Lecture 2. Tractable dependence models

Enkelejd Hashorva & Pavel Ievlev
Université de Lausanne

19 September, 2025

Leargning objectives

- Understand uniform bivariate distributions on different domains;
- Introduce the concept of **conditional independence**;
- Introduce important classes of dependent risks cush as **commonotonic**, **countermonotonic**, **implicit** and **explicit** risks;

Uniform distribution on a rectangle

Consider the rectangle

$$D = [a, b] \times [c, d] \subset \mathbb{R}^2.$$

Consider the following pdf[1]:

$$f_D(x, y) = \frac{1}{(b-a)(d-c)} \mathbb{1}\{(x, y) \in D\}.$$

Let F be the corresponding df.

> **Definition 1.**
> A random vector (X, Y) with df F is said to be **uniformly distributed** on the rectangle D. We denote this by $(X, Y) \sim \mathrm{Unif}(D)$.

[1] Why is this a valid pdf?

Independent risks

Let F be some df and $(X, Y) \sim F$.

Theorem 2.
*X and Y are independent **if and only if** F may be written as product of any*[2] *two functions G and H, i.e.*

$$F(x, y) = G(x)H(y).$$

Theorem 3.
*Assume F has pdf/pmf f. Then (X, Y) is independent **if and only if** f may be written as product of any*[3] *two functions g and h, i.e.*

$$f(x, y) = g(x)h(y).$$

[2] Check that $G(x) = cF_1(x)$ and $H(y) = F_2(y)/c$ for some $c > 0$. In other words, G and H are the marginal dfs *up to a constant*.

[3] Check that $g(x) = cf_1(x)$ and $h(y) = f_2(y)/c$ for some $c > 0$.

Example: independence of uniform risks on a rectangle

Recall that the pdf of $\mathrm{Unif}(D)$ with $D = [a, b] \times [c, d]$ is

$$f_D(x, y) = \frac{1}{(b-a)(d-c)} \mathbb{1}\{(x, y) \in D\}.$$

Note that

$$f_D(x, y) = \underbrace{\frac{1}{b-a}\mathbb{1}\{x \in [a, b]\}}_{g(x)} \cdot \underbrace{\frac{1}{d-c}\mathbb{1}\{y \in [c, d]\}}_{h(y)}.$$

Hence, X and Y are independent.

Uniform distribution on a set D

Let $D \subset \mathbb{R}^2$ be a set with **positive and finite** area

$$|D| = \iint_D dx\,dy \in (0, \infty).$$

Define the following pdf[4]:

$$f_D(x, y) = \frac{1}{|D|} \mathbb{1}\{(x, y) \in D\}.$$

Let F be the corresponding df.

> **Definition 4.**
> A random vector (X, Y) with df F is said to be **uniformly distributed** on the set D. We denote this by $(X, Y) \sim \text{Unif}(D)$.

[4] Why is this a valid pdf?

Uniform distribution on D is typically <u>not</u> independent

Theorem 5.
Let $(X, Y) \sim \mathrm{Unif}(D)$. *If* X *and* Y *are independent, then* D *is of the form*

$$D = A \times B \quad \textit{for some} \quad A, B \subset \mathbb{R}$$

up to a set of zero area.

Proof. Assume that X and Y are independent:

$$f_D(x, y) = g(x)h(y) = \frac{1}{|D|} \mathbb{1}\{(x, y) \in D\}.$$

Take any point $(x_0, y_0) \in D$ with $f_D(x_0, y_0) \neq 0$.

Uniform distribution on D is typically not independent

Theorem 5.
Let $(X, Y) \sim \mathrm{Unif}(D)$*. If* X *and* Y *are independent, then* D *is of the form*

$$D = A \times B \quad \textit{for some} \quad A, B \subset \mathbb{R}$$

up to a set of zero area.

Proof. We have: $f_D(x, y) = g(x)\, h(y)$, $g(x_0), h(y_0) \neq 0$.

Uniform distribution on D is typically <u>not</u> independent

Theorem 5.
Let $(X, Y) \sim \mathrm{Unif}(D)$. *If* X *and* Y *are independent, then* D *is of the form*

$$D = A \times B \quad \textit{for some} \quad A, B \subset \mathbb{R}$$

up to a set of zero area.

Proof. We have: $f_D(x, y) = g(x)\, h(y)$, $g(x_0), h(y_0) \neq 0$. Then,

$$f_D(x_0, y) = g(x_0)\, h(y) = \frac{1}{|D|} \mathbb{1}\{y : (x_0, y) \in D\} \implies h(y) = c\, \mathbb{1}\{y : (x_0, y) \in D\}$$

$$f_D(x, y_0) = g(x)\, h(y_0) = \frac{1}{|D|} \mathbb{1}\{x : (x, y_0) \in D\} \implies g(x) = c'\, \mathbb{1}\{x : (x, y_0) \in D\}$$

Uniform distribution on D is typically <u>not</u> independent

Theorem 5.
Let $(X,Y) \sim \mathrm{Unif}(D)$. *If* X *and* Y *are independent, then* D *is of the form*

$$D = A \times B \quad \textit{for some} \quad A, B \subset \mathbb{R}$$

up to a set of zero area.

