# Stochastic processes and overview of the course

Stochastic processes are everywhere! Whenever there is something evolving in time, with a probabilistic outcome, we have a stochastic process.

# Stochastic processes and overview of the course

Stochastic processes are everywhere! Whenever there is something evolving in time, with a probabilistic outcome, we have a stochastic process.

weather in different days

Stochastic processes are everywhere! Whenever there is something evolving in time, with a probabilistic outcome, we have a stochastic process.

number of job requests at a server

# Stochastic processes and overview of the course

Stochastic processes are everywhere! Whenever there is something evolving in time, with a probabilistic outcome, we have a stochastic process.

whether a factory machine is working or not

# Stochastic processes and overview of the course

Stochastic processes are everywhere! Whenever there is something evolving in time, with a probabilistic outcome, we have a stochastic process.

position of a pollen grain in a glass of water

# Stochastic processes and overview of the course

**Definition**

A stochastic process is a collection of random variables $\{X_i\}_{i \in I}$ for some set of ordered indices $I$, with the random variables sharing a common *state space* $S$.

- If $S$ and $I$ are discrete sets, we have a *discrete stochastic process with discrete time* (ex. weather day by day).
- If $S$ is discrete and $I$ is continuous, we have a *discrete stochastic process with continuous time* (ex. server requests, machine status).
- If $S$ and $I$ are continuous sets, we have a *continuous stochastic process with continuous time* (ex. pollen grain).

# Stochastic processes and overview of the course

## Definition

A stochastic process is a collection of random variables $\{X_i\}_{i \in I}$ for some set of ordered indices $I$, with the random variables sharing a common *state space* $S$.

- If $S$ and $I$ are discrete sets, we have a *discrete stochastic process with discrete time* (ex. weather day by day).
- If $S$ is discrete and $I$ is continuous, we have a *discrete stochastic process with continuous time* (ex. server requests, machine status).
- If $S$ and $I$ are continuous sets, we have a *continuous stochastic process with continuous time* (ex. pollen grain).

## Definition

A stochastic process is a collection of random variables $\{X_i\}_{i \in I}$ for some set of ordered indices $I$, with the random variables sharing a common *state space* $S$.

- If $S$ and $I$ are discrete sets, we have a *discrete stochastic process with discrete time* (ex. weather day by day).
- If $S$ is discrete and $I$ is continuous, we have a *discrete stochastic process with continuous time* (ex. server requests, machine status).
- If $S$ and $I$ are continuous sets, we have a *continuous stochastic process with continuous time* (ex. pollen grain).

# Stochastic processes and overview of the course

**Definition**

A stochastic process is a collection of random variables $\{X_i\}_{i \in I}$ for some set of ordered indices $I$, with the random variables sharing a common *state space* $S$.

- If $S$ and $I$ are discrete sets, we have a *discrete stochastic process with discrete time* (ex. weather day by day).
- If $S$ is discrete and $I$ is continuous, we have a *discrete stochastic process with continuous time* (ex. server requests, machine status).
- If $S$ and $I$ are continuous sets, we have a *continuous stochastic process with continuous time* (ex. pollen grain).

# Quiz time!

Which of the following ones are stochastic processes?

1. $\{X_t\}_{t\in[0,\infty)}$ with $X_t$ being the number of customers in a shop at time $t$;
2. $\{X_t\}_{t\in[0,\infty)}$ with $X_t$ being the status of your computer ("on", "off", or "broken") at time $t$;
3. $\{X_t\}_{t\in[0,\infty)}$ with $X_t$ being the position of a frog jumping around at time $t$.

# Quiz time!

Consider a sequence of i.i.d. random variables $\{X_i\}_{i=1}^{\infty}$. Is it a stochastic process?

Things become interesting if there is some “structure”

Things become interesting if there is some “structure”

**Preview 1** What happens in different interval time intervals is independent. For example the factory machine is as new whenever it is repaired.

Things become interesting if there is some “structure”

**Preview 2** What happens in the future is independent of the past given the current state.
For example tomorrow’s weather depends only on the current weather situation and not the past weather.

