Lecture 1. Bivariate Dependent Risks

Enkelejd Hashorva & Pavel Ievlev
Université de Lausanne

15 September, 2025

Leargning objectives

- Introduction to bivariate df's, pdf's, survival functions
- Introduction of copulas
- Calculation of moments for vectors of bivariate risks

Bivatiate random vectors and their joint df

Let X and Y be two random variables (rvs) with distribution functions (df's) F_1 and F_2, respectively.

Definition 1 (Bivariate random vector and its joint df).

- The pair (X, Y) is referred to as a **bivariate random vector (RV)**.
- The **joint df**[1] F of (X, Y) is a function $F : \mathbb{R}^2 \to [0, 1]$ defined by

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

- We denote this by $(X, Y) \sim F$.
- Functions $F_1(x) = \mathbb{P}\{X \leq x\}$ and $F_2(y) = \mathbb{P}\{Y \leq y\}$ are called the **marginal dfs**.

[1] **df** = distribution function

Survival function

Let $A = \{X \leq x\}$ and $B = \{Y \leq y\}$. Then

$$\begin{aligned} \mathbb{P}\{X \leq x \text{ or } Y \leq y\} &= \mathbb{P}\{A \cup B\} = \mathbb{P}\{A\} + \mathbb{P}\{B\} - \mathbb{P}\{A \cap B\} \\ &= F_1(x) + F_2(y) - F(x, y). \end{aligned}$$

Therefore,

$$\mathbb{P}\{X > x,\ Y > y\} = 1 - F_1(x) - F_2(y) + F(x, y).$$

Definition 2 (Survival function).
The function $\bar{F} : \mathbb{R}^2 \to [0, 1]$ defined by

$$\bar{F}(x, y) = \mathbb{P}\{X > x, Y > y\}$$

is called the **joint survival function** of (X, Y).

Alternative forms of the same formula

By replacing $F_i \rightsquigarrow 1 - \overline{F}_i$, we can easily derive from

$$\overline{F}(x, y) = 1 - F_1(x) - F_2(y) + F(x, y)$$

another equivalent formula

$$\overline{F}(x, y) = \overline{F}_1(x) + \overline{F}_2(y) - 1 + F(x, y).$$

Both forms are occasionally useful.

Product, upper and lower df's

If $X \sim F_1$ is independent of $Y \sim F_2$, then the joint df of (X, Y) is given by

$$F(x, y) = F_1(x)F_2(y).$$

Define further the **upper df**

$$H(x, y) = \min(F_1(x), F_2(y))$$

and the **lower df**

$$G(x, y) = (F_1(x) + F_2(y) - 1)_+,$$

where $a_+ = \max(0, a)$.

Not all functions are dfs

Question: A **df** is *by definition* a function such that $F(x, y) = \mathbb{P}\{X \le x,\ Y \le y\}$ for *some* random vector (X, Y). But if we are *given* a function F, how can we check whether it is a df of *some* random vector (X, Y)?

Clearly, F must satisfy the following properties:

- F is increasing (non-decreasing) in each argument
- F is right-continuous
- $F(x, -\infty) = \mathbb{P}\{X \le x,\ Y \le -\infty\} = 0$
- $F(-\infty, y) = \mathbb{P}\{X \le -\infty,\ Y \le y\} = 0$

Are these conditions **sufficient**? **Answer:** No! There is one more condition: probabilities of *all rectangles* must be non-negative, i.e.

$$\boxed{\mathbb{P}\{a < X \le b,\ c < Y \le d\} = F(b, d) - F(a, d) - F(b, c) + F(a, c) \ge 0}$$

Exercise

Exercise. Check that the functions G and H defined earlier are indeed dfs.

$$G(x, y) = (F_1(x) + F_2(y) - 1)_+, \quad H(x, y) = \min(F_1(x), F_2(y)).$$

Quick checks if something is not a df

- If F assumes negative values, then it is not a df.
- If F assumes values greater than 1, then it is not a df.
- If $F(x, y)$ is not increasing in x or y, then it is not a df.

Joint df $\to$ marginal dfs

If $(X, Y) \sim F$, then the marginal dfs F_1 and F_2 can be obtained from F by

$$F_1(x) = F(x, \infty), \quad F_2(y) = F(\infty, y).$$

Indeed, as $y \to \infty$, the event $\{Y \leq y\}$ becomes certain.

