# Stochastic Models (Lecture #3)

Thomas Verdebout

Université libre de Bruxelles (ULB)

## Independence

We would like to develop a concept of independence.

---

Intuitively, $A$ and $B$ are independent (notation: $A \perp\!\!\!\perp B$) if

$$\mathbb{P}[A|B] = \mathbb{P}[A],$$

or, equivalently, if

$$\mathbb{P}[A]\mathbb{P}[B] = \mathbb{P}[A|B]\mathbb{P}[B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}\mathbb{P}[B] = \mathbb{P}[A \cap B].$$

---

As a definition, we will say that

$$A \text{ and } B \text{ are independent } (A \perp\!\!\!\perp B) \Leftrightarrow \mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B].$$

(since this also covers the cases where $\mathbb{P}[B] = 0$).

# Independence

Examples:

$\mathcal{E}$: dice, $\Omega = \{1, 2, \ldots, 6\}$, $\mathcal{A} = \mathcal{P}(\Omega)$, $\mathbb{P}[A] = \frac{\#A}{\#\Omega}$.

- Are $A = \{6\}$ and $B = \{4, 5, 6\}$ independent?
  No, since
  $$\frac{1}{6} \times \frac{1}{2} = \mathbb{P}[A] \times \mathbb{P}[B] \neq \mathbb{P}[A \cap B] = \frac{1}{6}.$$
- Are $A =$"obtain an even result" and $B = \{4, 5, 6\}$ independent?

  No, since
  $$\frac{1}{2} \times \frac{1}{2} = \mathbb{P}[A] \times \mathbb{P}[B] \neq \mathbb{P}[A \cap B] = \frac{2}{3}.$$

# Independence

Another example:

$\mathcal{E}$: 2 dices, $\Omega = \{(1,1),(1,2),\ldots,(6,6)\}$, $\mathcal{A} = \mathcal{P}(\Omega)$, $\mathbb{P}[A] = \frac{\#A}{\#\Omega}$.

- Are $A =$"obtain an even result for dice 1" and $B =$"obtain an even result for dice 2" independent? Yes, since
$$\mathbb{P}[A]\mathbb{P}[B] = \frac{18}{36} \times \frac{18}{36} = \frac{9}{36} = \mathbb{P}[A \cap B].$$
- Are $A =$"obtain an even result for dice 1" and $C =$"obtain an odd sum" independent? Yes, since
$$\mathbb{P}[A]\mathbb{P}[C] = \frac{18}{36} \times \frac{18}{36} = \frac{9}{36} = \mathbb{P}[A \cap C].$$
- Are $B =$"obtain an even result for dice 2" and $C =$"obtain an odd sum" independent? Yes, since
$$\mathbb{P}[B]\mathbb{P}[C] = \frac{18}{36} \times \frac{18}{36} = \frac{9}{36} = \mathbb{P}[B \cap C].$$

## Independence

In this example, $A, B, C$ are pairwise independent.

But
$$\mathbb{P}[A]\mathbb{P}[B]\mathbb{P}[C] = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \neq 0 = \mathbb{P}[A \cap B \cap C].$$
We will say that $A, B, C$ are not mutually independent...

---

Definition: The events $A_1, A_2, \ldots A_n$ are (mutually) independent if and only if for all $k$, for all $1 \leq i_1 < i_2 < \ldots < i_k \leq n$, we have

$$\mathbb{P}[A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}] = \mathbb{P}[A_{i_1}] \times \mathbb{P}[A_{i_2}] \times \ldots \times \mathbb{P}[A_{i_k}].$$

---

It is clear that the above example violates this definition for $k = 3$. Good, since, intuitively, we didn't want to claim that those 3 events are independent.

## Independence

Let $X_1, X_2, \ldots$ be r.v.'s on $(\Omega, \mathcal{A}, P)$.

