Probability Space

80-646-08
Stochastic Calculus I

Geneviève Gauthier

HEC Montréal

Sample space

Definition

The *sample space* $\Omega = \{\omega_i : i \in \mathcal{I}\}$ is the set of all possible outcomes of a random experiment. Let $\mathcal{I}$ be a set of indices. For example, $\mathcal{I} = \{0, 1, 2, ..., T\}$, $\mathcal{I} = \{0, 1, 2, ...\}$, $\mathcal{I} = [0, \infty)$, etc.

Example. If the random experiment consists of rolling a dice, then

$$\Omega = \left\{ \boxed{1}, \boxed{2}, \boxed{3}, \boxed{4}, \boxed{5}, \boxed{6} \right\}.$$

Event

Definition

(incomplete). An *event* is a subset of Ω.

Example.

$$A = \text{the result is even} = \left\{ \boxed{2}, \boxed{4}, \boxed{6} \right\}.$$

Random variables I

Incomplete definition

Definition

(incomplete). A *random variable* $X : \Omega \to \mathbb{R}$ is a function mapping the sample space (its domain) to real numbers $\mathbb{R}$.

Beware! "Sample space" and "random variable" are concepts that should not be confused.

Random variables II

Incomplete definition

Example. If the random experiment consists of choosing a card at random from a deck of 52 cards, then the event "drawing the king of hearts" is not a random variable, since, among other things, "king of hearts" is not a real number. However, if drawing a certain card is associated with 10 points, then such a relationship is a random variable.
That's why we have chosen to use boxed numbers to denote the possible results of a dice roll: the idea is to distinguish between the event $\left\{\boxed{4}\right\} =$ "the side showing four dots " from the random variable that associates each of the sides of the dice to the number of dots on that side.

Random variables

Example

Example. X, Y, Z and W are random variables:

ω	$X(\omega)$	$Y(\omega)$	$Z(\omega)$	$W(\omega)$
1	0	0	0	5
2	0	5	0	5
3	5	5	0	5
4	5	5	5	5
5	10	5	10	0
6	10	10	10	10

Probability measures

Incomplete definition

Definition

(incomplete). $\mathbb{P}$ is a *probability measure* on the space Ω if :

P1 $\mathbb{P}(\Omega) = 1$.

P2 For any event A in Ω, $0 \leq \mathbb{P}(A) \leq 1$.

P3 For any mutually disjoint events A_1, A_2, ..., $\mathbb{P}\left(\bigcup_{i \geq 1} A_i\right) = \sum_{i \geq 1} \mathbb{P}(A_i)$ where two events A_i and A_j are disjoint if $A_i \cap A_j = \varnothing$.

Probability measures

Example

Example. The probability measure $\mathbb{P}$ represents the situation where the dice is well balanced, while $\mathbb{Q}$ models a case where the dice is loaded.

ω	$\mathbb{P}(\omega)$	$\mathbb{Q}(\omega)$	ω	$\mathbb{P}(\omega)$	$\mathbb{Q}(\omega)$
$\boxed{1}$	$\frac{1}{6}$	$\frac{4}{12}$	$\boxed{4}$	$\frac{1}{6}$	$\frac{1}{12}$
$\boxed{2}$	$\frac{1}{6}$	$\frac{1}{12}$	$\boxed{5}$	$\frac{1}{6}$	$\frac{1}{12}$
$\boxed{3}$	$\frac{1}{6}$	$\frac{1}{12}$	$\boxed{6}$	$\frac{1}{6}$	$\frac{4}{12}$

Probability measures

Equivalent definition

Theorem

When Card $(\Omega) < \infty$, say $\Omega = \{\omega_1, ..., \omega_n\}$ then the three conditions $(P1)$, $(P2)$ and $(P3)$ in the partial definition of a probability measure are equivalent to the three following conditions:

- P1* $\forall i \in \{1, ..., n\}$, $\mathbb{P}(\omega_i) \geq 0$.
- P2* For any event A in Ω, $\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega)$.
- P3* $\sum_{i=1}^{n} \mathbb{P}(\omega_i) = 1$.

The proof can be found in the appendix.

Probability measures

Properties

Theorem

Probability measures have the following properties:

- P4 For any event A in Ω, $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.
- P5 $\mathbb{P}(\varnothing) = 0$.
- P6 For any two events A and B in Ω (not necessarily disjoint), $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.
- P7 If $A \subseteq B \subseteq \Omega$ then $\mathbb{P}(A) \leq \mathbb{P}(B)$.

Probability measures I

Proofs

To be shown: *For any event A of Ω, $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.*

Proof of $(P4)$.

$$1 = \mathbb{P}(\Omega) = \mathbb{P}(A \cup A^c) = \mathbb{P}(A) + \mathbb{P}(A^c)$$

where the first equality comes from $(P1)$ and the third equality comes from $(P3)$. ■

Probability measures II

Proofs

To be shown: $\mathbb{P}(\varnothing) = 0$.