# What you should know

Things that you are welcome to ask me about during office hours, but that I will consider as known during classes:

- basic concepts of linear algebra;
- basic concepts of analysis: convergent and divergent series, how to calculate multi-dimensional integrals, monotone convergence theorem, Fubini and Tonelli theorems, Lebesgue's dominated convergence theorem, Fatou's lemma;
- basic concepts of probability: discrete and continuous random variables, $\sigma-$algebras, mean, variance, covariance;
- convergence of random variables: weak convergence, convergence in probability, strong convergence, convergence in mean, weak and strong law of large numbers, central limit theorem;
- dependence structure: conditional distribution, conditional expectation given a $\sigma-$algebra (which is a random variable) and its properties.

# Random variables we will be working with

During classes I will assume the law of the following random variables is known:

- exponential (my favourite one!);
- gaussian (or normal, also very cool);
- uniform (both discrete and continuous);
- Bernoulli;
- binomial;
- geometric;
- Poisson.

# Esame

La verifica dell'apprendimento per questo modulo (processi stocastici) consiste nelle seguenti parti:

- un orale obbligatorio;
- un'attività facoltativa.

# Esame: orale obbligatorio

L'esame orale durerà all'incirca mezz'ora, il candidato potrà presentarsi direttamente all'orario concordato per il suo esame.

L'esame orale verterà **principalmente** su un argomento (una lista completa verrà fornita a fine corso) pescato a sorte dalla commissione. L'argomento viene comunicato al candidato 30 minuti prima del suo esame. Durante questi trenta minuti, il candidato può consultare liberamente il materiale del corso.

# Esame: orale obbligatorio

Alla fine dell'esame orale verrà assegnato un punteggio massimo di 31 (corrispondente a 30 e lode). Verranno valutati:

Conoscenza della teoria : conoscenza delle definizioni, dei teoremi, delle dimostrazioni viste a lezioni (alcune saranno escluse);

Comprensione della teoria : comprensione del significato e della rilevanza dei teoremi (cosa permettono di concludere? perché ce n'era bisogno?) capacità di risolvere esercizi pratici e teorici non visti a lezione;

Capacità espositiva : Capacità dello studente di esporre i concetti in maniera chiara.

# Esame: attività facoltativa

L'attività facoltativa consiste in:

Risposte settimanali ai quiz : sulla pagina moodle del corso appariranno dei quiz settimanali (con scadenza di **due** settimane) ai quali gli studenti risponderanno da remoto, in autonomia. A settimane alterne, questi quiz saranno preparati dal docente e dagli studenti stessi, secondo le modalità descritte sotto;

Formulazione delle domande : ogni due settimane, gli studenti lavoreranno in classe in gruppo per formulare delle domande a risposta multipla, con una sola risposta corretta e 5 possibili opzioni. Mezz'ora prima della fine della lezione, il docente analizzerà la chiarezza e la correttezza delle domande, e richiederà evenutali modifiche. Le domande saranno poi caricate su moodle dal docente e assegnate casualmente la settimana successiva.

# Esame: attività facoltativa

L'attività facoltativa consiste in:

Risposte settimanali ai quiz : **Ogni settimana**, il martedì dopo lezione, sulla pagina moodle del corso appariranno dei quiz (**con scadenza di due settimane**) ai quali gli studenti risponderanno da remoto, in autonomia. A settimane alterne, questi quiz saranno preparati dal docente e dagli studenti stessi, secondo le modalità descritte sotto;

Formulazione delle domande : ogni due settimane, gli studenti lavoreranno in classe in gruppo per formulare delle domande a risposta multipla, con una sola risposta corretta e 5 possibili opzioni. Mezz'ora prima della fine della lezione, il docente analizzerà la chiarezza e la correttezza delle domande, e richiederà evenutali modifiche. Le domande saranno poi caricate su moodle dal docente e assegnate casualmente la settimana successiva.