- $X_1 \perp\!\!\!\perp X_2 \stackrel{\text{def}}{\Leftrightarrow} \forall B_1, B_2 \in \mathcal{B}$, $[X_1 \in B_1]$ and $[X_2 \in B_2]$ are $\perp\!\!\!\perp$ .
  $(\Leftrightarrow \forall B_1, B_2$, $\mathbb{P}[X_1 \in B_1, X_2 \in B_2] = \mathbb{P}[X_1 \in B_1]\mathbb{P}[X_2 \in B_2])$.
- $X_1, \ldots, X_n$ are $\perp\!\!\!\perp \stackrel{\text{def}}{\Leftrightarrow} \forall B_1, \ldots, B_n \in \mathcal{B}$,
  $[X_1 \in B_1], \ldots, [X_n \in B_n]$ are $\perp\!\!\!\perp$ .
- $X_1, X_2, \ldots$ are $\perp\!\!\!\perp \stackrel{\text{def}}{\Leftrightarrow}$ for every $n \geq 2$, $X_1, X_2, \ldots, X_n$ are $\perp\!\!\!\perp$ .

---

$X_1, \ldots, X_n$ are $\perp\!\!\!\perp \Leftrightarrow$

- $F^X(x_1, \ldots, x_n) = \prod_{i=1}^n F_{X_i}(x_i)$,
- $f^X(x_1, \ldots, x_n) = f^{X_1}(x_1) \ldots f^{X_n}(x_n)$ (provided existence...)

## Independence

If $X_1, \ldots, X_n$ are $\perp\!\!\!\perp$, $\mathbb{E}[X_1 \ldots X_n] = \mathbb{E}[X_1] \ldots \mathbb{E}[X_n]$.

---

**Corollary**: $X \perp\!\!\!\perp Y \Rightarrow \mathrm{Cov}[X, Y] = 0$.

Proof: $X \perp\!\!\!\perp Y \Rightarrow (X - \mathbb{E}[X]) \perp\!\!\!\perp (Y - \mathbb{E}[Y])$, so that $\mathrm{Cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[X - \mathbb{E}[X]]\mathbb{E}[Y - \mathbb{E}[Y]] = 0$.

Therefore, $\mathrm{Cov}[X, Y]$ can be considered as a measure of dependence between $X$ and $Y$.

---

Remarks:

- $\mathrm{Cov}[X, Y] = 0$ does not imply that $X \perp\!\!\!\perp Y$.
  Example: $X$, with pdf $f(.)$, symmetric about 0, and $Y = X^2$.
- $\mathrm{Var}[X + Y] = \mathrm{Var}[X] + \mathrm{Var}[Y] + 2\,\mathrm{Cov}[X, Y]$, so that $\mathrm{Var}[X + Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$ if $X \perp\!\!\!\perp Y$.

# Independence

Independence can be extended to $\sigma$-algebras:

---

Let $\mathcal{A}_1, \mathcal{A}_2, \ldots \subset \mathcal{A}$ be $\sigma$-algebras.

- $\mathcal{A}_1 \perp\!\!\!\perp \mathcal{A}_2 \overset{\text{def}}{\Leftrightarrow} \forall A_1 \in \mathcal{A}_1, \forall A_2 \in \mathcal{A}_2$, $A_1 \perp\!\!\!\perp A_2$.
- $\mathcal{A}_1, \ldots, \mathcal{A}_n$ are $\perp\!\!\!\perp \overset{\text{def}}{\Leftrightarrow} \forall A_1 \in \mathcal{A}_1, \ldots, \forall A_n \in \mathcal{A}_n$, $A_1, \ldots, A_n$ are $\perp\!\!\!\perp$ .
- $\mathcal{A}_1, \mathcal{A}_2, \ldots$ are $\perp\!\!\!\perp \overset{\text{def}}{\Leftrightarrow}$ for every $n \geq 2, \mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n$ are $\perp\!\!\!\perp$ .

# Independence

Denote by

$$\sigma(X) \stackrel{\text{def}}{=} X^{-1}(\mathcal{B})$$

the $\sigma$-algebra generated by the r.v. $X$. (See Lecture # 1.)