Proof of $(P5)$. Property $(P5)$ is nothing more than a special case of $(P4)$: let's replace A with Ω. ■

Probability measures III

Proofs

To be shown: *for any two events A and B of Ω not necessarily disjoint),*

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

Proof of $(P6)$. Since

$$A = (A \cap B) \cup (A \cap B^c) \text{ and } B = (B \cap A) \cup (B \cap A^c)$$

then, using $(P3)$, we get

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c) \text{ and } \mathbb{P}(B) = \mathbb{P}(B \cap A) + \mathbb{P}(B \cap A^c).$$

Probability measures IV

Proofs

On the other hand, $A \cup B = (A \cap B^c) \cup (B \cap A^c) \cup (A \cap B)$, which implies that

$$
\begin{aligned}
& \mathbb{P}(A \cup B) \\
= \; & \mathbb{P}(A \cap B^c) + \mathbb{P}(B \cap A^c) + \mathbb{P}(A \cap B) \\
= \; & [\mathbb{P}(A \cap B^c) + \mathbb{P}(A \cap B)] + [\mathbb{P}(B \cap A^c) + \mathbb{P}(A \cap B)] \\
& - \mathbb{P}(A \cap B) \\
= \; & \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) . \blacksquare
\end{aligned}
$$

Probability measures V

Proofs

To be shown: *If* $A \subseteq B \subseteq \Omega$ *then* $\mathbb{P}(A) \leq \mathbb{P}(B)$.

Proof of $(P7)$. Since $A \subseteq B$ then $A \cap B = A$. Using $(P3)$,

$$\mathbb{P}(B) = \mathbb{P}(A \cap B) + \underbrace{\mathbb{P}(A^c \cap B)}_{\geq 0 \text{ from } (P2)} \geq \mathbb{P}(A \cap B) = \mathbb{P}(A). \blacksquare$$

Probability measures built on Ω exist independently from the random variables and vice versa. What is the link between them? That's the topic of the next section.

Distribution (law)

Definition

Definition

The *distribution* or the *law* of a random variable X is characterized by its *(cumulative) distribution function*

$$F_X \quad : \quad \mathbb{R} \rightarrow [0,1]$$
$$x \rightarrow \text{probability that the r.v. } X \text{ is less than or equal to } x.$$

So, if $\mathbb{P}$ is the probability measure assigned to Ω then

$$\forall x \in \mathbb{R},\ F_X(x) = \mathbb{P}\{\omega \in \Omega \,|X(\omega) \leq x\}.$$

Distribution (law)

Alternative characterization

Theorem

In the case where Card $(\Omega) < \infty$, the distribution of a random variable is also characterized by its probability mass function

$$f_X : \mathbb{R} \to [0,1]$$
$$x \to \text{probability that the r.v. } X \text{ is equal to } x,$$

that is

$$\forall x \in \mathbb{R},\ f_X(x) = \mathbb{P}\{\omega \in \Omega \,|\, X(\omega) = x\}.$$

A proof of this result can be found in the appendix.

Distribution (law) I

Example

ω	$W(\omega)$	$\mathbb{Q}(\omega)$	ω	$W(\omega)$	$\mathbb{Q}(\omega)$
$\boxed{1}$	5	$\frac{4}{12}$	$\boxed{4}$	5	$\frac{1}{12}$
$\boxed{2}$	5	$\frac{1}{12}$	$\boxed{5}$	0	$\frac{1}{12}$
$\boxed{3}$	5	$\frac{1}{12}$	$\boxed{6}$	10	$\frac{4}{12}$

Distribution (law) II

Example

Let's find the probability mass function and the cumulative distribution function of the random variable W.

$$f_W(x) = \mathbb{Q}\{\omega \in \Omega \,|\, W(\omega) = x\}$$

$$= \begin{cases} \mathbb{Q}\{\boxed{5}\} & \text{if } x = 0 \\ \mathbb{Q}\{\boxed{1}, \boxed{2}, \boxed{3}, \boxed{4}\} & \text{if } x = 5 \\ \mathbb{Q}\{\boxed{6}\} & \text{if } x = 10 \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} \frac{1}{12} & \text{if } x = 0 \\ \frac{4}{12} + \frac{1}{12} + \frac{1}{12} + \frac{1}{12} = \frac{7}{12} & \text{if } x = 5 \\ \frac{4}{12} & \text{if } x = 10 \\ 0 & \text{otherwise} \end{cases}.$$

Distribution (law) III

Example

Recall that:

ω	$W(\omega)$	$\mathbb{Q}(\omega)$	ω	$W(\omega)$	$\mathbb{Q}(\omega)$
$\boxed{1}$	5	$\frac{4}{12}$	$\boxed{4}$	5	$\frac{1}{12}$
$\boxed{2}$	5	$\frac{1}{12}$	$\boxed{5}$	0	$\frac{1}{12}$
$\boxed{3}$	5	$\frac{1}{12}$	$\boxed{6}$	10	$\frac{4}{12}$

Distribution (law) IV

Example

Let's calculate the cumulative distribution function:

$$
\begin{aligned}
\text{if } x &< 0 \text{ then} \\
F_W(x) &= \mathbb{Q}\{\omega \in \Omega \,|\, W(\omega) \leq x\} = \mathbb{Q}(\varnothing) = 0; \\
\text{if } 0 &\leq x < 5 \text{ then} \\
F_W(x) &= \mathbb{Q}\{\omega \in \Omega \,|\, W(\omega) \leq x\} = \mathbb{Q}\left\{\boxed{5}\right\} = \frac{1}{12}; \\
\text{if } 5 &\leq x < 10 \text{ then} \\
F_W(x) &= \mathbb{Q}\{\omega \in \Omega \,|\, W(\omega) \leq x\} = \mathbb{Q}\left\{\boxed{1}, \boxed{2}, \boxed{3}, \boxed{4}, \boxed{5}\right\} = \frac{8}{12}; \\
\text{if } x &\geq 10 \text{ then} \\
F_W(x) &= \mathbb{Q}\{\omega \in \Omega \,|\, W(\omega) \leq x\} = \mathbb{Q}(\Omega) = 1.
\end{aligned}
$$

Distribution (law)

Cumulative distribution function properties

R1 It's a non-decreasing function, that is to say that if $x < y$ then $F_X(x) \leq F_X(y)$.

R2 It's a function that is right-continuous with left limits.

R3 $\lim_{x \downarrow -\infty} F_X(x) = 0$ and $\lim_{x \uparrow \infty} F_X(x) = 1$.

Distribution (law)

Notation

Remark. In order to alleviate the notation, it is common to write $\{X \leq x\}$ instead of $\{\omega \in \Omega \, | X(\omega) \leq x\}$ and $\mathbb{P}\{X \leq x\}$ instead of $\mathbb{P}\{\omega \in \Omega \, | X(\omega) \leq x\}$.

Distribution (law) I

Example

ω	$X(\omega)$	$Y(\omega)$	$Z(\omega)$	$W(\omega)$	$\mathbb{P}(\omega)$	$\mathbb{Q}(\omega)$
1	0	0	0	5	$\frac{1}{6}$	$\frac{4}{12}$
2	0	5	0	5	$\frac{1}{6}$	$\frac{1}{12}$
3	5	5	0	5	$\frac{1}{6}$	$\frac{1}{12}$
4	5	5	5	5	$\frac{1}{6}$	$\frac{1}{12}$
5	10	5	10	0	$\frac{1}{6}$	$\frac{1}{12}$
6	10	10	10	10	$\frac{1}{6}$	$\frac{4}{12}$

Distribution (law) II

Example

Distributions of the random variables X, Y, Z and W under the probability measure $\mathbb{P}$

x	$\mathbb{P}\{X = x\}$	$\mathbb{P}\{Y = x\}$	$\mathbb{P}\{Z = x\}$	$\mathbb{P}\{W = x\}$
0	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{6}$
5	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{1}{6}$	$\frac{2}{3}$
10	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{6}$

Remark. The distribution of the random variable X is said to be uniform since

$$\mathbb{P}\{X = 0\} = \mathbb{P}\{X = 5\} = \mathbb{P}\{X = 10\} = \frac{1}{3}.$$

Distribution (law) III

Example

Recall that:

Distributions of the random variables X, Y, Z and W under the probability measure $\mathbb{P}$

x	$\mathbb{P}\{X = x\}$	$\mathbb{P}\{Y = x\}$	$\mathbb{P}\{Z = x\}$	$\mathbb{P}\{W = x\}$
0	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{6}$
5	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{1}{6}$	$\frac{2}{3}$
10	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{6}$

Remark. The random variables Y and W have the same distribution, although they are not equal. Indeed,

$$Y\left(\boxed{1}\right) = 0 \neq 5 = W\left(\boxed{1}\right).$$

Distribution (law) IV

Example

Distributions of the random variables X, Y, Z and W under the probability measure $\mathbb{Q}$

x	$\mathbb{Q}\{X = x\}$	$\mathbb{Q}\{Y = x\}$	$\mathbb{Q}\{Z = x\}$	$\mathbb{Q}\{W = x\}$
0	$\frac{5}{12}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{12}$
5	$\frac{2}{12}$	$\frac{1}{3}$	$\frac{1}{12}$	$\frac{7}{12}$
10	$\frac{5}{12}$	$\frac{1}{3}$	$\frac{5}{12}$	$\frac{1}{3}$

Remark. Let's note that the distributions of the random variables have changed. Moreover, under the probability measure $\mathbb{Q}$, the random variables Y and W don't have the same distribution any more.