# Esame: attività facoltativa

| Sett. | Lezioni | | | Quiz su Moodle da |
|---|---|---|---|---|
| 1 | 22/09 | Sciopero | (da recuperare) | Docente |
| | 23/09 | 8:30-11:30 | Lezione | |
| 2 | 29/09 | 14:30-16:00 | Lezione | Docente |
| | 30/09 | 8:30-10:00 | Lezione | |
| | 30/09 | 10:00-11:30 | Attività di gruppo | |
| 3 | 06/10 | 14:30-16:00 | Lezione | Studenti! |
| | 07/10 | 8:30-11:30 | Lezione | |
| 4 | 13/10 | 14:30-16:00 | Lezione | Docente |
| | 14/10 | 8:30-10:00 | Lezione | |
| | 14/10 | 10:00-11:30 | Attività di gruppo | |

# Esame: attività facoltativa

| Sett. | Lezioni | | | Quiz su Moodle da |
|---|---|---|---|---|
| 5 | 20/10 | 14:30-16:00 | Lezione | Studenti! |
| | 21/10 | 8:30-11:30 | Lezione | |
| 6 | 27/10 | 14:30-16:00 | Lezione | Docente |
| | 28/10 | 8:30-10:00 | Lezione | |
| | 28/10 | 10:00-11:30 | Attività di gruppo | |
| 7 | 03/11 | 14:30-16:00 | Lezione | Studenti! |
| | 04/11 | 8:30-11:30 | Lezione | |
| 8 | 10/11 | 14:30-16:00 | Lezione | Docente |
| | 11/11 | 8:30-10:00 | Lezione | |
| | 11/11 | 10:00-11:30 | Attività di gruppo | |

# Esame: attività facoltativa

| Sett. | Lezioni | | | Quiz su Moodle da |
|---|---|---|---|---|
| 9 | 17/11 | 14:30-16:00 | Lezione | Studenti! |
| | 18/11 | 8:30-11:30 | Lezione | |
| 10 | 24/11 | 14:30-16:00 | Lezione | Docente |
| | 25/11 | 8:30-10:00 | Lezione | |
| | 25/11 | 10:00-11:30 | Attività di gruppo | |
| 11 | 01/12 | 14:30-16:00 | Lezione | Studenti! |
| | 02/12 | 8:30-11:30 | Lezione | |
| 12 | 08/12 | Festività | | Docente |
| | 09/12 | 8:30-10:00 | Lezione | |
| | 09/12 | 10:00-11:30 | Attività di gruppo | |

# Esame: attività facoltativa

| Sett. | Lezioni | | | Quiz su Moodle da |
|---|---|---|---|---|
| 13 | 15/12 | 14:30-16:00 | Lezione | Studenti! |
| | 16/12 | 8:30-11:30 | Lezione | |
| - | Vacanze di Natale | | | Docente (scadenza posticipata) |

# Esame: attività facoltativa

Punteggio: ogni settimana i quiz su Moodle avranno un punteggio massimo di 2 punti. Il punteggio finale di ogni singolo studente sarà calcolato così:

$$\sum_{i=1}^{14} P_i + \sum_{j=1}^{6} 5 * I_j * V_j$$

dove:

$0 \leq P_i \leq 2$ è il punteggio ottenuto con il quiz dell'$i-$esima settimana

$$I_j = \begin{cases} 0 & \text{se la } j-\text{esima domanda formulata non è chiara o corretta} \\ 1 & \text{altrimenti} \end{cases}$$

$0 \leq V_j \leq 1$ è la proporzione di studenti che hanno sbagliato la $j-$esima domanda formulata

I primi tre classificati vinceranno un gadget **bellissimo** scelto dal docente, oltre che gloria imperitura!

Per ottenere un punto extra all'esame invece, basta partecipare ad almeno 4 attività di gruppo e rispondere (anche in maniera errata) ad almeno il 90% dei quiz su moodle.

# Voti parziali

Questo corso è un modulo di un corso più grande, per un totale di 12 crediti. Il voto finale sarà calcolato come media dei due voti parziali.

- Manterrò i voti parziali fino a che posso (cioè fino a che il corso non cambia, e fino a che io sarò il titolare);
- Se si rifiuta il voto ad un appello, mi aspetto che venga ritentato l'esame di **almeno uno** dei due moduli;
- Il punto extra dell'attività facoltativa, se ottenuto, **non scade**.