- Then $X_1 \perp\!\!\!\perp X_2 \Leftrightarrow \sigma(X_1) \perp\!\!\!\perp \sigma(X_2)$.
- This also allows for defining independence between r.v. and $\sigma$-algebras:

$$X_1, \dots, X_n, \mathcal{A}_1, \dots, \mathcal{A}_n \text{ are } \perp\!\!\!\perp$$

$$\Updownarrow$$

$$\sigma(X_1), \dots, \sigma(X_n), \mathcal{A}_1, \dots, \mathcal{A}_n \text{ are } \perp\!\!\!\perp .$$

## Conditional expectation

In the previous lecture, we studied how the concept of expectation can be used to compute the "most typical value" or "best guess" for some r.v. $X$.

Conditional expectation allows for exploiting some information (e.g., the information that some event occurred) in order to improve this (unconditional) best guess.

This information can take various forms:

- (the occurrence of) an event.
- (the value of ) a r.v.
- (the occurrence of some event in) a $\sigma$-algebra.

## Conditional expectation (w.r.t. an event)

Let $X$ be an integrable r.v. on $(\Omega, \mathcal{A}, \mathbb{P})$.
Let $A \in \mathcal{A}$, with $\mathbb{P}[A] \neq 0$.

Then we let

$$\mathbb{E}[X|A] \stackrel{\text{def}}{=} \frac{1}{\mathbb{P}[A]} \int_A X(\omega)\, d\mathbb{P}(\omega) \left[= \int_\Omega X(\omega) d\mathbb{P}(\omega|A)\right].$$

Remarks:

- $\mathbb{E}[X|A] = \frac{1}{\mathbb{P}[A]} \int_\Omega X(\omega)\mathbb{I}_A(\omega)\, d\mathbb{P}(\omega). = \frac{\mathbb{E}[X\mathbb{I}_A]}{\mathbb{P}[A]}$.
- $\mathbb{E}[X|\Omega] = \mathbb{E}[X]$.
  Intuitively, the information that $\Omega$ has occurred is void, so that the "initial best guess" $\mathbb{E}[X]$ cannot be improved.
- $\mathbb{E}[\mathbb{I}_{A_1}|A_2] = \mathbb{P}[A_1|A_2]$ (exercise) .

$\rightsquigarrow$ Two examples...

# Conditional expectation (w.r.t. an event)

(A) assume you bet $b$ euros on "even" in one game of roulette...

$\mathcal{E}$: roulette, $\Omega = \{0, 1, 2, \ldots, 36\}$, $\mathcal{A} = \mathcal{P}(\Omega)$, $\mathbb{P}[A] = \frac{\#A}{\#\Omega}$.

Your (random) gain $X$ is given by

$$X(\omega) = \begin{cases} -b & \text{if} \quad \omega \in \{0, 1, 3, \ldots, 35\} \\ b & \text{if} \quad \omega \in \{2, 4, \ldots, 36\}, \end{cases}$$

whose distribution is

| distribution of $X$ | | |
|---|---|---|
| values | $-b$ | $b$ |
| probabilities | $\frac{19}{37}$ | $\frac{18}{37}$ |

with expected winning $-\frac{b}{37}$.

# Conditional expectation (w.r.t. an event)

(A) assume you bet $b$ euros on "even" in one game of roulette...

$\mathcal{E}$: roulette, $\Omega = \{0, 1, 2, \ldots, 36\}$, $\mathcal{A} = \mathcal{P}(\Omega)$, $\mathbb{P}[A] = \frac{\#A}{\#\Omega}$.

Now, we are given the information that the event $A$="result is not 0" occurred? What is then your expected winning, namely $\mathbb{E}[X|A]$?

