Welcome to the Advanced Statistics!

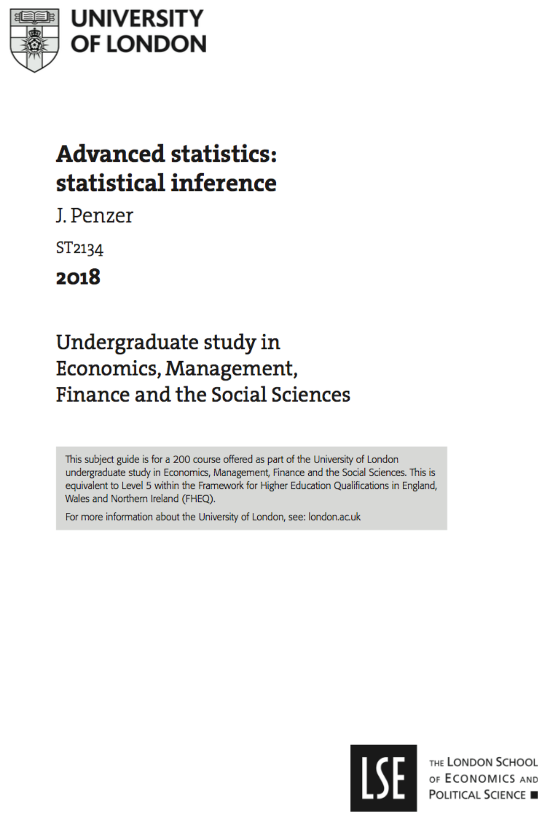
Peter Lukianchenko

2 September 2023

Course structure

Course Plan

- Stochastic Processes
- Time Series
- Advanced Statistics UoL



UNIVERSITY OF LONDON

Advanced statistics: statistical inference

J. Penzer

ST2134

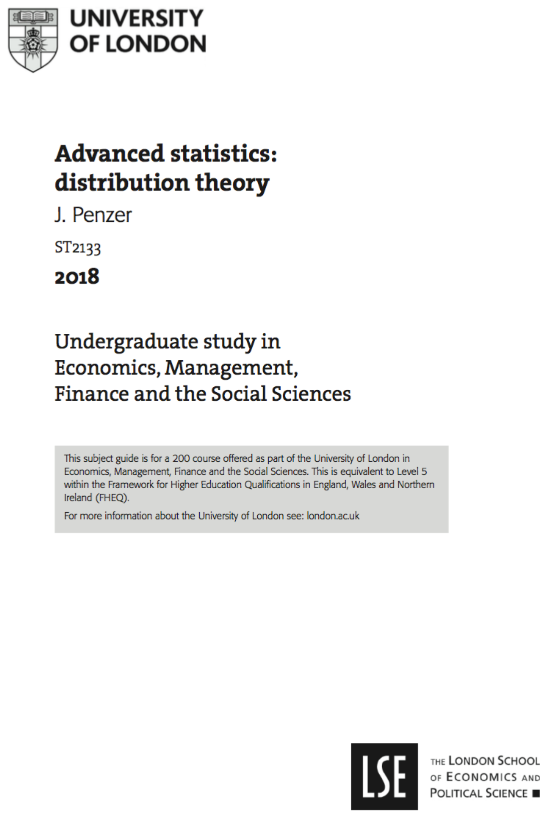
2018

Undergraduate study in Economics, Management, Finance and the Social Sciences

This subject guide is for a 200 course offered as part of the University of London undergraduate study in Economics, Management, Finance and the Social Sciences. This is equivalent to Level 5 within the Framework for Higher Education Qualifications in England, Wales and Northern Ireland (FHEQ).

For more information about the University of London, see: london.ac.uk

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE



UNIVERSITY OF LONDON

Advanced statistics: distribution theory

J. Penzer

ST2133

2018

Undergraduate study in Economics, Management, Finance and the Social Sciences

This subject guide is for a 200 course offered as part of the University of London in Economics, Management, Finance and the Social Sciences. This is equivalent to Level 5 within the Framework for Higher Education Qualifications in England, Wales and Northern Ireland (FHEQ).