# IMPORTANTE

Ciò che mi rende felice
è vedervi sbagliare

# Forum

Il forum è sulla pagina Moodle del corso, usatelo!

- Vi permette di avere tutte le domande e i dubbi raccolti in un'uncio posto, comodo per studiare!
- Vi permette di sviluppare dubbi! LEggendo una domanda potreste capire di non aver capito qualcosa.
- Fa partire discussioni fra voi studenti, spesso più efficaci dell'interazione docente-studente
- Alla fine controllo sempre che ci sia una risposta corretta.

# Ore di Q&A

Durante il corso, individueremo **due** slot da 1,5 ore ciascuno in cui risponderò ai vostri dubbi sul materiale e sugli esercizi.

Sceglieremo degli orari i più compatibili con i vostri impegni. Vi darò più dettagli più avanti.

# Nuovo modello didattico

Questo è un corso sperimentale, in cui in vista dell'AA 2027/28 1 CFU corrisponde a:

- 25 ore di lavoro (come sempre);
- 8 ore di lezione frontale (invece di 10);
- 2 ore di impegno individuale assistito (attività di gruppo, sessioni di Q&A);
- 15 ore di studio individuale.

# Materiale

Le lezioni frontali (NON le attività di gruppo) e le sessioni di Q&A saranno offerte anche in streaming e registrate.

Le slide saranno fornite sia **prima** che **dopo** la lezione (quindi con i gli scarabocchi del docente sopra).

Cerco di avere delle slide chiare, quindi slide + videoregistrazioni devono essere il vostro materiale di riferimento su cui studiare per l'esame.

# Libri di testo

Non è richiesto leggere nessun libro di testo in particolare. I seguenti libri sono libri che conosco e che a me piacciono, il consiglio per approfondimenti personali o per cercare di capire meglio qualcosa non troppo chiaro nelle mie spiegazioni (nel qual caso, comunque, chiedete!!!)

- R. Durrett, ”Essentials of Stochastic Processes”, Springer, 2021 (online available).
- D. A. Levin, Y. Peres, E. L. Wilmer, ”Markov chains and mixing times”, AMS, 2008 (online available).
- S.M. Ross, ”Introduction to Probability Models”, Elsevier, 2010.
- J.M. Steele ”Stochastic Calculus and Financial Applications”, Springer, 2000.

# Survey iniziale

Per favore compilate entro domani sera il survey su Moodle, così che possa avere un'idea della composizione studentesca e possa iniziare a formare i gruppi.

# Some technical definition

Recall what a $\sigma$-algebra is...

# Some technical definition

Why do we need $\sigma$-algebras?

# Some technical definition

Recall what random variables are...

# Some technical definition

A stochastic process is a random variable with infinite dimension: $\{X_i\}_{i \in I}$ is a random variable with image in

$$F = \{f : I \to S\}$$

# Some technical definition

## Definition

Let $\mathcal{F}$ be a σ-algebra. A *filtration* is a sequence $\{\mathcal{F}_i\}_{i\in I}$ of σ-algebras such that

1. $\mathcal{F}_i \subseteq \mathcal{F}$ for all $i \in I$;
2. $\mathcal{F}_i \subseteq \mathcal{F}_j$ for all $i, j \in I$ with $i < j$.

## Definition

A stochastic process $\{X_i\}_{i\in I}$ is *adapted* to a filtration $\{\mathcal{F}_i\}_{i\in I}$ if $X_i$ is $\mathcal{F}_i$-measurable for all $i \in I$.

In practice, $\mathcal{F}_i$ contains all the possible evolutions of the process up to time $i$ (it can contain more info).

# Some technical definition

## Definition

Let $\mathcal{F}$ be a $\sigma$-algebra. A *filtration* is a sequence $\{\mathcal{F}_i\}_{i\in I}$ of $\sigma$-algebras such that

1. $\mathcal{F}_i \subseteq \mathcal{F}$ for all $i \in I$;
2. $\mathcal{F}_i \subseteq \mathcal{F}_j$ for all $i, j \in I$ with $i < j$.