The r.v. $X\mathbb{I}_A$ is determined by

$$(X\mathbb{I}_A)(\omega) = \begin{cases} -b & \text{if} \quad \omega \in \{1, 3, \ldots, 35\} \\ 0 & \text{if} \quad \omega = 0 \\ b & \text{if} \quad \omega \in \{2, 4, \ldots, 36\}, \end{cases}$$

whose distribution is

| distribution of $X\mathbb{I}_A$ | | | |
|---|---|---|---|
| values | $-b$ | 0 | $b$ |
| probabilities | $\frac{18}{37}$ | $\frac{1}{37}$ | $\frac{18}{37}$ |

## Conditional expectation (w.r.t. an event)

Therefore,

your expected winning (conditional upon $A$) is

$$\mathbb{E}[X|A] = \frac{1}{\mathbb{P}[A]}\mathbb{E}[X\mathbb{I}_A] = \frac{1}{(\frac{36}{37})}\left((-b)\times\frac{18}{37}+0\times\frac{1}{37}+b\times\frac{18}{37}\right) = 0.$$

The fact that we know 0 won't be the result of the game, transforms this defavorable game ($\mathbb{E}[X] = -b/37$) into a fair one. $\mathbb{E}[X|A] = 0$, i.e., the game is neither favorable nor defavorable.

## Conditional expectation (w.r.t. an event)

(B) assume that three coins (10, 20, and 50 cents) are tossed and that you win the coins that show "head"...

$\mathcal{E}$: 3 coins, $\Omega = \{HHH, HHT, \ldots, TTT\}$, $\mathcal{A} = \mathcal{P}(\Omega)$, $\mathbb{P}[A] = \frac{\#A}{\#\Omega}$.

Denote by $X$ your gain or loss.

Denote by $A$ the event "two heads and one tail",

that is, $A = \{HHT, HTH, THH\}$.

What is your conditional expected winning $\mathbb{E}[X|A]$?

The r.v. $X\mathbb{I}_A$ is determined by

$$(X\mathbb{I}_A)(\omega) = \begin{cases} 10+20 & \text{if} \quad \omega = HHT \\ 10+50 & \text{if} \quad \omega = HTH \\ 20+50 & \text{if} \quad \omega = THH \\ 0 & \text{if} \quad \omega \notin A. \end{cases}$$

# Conditional expectation (w.r.t. an event)

The corresponding distribution is

| distribution of $X\mathbb{I}_A$ | | | | |
|---|---|---|---|---|
| values | 30 | 60 | 70 | 0 |
| probabilities | $\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{5}{8}$ |

Therefore,

your expected winning (conditional upon $A$) is

$$\mathbb{E}[X|A] = \frac{\mathbb{E}[X\mathbb{I}_A]}{\mathbb{P}[A]} = \frac{1}{(\frac{3}{8})}\left(30 \times \frac{1}{8} + \frac{60}{8} + \frac{70}{8} + 0 \times \frac{5}{8}\right) = 53.33\ldots$$

This is to be compared with (exercise) the unconditional expected winning $\mathbb{E}[X] = 40$ (interpretation).

# Conditional expectation

Conditional expectation allows for exploiting some information (e.g., the information that some event occurred) in order to improve the (unconditional) best guess $\mathbb{E}[X]$.

This available information can take various forms:

- (the occurrence of) an event.
- (the value of ) a r.v.
- (the occurrence of some event in) a $\sigma$-algebra.

## Conditional expectation (w.r.t. a r.v.)

Let $X$ be an integrable r.v. on $(\Omega, \mathcal{A}, \mathbb{P})$.
Let $Y$ be a discrete r.v. on $(\Omega, \mathcal{A}, \mathbb{P})$, say with distribution

| distribution of $Y$ | | | |
|---|---|---|---|
| values | $y_1$ | $y_2$ | $\ldots$ |
| probabilities | $p_1$ | $p_2$ | $\ldots$ |

$Y$ partitions $\Omega$ via

$$\Omega = \{\omega | Y(\omega) = y_1\} \cup \{\omega | Y(\omega) = y_2\} \cup \cdots$$

Then we define $\mathbb{E}[X|Y]$ as the r.v.

$$\mathbb{E}[X|Y](\omega') = E[X|\{\omega|Y(\omega) = y_i\}] \quad \text{if} \quad \omega' \in \{\omega|Y(\omega) = y_i\}.$$

The last conditional expectation is w.r.t. an event, and hence is well defined.