For more information about the University of London see: london.ac.uk

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

Course structure

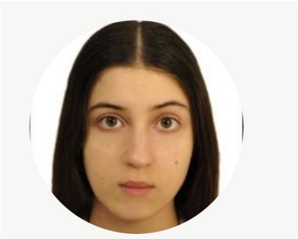
Formula for final grade – Stochastic Processes

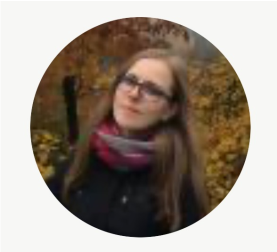
$$Final\ Score = 0.35 * \text{FALL} + 0.40 * \text{DEC} + 0.25 * \text{HW}(\text{m1} + \text{m2})$$

Formula for final grade – Time Series Analysis

$$Final\ Score = 0.30 * \text{Midterm Mimoza} + 0.45 * \text{Midterm Sakura} + 0.25 * \text{HW}(\text{m3} + \text{m4})$$

Course structure





Kirillova Maria

Popova Svetlana

Lecturer

Petr Lukianchenko

e-mail:
plukyanchenko@hse.ru or
lukianchenko.pierre@gmail.com

Class teacher

Boris Demeshev, Office S517

e-mail:
bdemeshev@hse.ru or boris.demeshev@gmail.com

Syllabus

Lecture 1
Recall basic of probability
MGF to start

Lecture 2
Markov chain

Lecture 3
Convergences

No memory

Memory $\exists$

Lecture 4
Conditional Expectations

Lecture 5
Poisson distribution

Lecture 6
G-algebra

Lecture 7
Filtration

How is about to recall ?

$N \quad \chi^2 \quad$ Uniform $\quad$ Poisson $\qquad \chi^2 = \sum_{i=1}^{n} X_i^2 , \; X_i \sim N(0;1)$

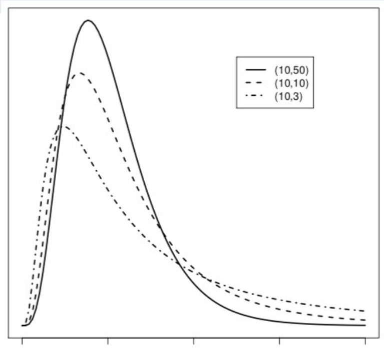
Definition

Let U and V be two independent random variables, where $U \sim \chi_p^2$ and $V \sim \chi_k^2$
Then the distribution of:

$$F = \frac{U/p}{V/k}$$

is the F distribution with degrees of freedom *(p, k)*, denoted $F \sim F_{p,k}$, or $F \sim F(p, k)$

$$t = \frac{N(0;1)}{\sqrt{\chi^2/n}}$$



Let’s recall rest of distributions

How is about to recall ?

The (random) interval (L, U) forms an interval estimator of θ:

For estimation to be as precise as possible, intuitively, the width of the interval, U − L, should be small. Then, typically, the coverage probability:

$$P(L(X_{1,} X_2, \ldots X_n)) < \theta < U(X_{1,} X_2, \ldots X_n)) < 0.95$$

Ideally, L and U should be chosen such that:

- the width of the interval is as small as possible;
- the coverage probability is as large as possible.

How is about to recall ?

Definition

Hence, supposing a 95% coverage probability:

$$
\begin{aligned}
0{,}95 &= P\left(\frac{\sqrt{n}|\bar{X}-\mu|}{\sigma} \leq 1{,}96\right) \\
&= P\left(|\bar{X}-\mu| \leq 1{,}96 * \frac{\sigma}{\sqrt{n}}\right) \\
&= P\left(|\bar{X}-\mu| \leq 1{,}96 * \frac{\sigma}{\sqrt{n}}\right) \\
&= P\left(-1{,}96 * \frac{\sigma}{\sqrt{n}} \leq \bar{X}-\mu \leq 1{,}96 * \frac{\sigma}{\sqrt{n}}\right) \\
&= P\left(\bar{X} - 1{,}96 * \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + 1{,}96 * \frac{\sigma}{\sqrt{n}}\right)
\end{aligned}
$$

How is about to recall ?