Random variables

Equality and equality in law

Definition

Two random variables X and Y are said to be *equal* if and only if $\forall \omega \in \Omega$, $X(\omega) = Y(\omega)$.
They are said to be *equal in distribution* (or in law) when they have the same distribution.

1. The concept of equality between two random variables is stronger than the concept of equality in distribution. Indeed, if two random variables are equal, then they are equal in distribution.
2. However, two random variables may be equal in distribution but not equal.
3. Moreover, two random variables may be equal in distribution under a certain probability measure but not be equal under another probability measure.

In the previous example, when the probability measure P is assigned to Ω, Y and W are equal in distribution but they are not equal.

Sigma-Algebra

Introduction

Question. If $Card\,(\Omega) = n < \infty$, how many distinct events are there?

Answer 2^n. An event is a subset of $\Omega = \{\omega_1, ..., \omega_n\}$. When we create an event $A \subseteq \Omega$, we have, for every ω_i, two alternatives: either $\omega_i \in A$ or $\omega_i \notin A$.

Example. If $n = 3$ then

$\omega_1 \in A$	$\omega_2 \in A$	$\omega_3 \in A$	$\{\omega_1, \omega_2, \omega_3\} = \Omega$
		$\omega_3 \notin A$	$\{\omega_1, \omega_2\}$
	$\omega_2 \notin A$	$\omega_3 \in A$	$\{\omega_1, \omega_3\}$
		$\omega_3 \notin A$	$\{\omega_1\}$
$\omega_1 \notin A$	$\omega_2 \in A$	$\omega_3 \in A$	$\{\omega_2, \omega_3\}$
		$\omega_3 \notin A$	$\{\omega_2\}$
	$\omega_2 \notin A$	$\omega_3 \in A$	$\{\omega_3\}$
		$\omega_3 \notin A$	$\varnothing$

Sigma-Algebra

Introduction

Usually, we don't need to know the probabilities associated with every event in Ω. Such a situation is particularly frequent when $\mathit{Card}\left(\Omega\right)$ is large or infinite.

Sigma-Algebra

Example

Example. The sample space is $\Omega = \left\{\boxed{1}, \boxed{2}, \boxed{3}, \boxed{4}, \boxed{5}, \boxed{6}\right\}$. Let's assume we are interested in the random variable X only.

ω	$X(\omega)$	ω	$X(\omega)$
$\boxed{1}$	0	$\boxed{4}$	5
$\boxed{2}$	0	$\boxed{5}$	10
$\boxed{3}$	5	$\boxed{6}$	10

The events that characterize the distribution of X are

$$\{X = 0\} = \left\{\boxed{1}, \boxed{2}\right\}; \{X = 5\} = \left\{\boxed{3}, \boxed{4}\right\}; \{X = 10\} = \left\{\boxed{5}, \boxed{6}\right\}.$$

So, in order to find the distribution of X, we only need to know the probabilities associated with the events $\left\{\boxed{1}, \boxed{2}\right\}$, $\left\{\boxed{3}, \boxed{4}\right\}$ and $\left\{\boxed{5}, \boxed{6}\right\}$. Knowing, in addition, the probability that event $\left\{\boxed{1}\right\}$ occurs doesn't give us any additional information about the distribution of the random variable X.

Sigma-Algebra I

Definition

The properties of a probability measure are such that, if we know the probability that an event A occurs, then we also know the probability associated with its complement A^c since $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$. We are also able to determine the probability associated with the union of a certain number of events characterizing the distribution of X because of property $(P6)$.

Sigma-Algebra II

Definition

Definition

A $\sigma-$algebra $\mathcal{F}$ of Ω is a subset of events such that

T1 $\Omega \in \mathcal{F}$.

T2 If $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$.

T3 If $A_1, A_2, ... \in \mathcal{F}$ then $\bigcup_{i \geq 1} A_i \in \mathcal{F}$. In the case where $Card\,(\Omega) < \infty$, the condition $(T3)$ is equivalent to

$$(T3^*) \quad \text{If } A_1, A_2, ..., A_n \in \mathcal{F} \text{ then } \bigcup_{i \geq 1}^{n} A_i \in \mathcal{F}.$$

Intuitively, the $\sigma-$algebra is the set of events in which we are interested.