# Conditional expectation

Note: we can interpret $\mathbb{E}[X|Y]$ as a function of $\omega$ or as a function of $y$:

$$E[X|Y](\omega) \quad \text{or} \quad E[X|Y=y].$$

It can be shown that:

**Theorem:**
(i) $\mathbb{E}[X|Y]$ is $\sigma(Y)$-measurable [that is: $\mathbb{E}[X|Y] = f(Y)$]
(ii) $\int_A \mathbb{E}[X|Y](\omega)\,dP(\omega) = \int_A X(\omega)\,dP(\omega)$, for all $A \in \sigma(Y)$.

Actually, this characterizes the r.v. $\mathbb{E}[X|Y]$.

---

This will be used to extend the concept of conditional expectation to more general setups...

## Conditional expectation

Some remarks:

- If a r.v. $Z$ is $\sigma(Y)$–measurable, then this means that $Z = f(Y)$ for some (measurable) function $f$.
- If $E[X^2]$ is finite, then an alternative way to define $E[X|Y]$ is this one:

  $E[X|Y] = f_0(Y)$, with $f_0$ is the function which minimizes the mean squared loss:

  $$E[X - f_0(Y)]^2 \leq E[X - f(Y)]^2, \quad \forall f \quad \text{such that } E[f(Y)]^2 < \infty.$$

  We perform a non-linear regression!

## Conditional expectation (w.r.t. a r.v.)

Let $X$ be an integrable r.v. on $(\Omega, \mathcal{A}, P)$.

Then we define quite generally $\mathbb{E}[X|Y]$ as the r.v. such that

(i) $\mathbb{E}[X|Y]$ is $\sigma(Y)$-measurable.

(ii) $\int_A \mathbb{E}[X|Y](\omega)\, dP(\omega) = \int_A X(\omega)\, dP(\omega)$, for all $A \in \sigma(Y)$.

---

In practice:

- If $(X, Y)$ is discrete, $\mathbb{E}[X|Y = y] = \sum_i x_i \mathbb{P}[X = x_i|Y = y]$.
- If $(X, Y)$ is absolutely continuous,

$$\mathbb{E}[X|Y = y] = \frac{1}{f^Y(y)} \int_{\mathbb{R}} x\, f^{(X,Y)}(x, y)\, dx = \int_{\mathbb{R}} x\, f^{(X|Y)}(x, y)\, dx,$$

  where $f^{(X|Y)}(x, y)$ denotes the pdf of $X$ conditionally on $Y = y$.

## Conditional expectation (w.r.t. a r.v.)

Same properties:

- $\mathbb{E}[X|Y]$ is $\sigma(Y)$-measurable
- $\mathbb{E}[X|Y] = \mathbb{E}[X]$ iff $X \perp\!\!\!\perp Y$.
- $\mathbb{E}[X|Y] = X$ iff $X$ is $\sigma(Y)$-measurable.
- $\mathbb{E}\Big[\mathbb{E}[X|Y]\Big] = \mathbb{E}[X]$.

---

Remark: $\mathbb{E}[X] = \mathbb{E}\Big[\mathbb{E}[X|Y]\Big] = \int_{\mathbb{R}} \mathbb{E}[X|Y=y]\, f^Y(y)\, dy$.
(// total probability formula).

We still define: $\mathbb{P}[A|Y] = \mathbb{E}[\mathbb{I}_A|Y]]$.