## Definition

A stochastic process $\{X_i\}_{i\in I}$ is *adapted* to a filtration $\{\mathcal{F}_i\}_{i\in I}$ if $X_i$ is $\mathcal{F}_i$-measurable for all $i \in I$.

In practice, $\mathcal{F}_i$ contains all the possible evolutions of the process up to time $i$ (it can contain more info).

# Example

For $i \in \{1, 2, 3, \dots\}$ let $Y_i \sim \text{Uniform}(0, 1)$, and

$$X_i = \begin{cases} 0 & \text{if } Y_i < 0.5 \\ 1 & \text{if } Y_i \geq 0.5 \end{cases}$$

Let $\mathcal{F}_i = \sigma\{Y_j : j \leq i\}$. Is $\{X_i\}_{i=1}^{\infty}$ adapted to $\{\mathcal{F}_i\}_{i=1}^{\infty}$?

# Some technical definition

## Definition

Let $\{X_i\}_{i \in I}$ be a stochastic process. Its *natural filtration* is given by

$$\mathcal{F}_i = \sigma\{X_j : j \leq i\}.$$

By definition, a stochastic process is always adapted to its natural filtration.

# Counting processes

If we count how many times something happened up to a certain time, then we have a counting process!
For example:

- $X_t$ is the number of waves hitting a rock by time $t$;
- $X_t$ is the number of customers entering a shop by time $t$;
- $X_t$ is the number of times a machine has been repaired by time $t$;

# Counting processes

Formal definition:

**Definition**

A counting process is a discrete stochastic process with state space $\{0, 1, 2, \dots\}$ that can only increase by one and can never decrease.

# Strong law of large numbers

**Theorem (Strong law of large numers)**

*Let $\{X_i\}_{i=1}^{\infty}$ be a collection of i.i.d. random variables for which $E[X_1]$ exists (it could be infinite). Then, with probability 1 we have*

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} X_i}{n} = E[X_1]$$

Extremely useful! For example in statistics to estimate the proportion of people voting for a given candidate we can take the average of $n$ trials.

# Strong law of large numbers

**Theorem (Strong law of large numers)**

*Let $\{X_i\}_{i=1}^{\infty}$ be a collection of i.i.d. random variables for which $E[X_1]$ exists (it could be infinite). Then, with probability 1 we have*

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} X_i}{n} = E[X_1]$$

Extremely useful! For example in statistics to estimate the proportion of people voting for a given candidate we can take the average of $n$ trials.

# Strong law of large numbers and stochastic processes

Independence is important to derive the law of large numbers, but a sequence of independent random variables is a boring stochastic process...

Sweet spot! We can consider independent inter-arrival times in a counting process!

# Strong law of large numbers and stochastic processes

Independence is important to derive the law of large numbers, but a sequence of independent random variables is a boring stochastic process...
Sweet spot! We can consider independent inter-arrival times in a counting process!

# Renewal processes

## Definition

A *renewal process* is a counting process such that the times in between consecutive arrivals are a sequence of i.i.d. random variables.

Formally, the arrival times are defined by $T_0 = 0$ and:

$$T_\ell = \inf\{i \in I : X_i = \ell\} \quad \text{for } i = 1, 2, 3, \dots$$

The inter-arrival times are defined by:

$$t_{\ell+1} = T_{\ell+1} - T_\ell$$

for all $\ell \in \{1, 2, 3, \dots\}$.

## Theorem

*Let $\{X_i\}_{i\in I}$ be a renewal process with inter-arrival times $\{t_i\}_{i=1}^{\infty}$, such that $0 < E[t_1]$ and $P(t_1 < \infty) = 1$. Then, with probability 1*

$$\lim_{i\to\infty} \frac{X_i}{i} = \frac{1}{E[t_1]}.$$

Why?

# Renewal processes

### Theorem

*Let* $\{X_i\}_{i \in I}$ *be a renewal process with inter-arrival times* $\{t_i\}_{i=1}^{\infty}$*, such that* $0 < E[t_1]$ *and* $P(t_1 < \infty) = 1$*. Then, with probability 1*

$$\lim_{i \to \infty} \frac{X_i}{i} = \frac{1}{E[t_1]}.$$

Why?