Sigma-Algebra

Examples

- **Example**. The trivial $\sigma-$algebra: $\{\varnothing, \Omega\}$.
- **Example.** The smallest $\sigma-$algebra containing the event A is $\{\varnothing, A, A^c, \Omega\}$.
- **Example.** $\Omega = \left\{\boxed{1}, \boxed{2}, \boxed{3}, \boxed{4}, \boxed{5}, \boxed{6}\right\}$. The smallest $\sigma-$algebra containing the events $\left\{\boxed{1}, \boxed{2}\right\}$, $\left\{\boxed{3}, \boxed{4}\right\}$ and $\left\{\boxed{5}, \boxed{6}\right\}$ is

$$\left\{ \begin{array}{c} \varnothing, \left\{\boxed{1}, \boxed{2}\right\}, \left\{\boxed{3}, \boxed{4}\right\}, \left\{\boxed{5}, \boxed{6}\right\}, \left\{\boxed{1}, \boxed{2}, \boxed{3}, \boxed{4}\right\}, \\ \left\{\boxed{1}, \boxed{2}, \boxed{5}, \boxed{6}\right\}, \left\{\boxed{3}, \boxed{4}, \boxed{5}, \boxed{6}\right\}, \Omega \end{array} \right\}$$

Sigma-Algebra I

Definitions

Definition

The pair $(\Omega, \mathcal{F})$ made up of a sample space and a $\sigma-$algebra is called *measurable space*.

Definition

A family $\mathcal{P} = \{A_1, ..., A_n\}$ of events in Ω is called a *finite partition* of Ω if

$$
\begin{array}{ll}
(PF1) & \forall i \in \{1, ..., n\}\,, \; A_i \neq \varnothing, \\
(PF2) & \forall i, j \in \{1, ..., n\} \text{ such that } i \neq j, \; A_i \cap A_j = \varnothing, \\
(PF3) & \bigcup_{i=1}^{n} A_i = \Omega.
\end{array}
$$

Sigma-Algebra II

Definitions

Definition

A $\sigma-$algebra $\mathcal{F}$ is said to be generated from the finite partition $\mathcal{P}$ if it is the smallest $\sigma-$algebra that contains all the elements of $\mathcal{P}$. In that case $\mathcal{F}$ is denoted $\sigma\left(\mathcal{P}\right)$ and the elements of $\mathcal{P}$ are the atoms of $\mathcal{F}$.

Random variables

Definition

Definition

A random variable X constructed on the measurable space $(\Omega, \mathcal{F})$, is a real-valued function $X : \Omega \to \mathbb{R}$ such that

$$(*)\ \forall x \in \mathbb{R},\ \{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}$$

Random variables

Definition

Exercise. Show that, if $Card(\Omega) < \infty$, then the condition $(*)$ is equivalent to

$$(**)\ \forall x \in \mathbb{R},\ \{\omega \in \Omega \,|X(\omega) = x\} \in \mathcal{F}.$$

So, $X : \Omega \to \mathbb{R}$ is a random variable on the measurable space $(\Omega, \mathcal{F})$ if and only if the events that characterize its distribution are elements of $\mathcal{F}$. If $\mathcal{F}$ and $\mathcal{G}$ are two $\sigma-$algebras of Ω then it is possible that X is a random variable on the measurable space $(\Omega, \mathcal{F})$ but that it is not a random variable on the space $(\Omega, \mathcal{G})$. To clearly express the fact that the $\sigma-$algebra $\mathcal{F}$ contains the events characterizing the distribution of X, we say that X is $\mathcal{F}-$*measurable*.

Random variables I

Example

$$\begin{aligned}\Omega &= \left\{\boxed{1},\boxed{2},\boxed{3},\boxed{4},\boxed{5},\boxed{6}\right\} \\ \text{and } \mathcal{F} &= \left\{\varnothing,\left\{\boxed{1},\boxed{3},\boxed{5}\right\},\left\{\boxed{2},\boxed{4},\boxed{6}\right\},\Omega\right\}.\end{aligned}$$

The function $U : \Omega \to \mathbb{R}$ that returns 1 if the result is even and 0 otherwise is a random variable on $(\Omega, \mathcal{F})$ whereas the function $V : \Omega \to \mathbb{R}$ that returns 1 if the result is less than 4 and 0 otherwise is not. Indeed,

$$\{\omega \in \Omega \,|\, U(\omega) = x\} = \begin{cases} \left\{\boxed{1},\boxed{3},\boxed{5}\right\} \in \mathcal{F} & \text{if } x = 0 \\ \left\{\boxed{2},\boxed{4},\boxed{6}\right\} \in \mathcal{F} & \text{if } x = 1 \\ \varnothing \in \mathcal{F} & \text{otherwise} \end{cases}$$

and

$$\{\omega \in \Omega \,|\, V(\omega) = x\} = \begin{cases} \left\{\boxed{4},\boxed{5},\boxed{6}\right\} \notin \mathcal{F} & \text{if } x = 0 \\ \left\{\boxed{1},\boxed{2},\boxed{3}\right\} \notin \mathcal{F} & \text{if } x = 1 \\ \varnothing \in \mathcal{F} & \text{otherwise} \end{cases}.$$

U is said to be $\mathcal{F}-$measurable whereas V is not.