Definition

The likelihood function is defined as:

$$L(\theta) = \prod_{i=1}^{n} f(X_i, \theta)$$

- the likelihood function is the function of θ, while $X_{1,} X_2, \ldots X_n$ are treated as constants (as given observations);
- the likelihood function reflects the information about the unknown parameter θ in the data $X_{1,} X_2, \ldots X_n$.

Moment Generating Function

Definition

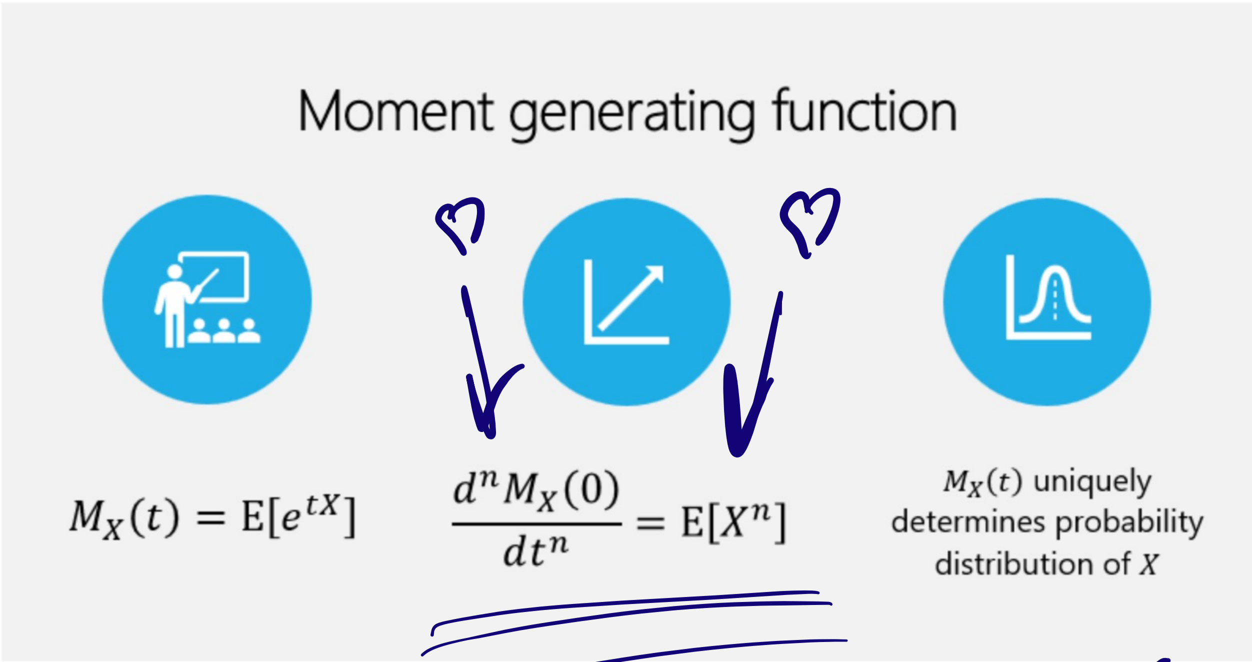
For many distributions, all the moments *E(X)*, $E(X^2)$ ... Can be encapsulated in a single function, which is called **moment generation function**. It exists for many commonly used distributions and often provides the most efficient way to calculate moments.

The moment generating function of a random variable X is a function $M_x : R \rightarrow [0; \infty)$ given by:

$$M_{X(t)} = E(e^{tX}) = \begin{cases} \sum_{x} e^{tX} f_X(x) & if\ X\ discrete, \\ \int_{-\infty}^{\infty} e^{tX} f_X(x) dx & if\ X\ continuous \end{cases}$$

Where, to be well-defined, we require some *h > 0* such that $M_{X(t)} < \infty$ for all $t \in [-h, h]$.