$\rightsquigarrow \mathbb{P}[A] = \mathbb{E}[\mathbb{I}_A] = \mathbb{E}\Big[\mathbb{E}[\mathbb{I}_A|Y]\Big] = \int_{\mathbb{R}} \mathbb{E}[\mathbb{I}_A|Y=y]\, f^Y(y)\, dy = \int_{\mathbb{R}} \mathbb{P}[A|Y=y]\, f^Y(y)\, dy$.

## Conditional variances

In time series analysis, many models are designed to explain the dynamics of the conditional variance $\mathrm{Var}[X_t|X_{t-1}, X_{t-2}, \ldots]$ (e.g., the so-called stochastic volatility models). We define

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

---

Remarks:

- Similarly (exercise) as for $\mathrm{Var}[X]$,

$$\mathrm{Var}[X|Y] = \mathbb{E}[X^2|Y] - \big(\mathbb{E}[X|Y]\big)^2.$$

- But $\mathrm{Var}[X] \neq \mathbb{E}[\mathrm{Var}[X|Y]]$, rather (exercise) we have

$$\mathrm{Var}[X] = \mathbb{E}\big[\mathrm{Var}[X|Y]\big] + \mathrm{Var}\big[\mathbb{E}[X|Y]\big].$$

Example:

Let $X \sim \mathrm{Unif}(0,1)$. If $X = x$, you flip $m$ times a coin with property $\mathbb{P}[\mathrm{Head}] = x$. Let $N$ be the number of "heads". $\mathbb{E}[N] =$? Distribution of $N$?

---

$$\begin{aligned}\mathbb{E}[N|X=x] &= \sum_{n=0}^{m} n\,\mathbb{P}[N=n|X=x] = \sum_{n=0}^{m} n\binom{m}{n}x^n(1-x)^{m-n} \\ &= mx \quad \text{expectation of a Binomial random variable.}\end{aligned}$$

$\Rightarrow \mathbb{E}[N|X] = mX$ ($\sigma(X)$-measurable).

Consequently, $\mathbb{E}[N] = \mathbb{E}\Big[\mathbb{E}[N|X]\Big] = \mathbb{E}[mX] = m\,\mathbb{E}[X] = \frac{m}{2}$.

# Conditional expectation (w.r.t. a r.v.)

Distribution of $N$?

---

$$\begin{aligned}\mathbb{P}[N=n] &= \int_{\mathbb{R}} \mathbb{P}[N=n|X=x]\, f^X(x)\, dx = \int_0^1 \mathbb{P}[N=n|X=x]\, dx \\ &= \int_0^1 \binom{m}{n} x^n (1-x)^{m-n}\, dx = \binom{m}{n} \int_0^1 x^n (1-x)^{m-n}\, dx \\ &= \ldots = \frac{1}{m+1},\end{aligned}$$

for all $n = 0, 1, \ldots, m$.

Hence, $N$ is uniformly distributed on $\{0, 1, \ldots, m\}$.

(which now makes clear why $\mathbb{E}[N] = \frac{m}{2}$).

## Conditional expectation (w.r.t. a r.v.)

Example:

A hen lays $N$ eggs, where $N \sim \mathcal{P}(\lambda)$. Each egg hatches out with probability $p$. Let $K$ be number of born chicken.

$$\mathbb{E}[K|N] =? \text{ and } \mathbb{E}[K] =?.$$

---

$\mathbb{E}[K|N](\omega) = \mathbb{E}[K|N = n]$ if $N(\omega) = n$, where

$$\mathbb{E}[K|N = n] = \sum_k k\mathbb{P}[K = k|N = n] = \sum_{k=0}^{n} k \binom{n}{k} p^k q^{n-k}$$
$$= np \quad \text{expectation of a Binomial r.v.}$$

$\Rightarrow \mathbb{E}[K|N] = Np$ ($\sigma(N)$-measurable).

Consequently, $\mathbb{E}[K] = \mathbb{E}\Big[\mathbb{E}[K|N]\Big] = \mathbb{E}[Np] = p\,\mathbb{E}[N] = p\lambda$.

