}(1).$$

Define

$$S = X + Y.$$

Determine the density of S.

The densities are

$$f_X(x) = \begin{cases} 1 & \text{if} \quad x \in [0,1], \\ 0 & \text{otherwise.} \end{cases} \qquad f_Y(y) = \begin{cases} e^{-y} & \text{if} \quad x \in [0,+\infty), \\ 0 & \text{otherwise.} \end{cases}$$

Since $X \in [0,1]$ and $Y \geq 0$, we have $S = X + Y \in [0, +\infty)$. Hence $f_S(s) = 0$ for $s < 0$.

By independence, for $s \in \mathbb{R}$,

$$f_S(s) = \int_{-\infty}^{+\infty} f_X(x)\, f_Y(s-x)\, dx.$$

Here $f_X(x) = 0$ if $x \notin [0,1]$, and $f_Y(s-x) = 0$ if $s - x < 0$, i.e. if $x > s$.

Therefore the integrand is nonzero only when

$$x \in [0,1] \cap (-\infty, s].$$

So

$$f_S(s) = \int_0^{\min(1,s)} e^{-(s-x)}\, dx, \qquad (s \geq 0).$$

We thus need ton consider two cases:

1. If $0 \leq s < 1$, then $\min(1, s) = s$, so

$$f_S(s) = \int_0^s e^{-(s-x)}\, dx = e^{-s} \int_0^s e^x\, dx = e^{-s}(e^s - 1) = 1 - e^{-s}.$$

2. If $s \geq 1$, then $\min(1, s) = 1$, so

$$f_S(s) = \int_0^1 e^{-(s-x)}\, dx = e^{-s} \int_0^1 e^x\, dx = e^{-s}(e - 1).$$

Hence,

$$f_S(s) = \begin{cases} 0, & s < 0, \\ 1 - e^{-s}, & 0 \leq s < 1, \\ (e-1)e^{-s}, & s \geq 1. \end{cases}$$

5. Exercise

Exercise: Marginals and Conditionals

Let (X, Y) be a pair of continuous random variables with joint density

$$f_{X,Y}(x, y) = \begin{cases} 4e^{-2x}, & x \geq y \geq 0, \\ 0, & \textit{otherwise.} \end{cases}$$

1. *Compute the marginal densities of X and Y.*
2. *Compute the conditional density of X given $Y = y$.*
3. *Compute the conditional density of Y given $X = x$.*

Support of (X, Y)

The density is nonzero on

$$\mathcal{S} = \{(x, y) \in \mathbb{R}^2 : \ 0 \leq y \leq x\}.$$

Equivalently:

$$x \in [0, +\infty), \qquad y \in [0, x].$$

This support will determine the bounds of integration for the marginals and conditionals.

Marginal Density of X

For $x \geq 0$,

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\, dy = \int_0^x 4e^{-2x}\, dy.$$

Since $4e^{-2x}$ does not depend on y,

$$f_X(x) = 4e^{-2x} \int_0^x 1\, dy = 4xe^{-2x}, \qquad x \geq 0.$$

Therefore,

$$f_X(x) = \begin{cases} 4xe^{-2x} & \text{if} \quad x \geq 0, \\ 0 & \text{if} \quad x < 0. \end{cases}$$

Marginal Density of Y

Recall that the support of (X, Y) is $\mathcal{S} = \{(x, y) \in \mathbb{R}^2 : 0 \leq y \leq x\}$. Fix $y \in \mathbb{R}$.

- If $y < 0$, then there is no x such that $(x, y) \in \mathcal{S}$. Hence $f_Y(y) = 0$.
- If $y \geq 0$, then $(x, y) \in \mathcal{S}$ iff $x \geq y$. (No additional condition is needed since $x \geq y \geq 0$ implies $x \geq 0$.)

Therefore, for $y \geq 0$,

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{+\infty} f_{X,Y}(x, y)\, dx = \int_y^{+\infty} 4e^{-2x}\, dx \\ &= 4\left[-\frac{1}{2}e^{-2x}\right]_y^{+\infty} = 2e^{-2y}. \end{aligned}$$

Hence,

$$f_Y(y) = \begin{cases} 2e^{-2y}, & y \geq 0, \\ 0, & y < 0. \end{cases}$$

Conditional Density of X given $Y = y$

For $y \geq 0$ and $x \geq y$,

$$f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{4e^{-2x}}{2e^{-2y}} = 2e^{-2(x-y)}.$$

Thus,

$$f_{X|Y}(x \mid y) = \begin{cases} 2e^{-2(x-y)} & \text{if} \quad x \geq y, \\ 0 & \text{if} \quad x < y, \end{cases} \qquad (y \geq 0).$$

Hence, $X - y \mid (Y = y) \sim \mathrm{Exp}(2)$.

Conditional Density of Y given $X = x$

For $x > 0$ and $0 \leq y \leq x$,

$$f_{Y|X}(y \mid x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{4e^{-2x}}{4xe^{-2x}} = \frac{1}{x}.$$

Hence,

$$f_{Y|X}(y \mid x) = \begin{cases} \frac{1}{x} & \text{if} \quad 0 \leq y \leq x, \\ 0 & \text{otherwise,} \end{cases} \qquad (x > 0).$$

Hence, $Y \mid (X = x) \sim \mathcal{U}(0, x)$ (and if $x = 0$, then $Y = 0$ a.s.).

References I

Hansen, B. (2022). *Probability and statistics for economists*. Princeton University Press.