Moment Generating Function



$$(M_X)'_t\Big|_{t=0} = E(X) \qquad (M_X)''_{tt}\Big|_{t=0} = E(X^2)$$

Moment Generating Function

Example Let X be a continuous random variable with support

$$R_X = [0, \infty)$$

and probability density function

$$f_X(x) = \begin{cases} \lambda \exp(-\lambda x) & \text{if } x \in R_X \\ 0 & \text{if } x \notin R_X \end{cases}$$

$f = \lambda e^{-\lambda x}$

where λ is a strictly positive number. The expected value $\mathrm{E}[\exp(tX)]$ can be computed as follows:

$$\begin{aligned}
\mathrm{E}[\exp(tX)] &= \int_{-\infty}^{\infty} \exp(tx) f_X(x)\,dx \\
&= \int_0^{\infty} \exp(tx)\lambda \exp(-\lambda x)\,dx \\
&= \lambda \int_0^{\infty} \exp((t-\lambda)x)\,dx \qquad \text{(which is finite only if } t < \lambda\text{)} \\
&= \lambda \left[\frac{1}{t-\lambda} \exp((t-\lambda)x) \right]_0^{\infty} \\
&= \lambda \left[0 - \frac{1}{t-\lambda} \right] \\
&= \frac{\lambda}{\lambda - t}
\end{aligned}$$

$f = \lambda e^{-\lambda x}$

$M_x(t) = E e^{tx} = \int^{+\infty} e^{tx} \cdot \lambda e^{-\lambda x}\, dx$ $\quad \boxed{e^{-\infty} = 0}$

$= \lambda \int_0^{+\infty} e^{(t-\lambda)x}\, dx$

$t \neq \lambda$: $\lambda \cdot \frac{1}{t-\lambda} e^{(t-\lambda)x} \Big|_0^{+\infty} =$

$t = \lambda$: $= \lambda \cdot x \Big|_0^{+\infty} = +\infty$

$t > \lambda$: $= +\infty$

$t < \lambda$: $= 0 - \frac{\lambda}{t-\lambda} e^{0} = \frac{\lambda}{\lambda - t}$

Furthermore, the ~~above~~ expected value exists and is finite for any $t \in [-h, h]$, provided $0 < h < \lambda$. As a consequence, X possesses a mgf:

$$M_X(t) = \frac{\lambda}{\lambda - t}$$

$\frac{\partial}{\partial t}$

$EX = \frac{-\lambda}{(\lambda-t)^2} \cdot (-1) \Big|_{t=0} = \frac{\lambda}{\lambda^2} = \frac{1}{\lambda}$

$E(X^2) = \frac{\partial^2 M}{\partial t^2} = \frac{\partial}{\partial t} \frac{\lambda}{(\lambda-t)^2} = \frac{-2\cdot\lambda}{(\lambda-t)^3} \cdot (-1) \Big|_{t=0} = \frac{2}{\lambda^2}$

$\mathrm{Var}\, X = EX^2 - E^2X = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$

Moment Generating Function – Linear Transformation

Let X be a random variable possessing a mgf $M_X(t)$.

Define

$$Y = a + bX$$

M_X

where $a, b \in \mathbb{R}$ are two constants and $b \neq 0$.

Then, the random variable Y possesses a mgf $M_Y(t)$ and

$$M_Y(t) = \exp(at) M_X(bt)$$

Proof

By the very definition of mgf, we have

$y = a + bX$ $\quad V_y = b^2 \cdot V_X$

$$\begin{aligned} M_Y(t) &= \mathrm{E}[\exp(tY)] \\ &= \mathrm{E}[\exp(at + btX)] \\ &= \mathrm{E}[\exp(at)\exp(btX)] \\ &= \exp(at)\mathrm{E}[\exp(btX)] \\ &= \exp(at) M_X(bt) \end{aligned}$$

Obviously, if $M_X(t)$ is defined on a closed interval $[-h, h]$, then $M_Y(t)$ is defined on the interval $\left[-\frac{h}{b}, \frac{h}{b}\right]$.

Moment Generating Function

Let X_1, ..., X_n be n mutually independent random variables.