- **Question:** Can we determine the joint df F from the marginal dfs F_1 and F_2?
- **Answer:** In general, no!
- There are infinitely many joint dfs with the same marginals, each corresponding to a different **dependence structure** between X and Y.
- These dependence structures are encoded by **copulas** and will be the main subject of this course.

Exercise

Exercise. Given marginal dfs F_1 and F_2, show that the function

$$F(x, y) = F_1(x)\, F_2(y) \left(1 + a\overline{F}_1(x)\overline{F}_2(y)\right), \quad a \in [-1, 1]$$

is a joint df with marginals F_1 and F_2.

This family of dfs is known as the **FGM distribution/family of distributions**.

Joint df → joint pdf/pmf

Definition 3 (Joint pdf).

We say that a function $f \geq 0$ is the **joint pdf**[2] of a given df F if F can be represented as

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(z, w)\, dz dw.$$

Definition 4 (Joint pmf).

We say that a function $f \geq 0$ is the **joint pmf**[3] of a given df F if F can be represented as

$$F(x, y) = \sum_{z \leq x, w \leq y} f(z, w).$$

[2] **pdf** = probability density function
[3] **pmf** = probability mass function

Examples of a joint pdf/pmf

Let $f(x,y) = 1$ for all $x, y \in [0,1]^2$. Then the corresponding df is

$$F(x,y) = \int_0^x \int_0^y f(s,t)\,dsdt = \int_0^x \int_0^y ds\,dt = xy = F_1(x)F_2(y),$$

where

$$F_1(x) = x \quad \text{and} \quad F_2(y) = y \quad \text{for} \quad x, y \in [0,1].$$

Thus, F is the joint df of two independent $\text{Unif}(0,1)$ random variables.

Similarly, let $f(x,y) = \frac{1}{4}$ for all $x, y \in \{0,1\}^2$. Then the corresponding df is

$$F(x,y) = F_1(x)\,F_2(y),$$

where F_1 and F_2 are the dfs of two independent $\text{Ber}(1/2)$ random variables.

pdf (pmf) $\rightarrow$ joint df

If $f \geq 0$ is an integrable (summable) function such that

$$\int_{\mathbb{R}^2} f(x,y)\,dxdy = 1, \quad \left(\sum_{i=1,j=1}^{\infty} f(x_i,x_j) = 1\right),$$

then the function F **defined** by

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(z,w)\,dzdw, \quad \left(F(x,y) = \sum_{z \leq x, w \leq y} f(z,w)\right)$$

is a df.

Differentiable joint df → joint pdf/pmf

Given an joint df F, such that the following mixed derivative

$$f(x, y) = \frac{\partial^2 F}{\partial x \, \partial y}(x, y)$$

exists for almost all $x, y \in \mathbb{R}$, can we conclude that f is the joint pdf (pmf) of F?

- If F is a df and the mixed partial derivative f exists, then $f \geq 0$ automatically (follows from the rectangle property), so we don't need to check positivity.
- However, even if $f \geq 0$, it may not be a pdf due to a **loss of mass**[4]: it is possible to have $\int_{\mathbb{R}^2} f(x, y) \, dx \, dy < 1$.
- If f is continuous, then loss of mass cannot happen, so f is indeed the joint pdf of F.
- If f is discontinuous, we have to check that $\int_{\mathbb{R}^2} f(x, y) \, dx \, dy = 1$.

[4] Where does mass go?

Conditional pdf/pmf and marginal pdf/pmf

If (X, Y) with joint df F has a joint pdf (pmf) f, then the conditional pdf (pmf) of $X \mid Y = y$ satisfies

$$f(x, y) = f_2(y)\, f_{1|2}(x \mid y), \quad x, y \in \mathbb{R},$$

where f_2 is the marginal pdf (pmf) of Y.

Moreover, from f we can calculate the marginal pdf's (pmf's) f_i's, namely

$$f_1(x) = \int_{\mathbb{R}} f(x, z) dz, \quad f_2(y) = \int_{\mathbb{R}} f(z, y) dz, \quad x, y \in \mathbb{R}.$$

These are pdfs of F_1 and F_2.