# Conditional expectation (w.r.t. a r.v.)

Example:

Let $X \sim \mathrm{Exp}(\lambda_1)$ and $Y \sim \mathrm{Exp}(\lambda_2)$ be two independent exponential random variables. What is the distribution of $V = X + Y$?

---

We use total probability formula and assume $z \geq 0$. (The case $z < 0$ is trivial.)

$$\begin{aligned}
\mathbb{P}[V \leq z] &= \int_{\mathbb{R}} P[V \leq z | Y = y] f^Y(y) dy \\
&= \int_{\mathbb{R}} P[X + Y \leq z | Y = y] f^Y(y) dy \\
&= \int_{\mathbb{R}} P[X \leq z - y] f^Y(y) dy \\
&= \int_{\mathbb{R}} (1 - e^{-\lambda_1 (z-y)}) \mathbb{I}[z - y \geq 0] \lambda_2 e^{-\lambda_2 y} \mathbb{I}[y \geq 0] dy \\
&= \int_0^z (1 - e^{-\lambda_1 (z-y)}) \lambda_2 e^{-\lambda_2 y} dy = \cdots
\end{aligned}$$

## Conditional expectation

Conditional expectation allows for exploiting some information (e.g., the information that some event occurred) in order to improve the (unconditional) best guess $\mathbb{E}[X]$.

This available information can take various forms:

- (the occurrence of) an event.
- (the value of ) a r.v.
- (the occurrence of some event in) a $\sigma$-algebra.

## Conditional expectation (w.r.t. a $\sigma$-algebra)

Let $X$ be an integrable r.v. on $(\Omega, \mathcal{A}, \mathbb{P})$.

Let $\mathcal{F}$ be a $\sigma$-algebra $\subset \mathcal{A}$.

Then we define $\mathbb{E}[X|\mathcal{F}]$ as <u>the</u> r.v. such that

(i) $\mathbb{E}[X|\mathcal{F}]$ is $\mathcal{F}$-measurable.

(ii) $\int_A \mathbb{E}[X|\mathcal{F}](\omega)\, d\mathbb{P}(\omega) = \int_A X(\omega)\, d\mathbb{P}(\omega)$, for all $A \in \mathcal{F}$.

---

Same kind of properties as for $\mathbb{E}[X|Y]$:

- $\mathbb{E}[X|\mathcal{F}]$ is $\mathcal{F}$-measurable
- $\mathbb{E}[X|\mathcal{F}] = \mathbb{E}[X]$ iff $X \perp\!\!\!\perp \mathcal{F}$.
- $\mathbb{E}[X|\mathcal{F}] = X$ iff $X$ is $\mathcal{F}$-measurable.
- $\mathbb{E}\Big[\mathbb{E}[X|\mathcal{F}]\Big] = \mathbb{E}[X]$.

## Conditional expectation (w.r.t. a $\sigma$-algebra)

Let $\mathcal{F}_1 \subset \mathcal{F}_2$ be two $\sigma$-algebras in $\mathcal{A}$.

Extra properties:

- $\mathbb{E}\Big[\mathbb{E}[X|\mathcal{F}_2]\Big|\mathcal{F}_1\Big] = \mathbb{E}[X|\mathcal{F}_1]$.
- $\mathbb{E}\Big[\mathbb{E}[X|\mathcal{F}_1]\Big|\mathcal{F}_2\Big] = \mathbb{E}[X|\mathcal{F}_1]$.

---

Conditional expectation w.r.t. to a $\sigma$-algebra also allows for defining conditional expectation w.r.t. a collection of r.v.'s. More specifically, we define

$$\mathbb{E}[X|Y_1, \ldots, Y_n] := \mathbb{E}[X|\sigma(Y_1, \ldots, Y_n)],$$

where $\sigma(Y_1, \ldots, Y_n)$ is the smallest $\sigma$-algebra containing $\{Y_i^{-1}(B) \,|\, B \in \mathcal{B},\, i = 1, \ldots, n\}$.