Let Z be their sum:

$$Z = \sum_{i=1}^{n} X_i$$

Then, the mgf of Z is the product of the mgfs of X_1, ..., X_n:

$$M_Z(t) = \prod_{i=1}^{n} M_{X_i}(t)$$

Proof

This is easily proved by using the definition of mgf and the properties of mutually independent variables:

$$\begin{aligned}
M_Z(t) &= \mathrm{E}[\exp(tZ)] \\
&= \mathrm{E}\left[\exp\left(t\sum_{i=1}^{n} X_i\right)\right] \\
&= \mathrm{E}\left[\exp\left(\sum_{i=1}^{n} tX_i\right)\right] \\
&= \mathrm{E}\left[\prod_{i=1}^{n} \exp(tX_i)\right] \\
&= \prod_{i=1}^{n} \mathrm{E}[\exp(tX_i)] && \text{(by mutual independence)} \\
&= \prod_{i=1}^{n} M_{X_i}(t) && \text{(by the definition of mgf)}
\end{aligned}$$

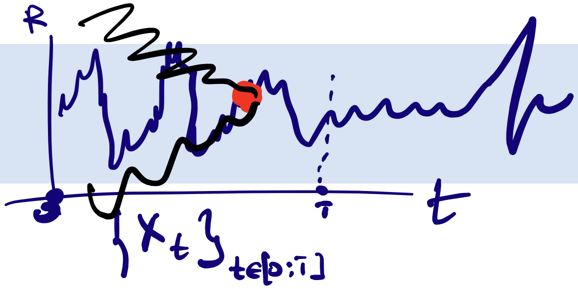
Moment Generating Function - Problem

Let X be a random variable with moment generating function

$$M_X(t) = \frac{1}{2}(1 + \exp(t))$$

Derive the variance of X.

Stochastic Processes: Basic Definitions



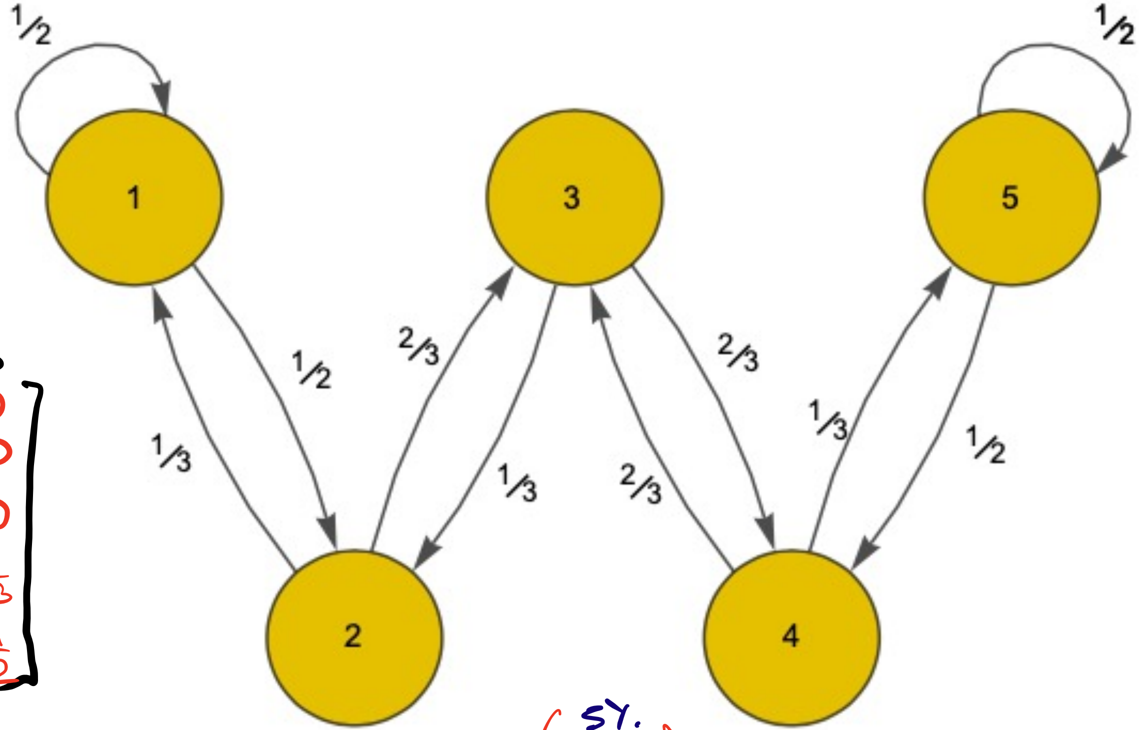
Stochastic process	The value of a variable changes in an uncertain way
Discrete vs. continuous time	When can a variable change? What values can a variable take?
Markov property	Only the current value of a variable is relevant for future predictions No information from past prices or path