Expectations

Let $(X, Y) \sim F$ and $J : \mathbb{R}^2 \to \mathbb{R}$ be some function. Then, the expectation of $J(X, Y)$ is defined by

$$\mathbb{E}\{J(X,Y)\} = \int_{\mathbb{R}^2} J(x,y)\ dF(x,y).$$

If F has a joint pdf f, then

$$\mathbb{E}\{J(X,Y)\} = \int_{\mathbb{R}^2} J(x,y)\, f(x,y)\, dxdy.$$

If F has a joint pmf f, then

$$\mathbb{E}\{J(X,Y)\} = \sum_{i,j} J(x_i,x_j)\, f(x_i,x_j).$$

Smirnov transform/Inverse sampling method

Theorem 5.

- *Let X be any random variable with continuous df F.*
- *Apply the df F to X: $Y = F(X)$.*
- ***Claim:*** $Y \sim \text{Unif}(0,1)$.

Proof in the case when F is invertible:

$$\mathbb{P}\{Y \leq y\} = \mathbb{P}\{F(x) \leq y\} = \mathbb{P}\left\{X \leq F^{-1}(y)\right\} = F(F^{-1}(y)) = y,$$

which is the df of $\text{Unif}(0,1)$.

Copulas

Definition 6 (Copula of a continuous bivariate df).

- Let $(X, Y) \sim F$ be a bivariate random vector with continuous marginal dfs F_1 and F_2.
- Then the **copula** C of F is the joint df of the random vector $(F_1(X), F_2(Y))$.

Note that the marginals of $(F_1(X), F_2(Y))$ are $\text{Unif}(0, 1)$ by the Smirnov transform theorem. This motivates the following alternative definition of a copula:

Definition 7 (Copula).

A **copula** is a bivariate df C whose marginals are $\text{Unif}(0, 1)$.

The two definitions are equivalent in the sense that every copula is the copula of some bivariate df and vice versa.

Idea of copulas

- A copula C is a compact way to encode the **dependence structure** between X and Y.
- By passing from (X, Y) to $(F_1(X), F_2(Y))$, we **throw away** the information about the marginal distributions of X and Y. The resulting vector does not know anything about the laws of X and Y, but knows everything about how X and Y depend on each other.
- If we know the copula C of F and the marginal dfs F_1 and F_2, then we can reconstruct the joint df F. This is done by **Sklar's theorem**, which will be discussed later.

Product, upper and lower copulas

Here are the most basic examples of copulas:

- The **product copula** $C_I(u_1, u_2) = u_1 u_2$ encodes *independence*. This is the copula of two independent $\mathrm{Unif}(0, 1)$ random variables.
- The **upper copula** $C_U(u_1, u_2) = \min(u_1, u_2)$ encodes *perfect positive dependence*. This is the copula of (U, U) where $U \sim \mathrm{Unif}(0, 1)$.
- The **lower copula** $C_L(u_1, u_2) = (u_1 + u_2 - 1)_+$ encodes *perfect negative dependence*. This is the copula of $(U, 1 - U)$ where $U \sim \mathrm{Unif}(0, 1)$.

Not every function is a copula

Question: How can we check whether a given function C is a copula?

Clearly, C must satisfy the properties of a bivariate df. In addition, the marginals of C must be $\text{Unif}(0,1)$, i.e.

$$C(u,1) = u, \quad C(1,u) = u, \quad u \in [0,1].$$

Are these conditions sufficient?

Answer: Yes! These conditions are necessary and sufficient for C to be a copula.

Exercise

Exercise. Check that C_L is the copula of $(U, 1-U)$ where $U \sim \mathrm{Unif}(0,1)$ by directly calculating its joint df.

Upper and lower copulas do not have pdfs

Note that the functions

$$C_U(u_1, u_2) = \min(u_1, u_2) \quad \text{and} \quad C_L(u_1, u_2) = (u_1 + u_2 - 1)_+$$

are **not** differentiable on the diagonal $u_1 = u_2$, which is a set of Lebesgue measure zero. However, their mixed partial derivatives *away from the diagonal* are equal to zero[5]. Therefore,

$$\int_{[0,1]^2 \setminus \{u_1 = u_2\}} \frac{\partial^2 C_U}{\partial u_1 \, \partial u_2}(u_1, u_2) \, du_1 du_2 = 0,$$

and similarly for C_L. Therefore, C_U and C_L do not have joint pdfs.

[5] Check this!