Proof. We have: $f_D(x,y) = g(x)\,h(y)$, $g(x_0), h(y_0) \neq 0$. Then,

$$h(y) = c\,\mathbb{1}\{y : (x_0, y) \in D\} = c\,\mathbb{1}\{y \in B\} \quad \text{with} \quad B = \{y : (x_0, y) \in D\}$$

$$g(x) = c'\,\mathbb{1}\{x : (x, y_0) \in D\} = c'\,\mathbb{1}\{x \in A\} \quad \text{with} \quad A = \{x : (x, y_0) \in D\}$$

Uniform distribution on D is typically <u>not</u> independent

Theorem 5.
Let $(X, Y) \sim \mathrm{Unif}(D)$. If X and Y are independent, then D is of the form

$$D = A \times B \quad \textit{for some} \quad A, B \subset \mathbb{R}$$

up to a set of zero area.

Proof. We have: $f_D(x, y) = g(x)\, h(y)$, $g(x_0), h(y_0) \neq 0$. Then,

$$f_D(x, y) = cc'\, \mathbb{1}\{x \in A\}\, \mathbb{1}\{y \in B\} \implies D = A \times B$$

up to a set of zero area. □

Conditional independence

Definition 6 (Conditional independence).
$X_1, \dots, X_d$ are **conditionally independent** given $W = w$ if for all real numbers $x_1, \dots, x_d$ holds

$$\mathbb{P}\left\{X_1 \le x_1, \dots, X_d \le x_d \mid W = w\right\} = \prod_{i=1}^{d} \mathbb{P}\left\{X_i \le x_i \mid W = w\right\}.$$

If W is independent of $X_1, \dots, X_d$, then conditional independence implies independence.

“Common deflator/inflator” dependence structure

Let Θ, Z_1, $Z_2, \ldots, Z_d$ be **independent** exponentially distributed random variables. Define

$$X_1 = \frac{Z_1}{\Theta}, \quad X_2 = \frac{Z_2}{\Theta}, \quad \ldots, \quad , X_d = \frac{Z_d}{\Theta}.$$

Then, **conditionally on** $\Theta = \theta$, $X_1, \ldots, X_d$ are **independent** exponentially distributed random variables.

Comonotonic risks

Definition 7 (Comonotonic risks).

Random variables X and Y are said to be **comonotonic** if there exists **one** random variable $U \sim \text{Unif}(1)$ (common risk) such that $X = F_1^{-1}(U)$ and $Y = F_2^{-1}(U)$. We say that (X, Y) is a **comonotonic vector**.

Example. Let $X = c_1 U$ and $Y = c_2 U$, where $c_1, c_2 > 0$. Then (X, Y) is a comonotonic vector with $\text{Unif}(0, c_i)$ marginals.

Theorem 8 (Comonotonicity conditions).

The following conditions are equivalent:

- (X, Y) *is comonotonic*
- $F(x, y) = \min(F_1(x), F_2(y))$ *(F is the* ***upper df****)*
- *There exists a random variable Z (common risk) and two non-decreasing functions* h_1, h_2 *such that* $X = h_1(Z)$ *and* $Y = h_2(Z)$.

Stop-loss transform

Definition 9 (Stop-loss).
Let X be a random variable with df F_1. Its **stop-loss transform** is defined as

$$Y_s = \max\{0, X - s\}, \quad s \in \mathbb{R}.$$

Let F_2 be the df of Y_s. Then

$$\mathbb{P}\{X \leq x,\ (X-s)_+ \leq y\} = \min\{F_1(x), F_2(y)\}.$$

Hence, (X, Y_s) is a comonotonic vector.

Countermonotonic risks

Definition 10 (Countermonotonic risks).

Random variables X and Y are said to be **countermonotonic** if there exists **one** random variable $U \sim \mathrm{Unif}(1)$ such that $X = F_1^{-1}(U)$ and $Y = F_2^{-1}(\mathbf{1 - \boldsymbol{U}})$. We say that (X, Y) is a **countermonotonic vector**.

Theorem 11 (Countermonotonicity conditions).

The following conditions are equivalent:

- (X, Y) *is countermonotonic*
- $F(x, y) = \max(F_1(x) + F_2(y) - 1, 0)$ *(F is the* ***lower df****)*
- *There exists a random variable Z and two non-increasing functions h_1, h_2 such that $X = h_1(Z)$ and $Y = h_2(-Z)$.*

Explicit functional dependence with independent generators

Definition 12.
Let $Z_1, \ldots, Z_d$ be **independent** random variables and $q : \mathbb{R}^d \to \mathbb{R}^k$ some function. Define

$$(X_1, \ldots, X_k) = q(Z_1, \ldots, Z_d).$$

Then $(X_1, \ldots, X_k)$ is said to be an **explicit functional dependence model** with **independent generators** $Z_1, \ldots, Z_d$.

Example. The path $(X_1, \ldots, X_n)$ of a **random walk**

$$X_n = \sum_{i=1}^{n} Z_i.$$

Implicit functional dependence

Idea: given independent generators[5] $Z_1, \ldots, Z_d$, we can define a model **implicitly via a system of equations**. For example, **recursively**.

Example. The path $(X_1, \ldots, X_n)$ of a **moving average time series**

$$X_{n+1} = a_n X_n + a_{n-1} X_{n-1}$$

with $X_1 = V_1$ and $X_2 = V_2$.

Example. Define X_{n+1} recursively as the **solution** of the following equation:

$$X_{n+1} + \sum_{i=1}^{n} q_i(X_i) = 0,$$

where q_i are some known functions.

[5] Think of them as *sources of randomness* in the model.

Exercises/questions

- Is the df F corresponding to the following pdf

$$f(x, y) = \frac{4xy + 2x + 2y + 1}{4}, \quad (x, y) \in [0, 1]^2$$

  a **product** df?
- Is the df F corresponding to the pdf $f(x, y) = x + y$ a product df?
- If X and Y have correlation $\rho = \pm 1$, then (X, Y) does not possess a pdf. Why? How is this related to functional dependence?