A Chain

$P_{i \to i+1}$

$$
P_{i+1}
$$

$$
P_i \quad \begin{array}{c|ccccc} & 1 & 2 & 3 & 4 & 5 \\ \hline 1 & 1/2 & 1/2 & 0 & 0 & 0 \\ 2 & 1/3 & 0 & 2/3 & 0 & 0 \\ 3 & 0 & 1/3 & 0 & 2/3 & 0 \\ 4 & 0 & 0 & 2/3 & 0 & \frac{1}{3} \\ 5 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{array}
$$

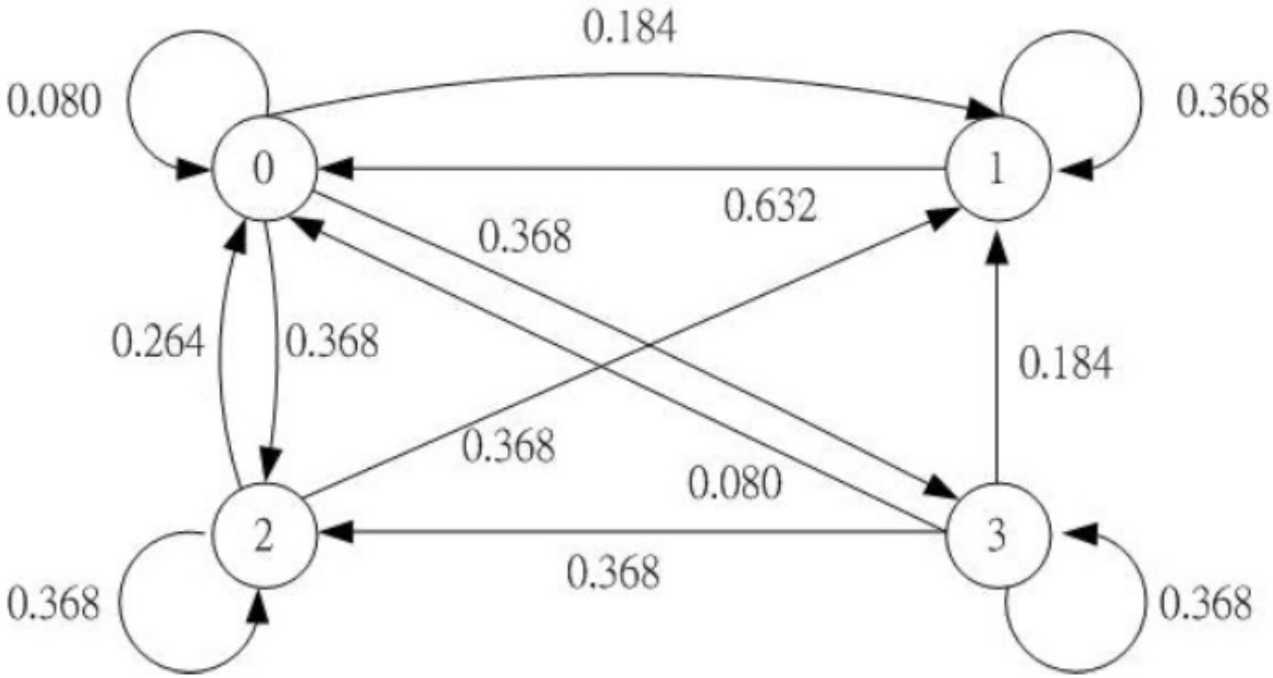


$V_2 = P_{i \to i+1} \cdot V_1$

$$
V_1 = \begin{pmatrix} 5\% \\ 25\% \\ 7\% \\ 13\% \\ 50\% \end{pmatrix}
$$

Markov Chain

- The state transition diagram:



Markov Chain

- Consider time index $n = 0, 1, 2, \ldots$ & time dependent random state X_n
- State X_n takes values on a countable number of states
  - In general denotes states as $i = 0, 1, 2, \ldots$
  - Might change with problem
- Denote the history of the process $\mathbf{X}_n = [X_n, X_{n-1}, \ldots, X_0]^T$
- Denote stochastic process as $\mathbf{X}_{\mathbb{N}}$
- The stochastic process $X_{\mathbb{N}}$ is a Markov chain (MC) if

$$\mathrm{P}\left[X_{n+1} = j \,\middle|\, X_n = i, \mathbf{X}_{n-1}\right] = \mathrm{P}\left[X_{n+1} = j \,\middle|\, X_n = i\right] = P_{ij}$$

- Future depends only on current state X_n

Observations

- Process's history $\mathbf{X}_{n-1}$ irrelevant for future evolution of the process
- Probabilities P_{ij} are constant for all times (time invariant)

- From the definition we have that for arbitrary m

$$\mathrm{P}\left[X_{n+m} \,\middle|\, X_n, \mathbf{X}_{n-1}\right] = \mathrm{P}\left[X_{n+m} \,\middle|\, X_n\right]$$

- X_{n+m} depends only on X_{n+m-1}, which depends only onX_{n+m-2}, $\ldots$ which depends only on X_n

- Since P_{ij}'s are probabilities they're positive and sum up to 1

$$P_{ij} \geq 0 \qquad \sum_{j=1}^{\infty} P_{ij} = 1$$

Matrix Representation

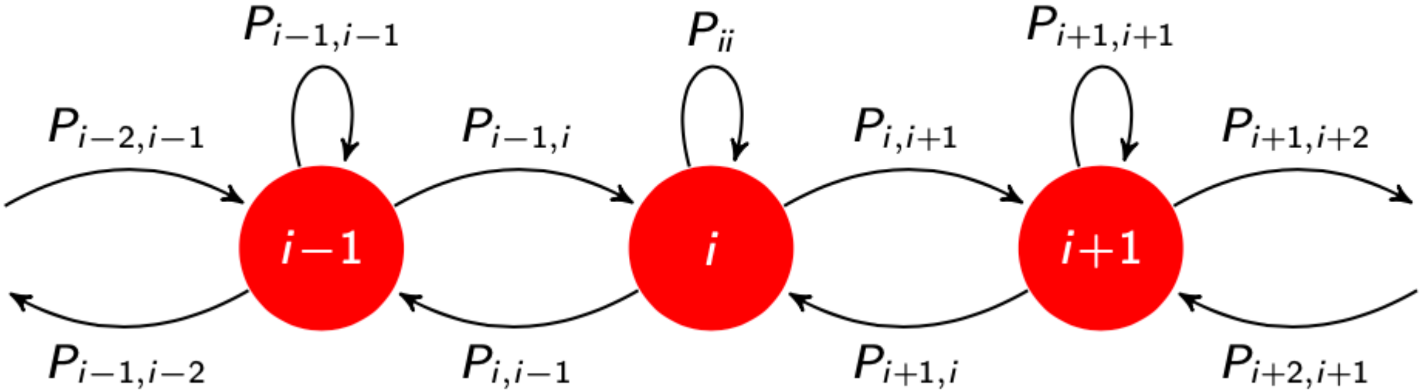
- Group transition probabilities P_{ij} in a "matrix" $\mathbf{P}$

$$
\mathbf{P} := \begin{pmatrix}
P_{00} & P_{01} & P_{02} & \cdots \\
P_{10} & P_{11} & P_{12} & \cdots \\
\vdots & \vdots & \vdots & \vdots \\
P_{i0} & P_{i1} & P_{i2} & \cdots \\
\vdots & \vdots & \vdots & \ddots
\end{pmatrix}
$$

- Not really a matrix if number of states is infinite

Graph Representation

- A graph representation is also used



- Useful when number of states is infinite

Thank you for your attention!

See next week!