Random variables II

Example

By contrast, U and V are $\mathcal{G}-$measurable where

$$
\begin{aligned}
\mathcal{G} &= \sigma\left\{\left\{\boxed{2}\right\},\left\{\boxed{5}\right\},\left\{\boxed{1},\boxed{3}\right\},\left\{\boxed{4},\boxed{6}\right\}\right\} \\
&= \left\{\begin{array}{c}
\varnothing,\left\{\boxed{2}\right\},\left\{\boxed{5}\right\},\left\{\boxed{1},\boxed{3}\right\},\left\{\boxed{4},\boxed{6}\right\},\left\{\boxed{2},\boxed{5}\right\} \\
\left\{\boxed{1},\boxed{2},\boxed{3}\right\},\left\{\boxed{1},\boxed{3},\boxed{5}\right\},\left\{\boxed{2},\boxed{4},\boxed{6}\right\},\left\{\boxed{4},\boxed{5},\boxed{6}\right\}, \\
\left\{\boxed{1},\boxed{2},\boxed{3},\boxed{5},\right\},\left\{\boxed{1},\boxed{3},\boxed{4},\boxed{6}\right\},\left\{\boxed{2},\boxed{4},\boxed{5},\boxed{6}\right\} \\
\left\{\boxed{1},\boxed{3},\boxed{4},\boxed{5},\boxed{6}\right\},\left\{\boxed{1},\boxed{2},\boxed{3},\boxed{4},\boxed{6}\right\},\Omega
\end{array}\right\}
\end{aligned}
$$

The next results will enable us to identify the smallest $\sigma-$algebra that make one or several random variables measurable.

Random variables

Measurability

Theorem

Let $(\Omega, \mathcal{F})$, *Card* $(\Omega) < \infty$, *be a measurable space and* $\mathcal{P} = \{A_1, ..., A_n\}$, *be the finite partition of* Ω *that generates* $\mathcal{F}$. *The function* $X : \Omega \to \mathbb{R}$ *is a random variable on that space (*X *is* $\mathcal{F}-$*measurable) if and only if* X *is constant on the atoms of* $\mathcal{F}$.

Random variables I

Proof

Proof. Let's first verify that, *if* X *is constant on the atoms of* $\mathcal{F}$ *then* X *is a random variable on that space* $(\Omega, \mathcal{F})$.

If X is constant on the atoms of $\mathcal{F}$ then X may only take a finite number of values that we will denote $x_1, ..., x_m$. So, $\forall i \in \{1, ..., m\}$, the event $\{\omega \in \Omega \,|X(\omega) = x_i\}$ may be represented as a union of atoms of $\mathcal{F}$ and, since a $\sigma-$algebra is closed under finite unions, then $\{\omega \in \Omega \,|X(\omega) = x_i\} \in \mathcal{F}$. ■

Random variables II

Proof

Let's now verify that *if X is a random variable, then X is constant on the atoms of $\mathcal{F}$.*

We're going to use a proof by contradiction. Let's assume that there exists an atom A_k of $\mathcal{F}$ on which X is not constant. Then there exists at least two values for X on A_k. Let x_0, be one of those values. The event $A_0 = \{\omega \in \Omega \,|X(\omega) = x_0\}$ is an element of the $\sigma-$algebra because, by hypothesis, X is a random variable. As a result, A_0 is a strict subset of A_k that belongs to $\mathcal{F}$, which contradicts the fact that A_k is an atom of $\mathcal{F}$. ■

The proof of the theorem is complete. ■

Random variables

Corollaries

Corollary

Any random variable on a measurable space equipped with the trivial $\sigma-$algebra is constant.

Corollary

If $\mathcal{F} =$ the set of all possible events in Ω then any real-valued function on Ω $(X : \Omega \to \mathbb{R})$ is $\mathcal{F}-$measurable, that is to say that is is a random variable on $(\Omega, \mathcal{F})$.

Random variables

Generated sigma-algebra

Definition

Let $X : \Omega \to \mathbb{R}$. The smallest $\sigma-$algebra $\mathcal{F}$ that make X a random variable on the measurable space $(\Omega, \mathcal{F})$ is called the $\sigma-$algebra generated *by* X and is denoted $\sigma(X)$.

If $Card(\Omega) < \infty$ then X can only take a finite number of values, let's say $x_1, \ldots, x_m$. For any $i \in \{1, \ldots, m\}$, let's define $A_i = \{\omega \in \Omega \,|X(\omega) = x_i\}$. Then $\mathcal{P} = \{A_1, \ldots, A_m\}$ is a finite partition of Ω and $\sigma(X) = \sigma(\mathcal{P})$.