# Application

Assume that the times between requests to a server are i.i.d. with mean $\mu = 5$ seconds. After a day, how many requests per second have been received?

# Renewal-reward processes

To extend the possible applications, we assume that in each inter-arrival interval $[T_\ell, T_{\ell+1})$ a reward $r_{\ell+1}$ is given.

**Example:** if the factory machine does not break for at least one month, the tecnician receives a bonus of 5 euros.

In the long run, how much does the technician gain in bonuses per week?

# Renewal-reward processes

To extend the possible applications, we assume that in each inter-arrival interval $[T_\ell, T_{\ell+1})$ a reward $r_{\ell+1}$ is given.
**Example:** if the factory machine does not break for at least one month, the tecnician receives a bonus of 5 euros.

In the long run, how much does the technician gain in bonuses per week?

To extend the possible applications, we assume that in each inter-arrival interval $[T_\ell, T_{\ell+1})$ a reward $r_{\ell+1}$ is given.
**Example:** if the factory machine does not break for at least one month, the tecnician receives a bonus of 5 euros.

In the long run, how much does the technician gain in bonuses per week?

# Renewal-reward processes

Rewards can be:

- discrete (ex. fixed amounts of 5 euros);
- continuous (ex. the time the machine works);
- negative (ex. the repair may have a cost of 10 euros).

# Renewal-reward processes

Rewards can be:

- discrete (ex. fixed amounts of 5 euros);
- continuous (ex. the time the machine works);
- negative (ex. the repair may have a cost of 10 euros).

# Renewal-reward processes

Rewards can be:

- discrete (ex. fixed amounts of 5 euros);
- continuous (ex. the time the machine works);
- negative (ex. the repair may have a cost of 10 euros).

# Renewal-reward processes

Rewards can be:

- discrete (ex. fixed amounts of 5 euros);
- continuous (ex. the time the machine works);
- negative (ex. the repair may have a cost of 10 euros).

# Renewal-reward processes

## Definition

Let $\{X_i\}_{i\in I}$ be a renewal process with inter-arrival times $\{t_\ell\}_{\ell=1}^\infty$ and rewards $\{r_\ell\}_{\ell=1}^\infty$. If the random variables $\{(t_\ell, r_\ell)\}_{\ell=1}^\infty$ are i.i.d. then we have a *renewal-reward process*.

## Theorem

*Let $\{X_i\}_{i\in I}$ be a renewal process with $E[t_1] > 0$, and $P(t_1 < \infty) = 1$. Then, with probability 1,*

$$\lim_{i\to\infty} \frac{R_i}{i} = \frac{E[r_1]}{E[t_1]},$$

*where*

$$R_i = \sum_{\ell=1}^{X_i} r_\ell$$

*is the cumulative reward up to time $i$.*

# Renewal-reward processes

## Definition

Let $\{X_i\}_{i \in I}$ be a renewal process with inter-arrival times $\{t_\ell\}_{\ell=1}^\infty$ and rewards $\{r_\ell\}_{\ell=1}^\infty$. If the random variables $\{(t_\ell, r_\ell)\}_{\ell=1}^\infty$ are i.i.d. then we have a *renewal-reward process*.

## Theorem

*Let $\{X_i\}_{i \in I}$ be a renewal process with $E[t_1] > 0$, and $P(t_1 < \infty) = 1$. Then, with probability 1,*

$$\lim_{i \to \infty} \frac{R_i}{i} = \frac{E[r_1]}{E[t_1]},$$

*where*

$$R_i = \sum_{\ell=1}^{X_i} r_\ell$$

*is the cumulative reward up to time $i$.*

# Application

Assume a machine create a revenue of 10 euros per day when it works, repairing it costs 50 euros of transportation plus a repairing cost of approximately 40 euros per day. The machine does not break down for times with mean 30 days and repairing it takes on average 5 days. Assume that when the machine is repaired, it can be considered as new.
After 10 years of usage, how much revenue per day has the machine produced?

# Application

# Application

In the previous question, how does the answer change if repair times are geometric random variables, and times until break-downs are Poisson?