Random variables

Transformations

Theorem

Let's assume that $Card(\Omega) < \infty$*. If* $X : \Omega \to \mathbb{R}$ *and* $Y : \Omega \to \mathbb{R}$ *are* $\mathcal{F}-$*measurable, then* $\forall a, b \in \mathbb{R}$, $aX + bY$ *is also* $\mathcal{F}-$*measurable, which is to say that any linear combination of random variables built on the same measurable space is a random variable of that space.*

Random variables

Transformations

Proof. Since $Card(\Omega) < \infty$, the random variables X and Y can only take a finite number of values, let's say $x_1 < \ldots < x_m$ and $y_1 < \ldots < y_n$ respectively. $\forall z \in \mathbb{R}$,

$$\begin{aligned} & \{\omega \in \Omega \,|\, aX(\omega) + bY(\omega) \leq z\} \\ = & \bigcup_{ax_i + by_j \leq z} \{\omega \in \Omega \,|\, X(\omega) = x_i \text{ and } Y(\omega) = y_j\} \\ = & \bigcup_{ax_i + by_j \leq z} \underbrace{\underbrace{\{\omega \in \Omega \,|\, X(\omega) = x_i\}}_{\in \mathcal{F}} \cap \underbrace{\{\omega \in \Omega \,|\, Y(\omega) = y_j\}}_{\in \mathcal{F}}}_{\in \mathcal{F}} \in \mathcal{F}. \ \blacksquare \end{aligned}$$

Probability measures

Definitions

Definition

Let $(\Omega, \mathcal{F})$ be a measurable space. The function $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$ is a *probability measure* on $(\Omega, \mathcal{F})$ if

P1 $\mathbb{P}(\Omega) = 1$.

P2 $\forall A \in \mathcal{F}$, $0 \leq \mathbb{P}(A) \leq 1$.

P3 $\forall A_1, A_2, \ldots \in \mathcal{F}$ such that $A_i \cap A_j = \varnothing$ if $i \neq j$, $\mathbb{P}\left(\bigcup_{i \geq 1} A_i\right) = \sum_{i \geq 1} \mathbb{P}(A_i)$.

Definition

The triple $(\Omega, \mathcal{F}, \mathbb{P})$ made up of a sample space, a $\sigma-$algebra on Ω and a probability measure built on $(\Omega, \mathcal{F})$ is called *probability space*.

References

1. BILLINGSLEY, Patrick (1986). *Probability and Measure,* Second Edition, Wiley, New York.
    - That book is meant for people with a strong background in mathematics. All important results in probability theory can be found in that book with their proofs. The book offers a wide selection of exercises, at all difficulty levels. That book is well above the level set for this course.
2. ROSS, Sheldon, M. (1997). *Introduction to probability Models*, sixth edition, Academic Press, New York.

Probability measures I

Appendix I

Theorem

When Card $(\Omega) < \infty$, *let's say* $\Omega = \{\omega_1, ..., \omega_n\}$ *then the three conditions* $(P1)$, $(P2)$ *and* $(P3)$ *of the partial definition of a probability measure are equivalent to the following three conditions:*

Definition

P1* $\forall i \in \{1, ..., n\}$, $\mathbb{P}(\omega_i) \geq 0$.

P2* For any event A in Ω, $\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega)$.

P3* $\sum_{i=1}^{n} \mathbb{P}(\omega_i) = 1$.

Probability measures II

Appendix I

Proof. Let's first assume that $\mathbb{P}$ satisfies the three conditions $(P1)$, $(P2)$ and $(P3)$ of the definition of a probability measure and let's show that, then, $\mathbb{P}$ also satisfies the conditions $(P1^*)$, $(P2^*)$ and $(P3^*)$.

- $(P1^*)$ Since $\forall i \in \{1, ..., n\}$, $\{\omega_i\}$ is an event in Ω then, from condition $(P2)$, $0 \leq \mathbb{P}(\omega_i) \leq 1$.
- $(P2^*)$ Since $\{\omega_1\}, ..., \{\omega_n\}$ are mutually disjoint events, then, using condition $(P3)$, we obtain that, for any event A in Ω, $\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega)$.
- $(P3^*)$ Since Ω is itself an event, by using the equality that we have just established and property $(P1)$, we have that $\sum_{i=1}^{n} \mathbb{P}(\omega) = \mathbb{P}(\Omega) = 1$.

Probability measures III

Appendix I

Let's now assume that $\mathbb{P}$ satisfies conditions $(P1^*)$, $(P2^*)$ and $(P3^*)$ and let's show that, then, $\mathbb{P}$ also satisfies the three conditions of the partial definition of a probability measure.

- $(P1)$ Using successively $(P2^*))$ and $(P3^*)$, $\mathbb{P}(\Omega) = \sum_{i=1}^{n} \mathbb{P}(\omega_i) = 1$.
- $(P2)$ Since condition $(P1^*)$ implies that $\forall \omega \in \Omega$, $\mathbb{P}(\omega) \geq 0$, then for any event A in Ω, $\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega) \geq 0$ where the equality is warranted by $(P2^*)$.
  On the other hand, by using successively $(P2^*)$, the inequality above and $(P3^*)$, $\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega) \leq \sum_{\omega \in A} \mathbb{P}(\omega) + \underbrace{\sum_{\omega \in A^c} \mathbb{P}(\omega)}_{\geq 0} = \sum_{\omega \in \Omega} \mathbb{P}(\omega) = 1$.

Probability measures IV

Appendix I

- $(P3)$ For all events A_1, A_2, ... mutually disjoint,

$$\begin{aligned}
\mathbb{P}\left(\bigcup_{i\geq 1} A_i\right) &= \mathbb{P}\left(\bigcup_{\omega\in\bigcup_{i\geq 1} A_i} \omega\right) \\
&= \sum_{\omega\in\bigcup_{i\geq 1} A_i} \mathbb{P}(\omega) \\
&= \sum_{i\geq 1}\sum_{\omega\in A_i} \mathbb{P}(\omega) \\
&= \sum_{i\geq 1} \mathbb{P}(A_i) . \blacksquare
\end{aligned}$$

Probability measures I
Appendix II

Theorem

In the case where Card $(\Omega) < \infty$, *the distribution of a random variable is also characterized by its probability mass function*

$$f_X \quad : \quad \mathbb{R} \longrightarrow [0,1]$$
$$x \longrightarrow \textit{probability that the r.v. } X \textit{ is equal to } x,$$

which is to say that

$$\forall x \in \mathbb{R},\ f_X(x) = \mathbb{P}\{\omega \in \Omega \,|\, X(\omega) = x\}.$$

Probability measures II

Appendix II

Proof. We must prove that

- (i) Given a probability mass function f_X, there exists one and only one cumulative distribution function,
- (ii) for any given cumulative distribution function F_X, there exists one and only one probability mass function.

Since $Card\,(\Omega) < \infty$ then the random variable X can only take a finite number of values, let's say $x_1 < ... < x_n$.

Probability measures III

Appendix II

Proof of point (i). Let f_X be the probability mass function of the random variable X. Then, $\forall x \in \mathbb{R}$,

$$
\begin{aligned}
F_X(x) &= \mathbb{P}\{\omega \in \Omega \,|\, X(\omega) \leq x\} \\
&= \mathbb{P}\left[\bigcup_{x_i \leq x} \{\omega \in \Omega \,|\, X(\omega) = x_i\}\right] \\
&= \sum_{x_i \leq x} \mathbb{P}\{\omega \in \Omega \,|\, X(\omega) = x_i\}
\end{aligned}
$$

because the events in question are disjoint.

$$
= \sum_{x_i \leq x} f_X(x_i).
$$

Since, by construction, the probability mass function is unique, then, the cumulative distribution function is also unique.

Probability measures IV

Appendix II

Proof of point (ii). Let F_X, be the cumulative distribution function of the random variable X. Then, $\forall i \in \{2, ..., n\}$,

$$\begin{aligned} & F_X(x_i) \\ = \ & \mathbb{P}\{\omega \in \Omega \,|\, X(\omega) \leq x_i\} \\ = \ & \mathbb{P}[\{\omega \in \Omega \,|\, X(\omega) = x_i\} \cup \{\omega \in \Omega \,|\, X(\omega) \leq x_{i-1}\}] \\ = \ & \mathbb{P}\{\omega \in \Omega \,|\, X(\omega) = x_i\} + \mathbb{P}\{\omega \in \Omega \,|\, X(\omega) \leq x_{i-1}\} \\ & \text{because both events in question are disjoint.} \\ = \ & f_X(x_i) + F_X(x_{i-1}). \end{aligned}$$

Which implies that $\forall i \in \{2, ..., n\}$,

$$f_X(x_i) = F_X(x_i) - F_X(x_{i-1}).$$

Probability measures V

Appendix II

What happens to x_1? Since x_1 is the smallest possible value for the random variable X,

$$\begin{aligned} f_X(x_1) &= \mathbb{P}\{\omega \in \Omega \,|\, X(\omega) = x_1\} \\ &= \mathbb{P}\{\omega \in \Omega \,|\, X(\omega) \leq x_1\} \\ &= F_X(x_1). \end{aligned}$$

Now, $\forall x \in \mathbb{R}$, $x \notin \{x_1, ..., x_n\}$,

$$f_X(x) = \mathbb{P}\{\omega \in \Omega \,|\, X(\omega) = x\} = \mathbb{P}\{\varnothing\} = 0.$$

Since, by construction, the cumulative distribution function is unique, then, the probability mass function is also unique. ■
That last argument completes the proof of the proposition. ■