Probability Calculus

Anna Janicka

lecture XIV, 15.01.2019

Markov Chains

Plan for Today

CLT – examples cont.

Markov chains

- introduction
- basic definitions
- some more definitions
- ergodic theorem
- some more definitions and problems

Central Limit Theorem – reminder

1. Classical version:

Let $X_1, X_2, \ldots$ be identically distributed independent random variables, such that $\mathbb{E}X_1^2 < \infty$. If by $m = \mathbb{E}X_1$ we denote the mean, and by $\sigma^2 = \mathrm{Var}X_1$ the variance of this distribution, then for any $t \in \mathbb{R}$, we have that

$$\mathbb{P}\left(\frac{X_1+X_2+\ldots+X_n-nm}{\sigma\sqrt{n}} \leqslant t\right) \xrightarrow{n\to\infty} \Phi(t),$$

where $\Phi(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)dx$

is the CDF of the standard normal distribution.

also:

$$\mathbb{P}\left(s \leqslant \frac{X_1+X_2+\ldots+X_n-nm}{\sigma\sqrt{n}}\right) \xrightarrow{n\to\infty} 1-\Phi(s)$$

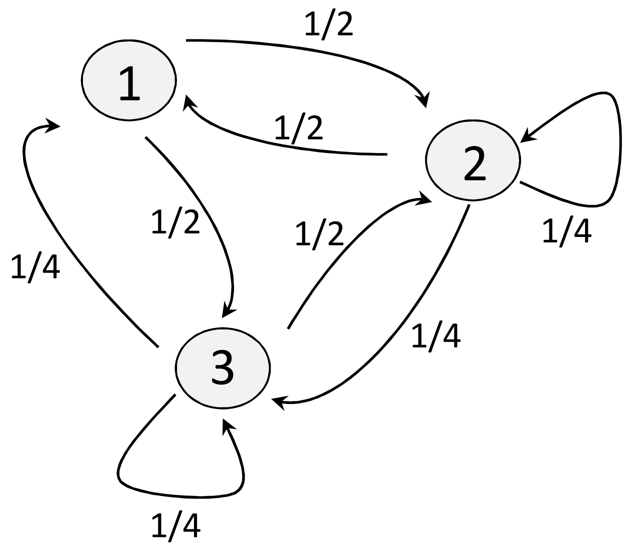
$$\mathbb{P}\left(s \leqslant \frac{X_1+X_2+\ldots+X_n-nm}{\sigma\sqrt{n}} \leqslant t\right) \xrightarrow{n\to\infty} \Phi(t)-\Phi(s)$$

Central Limit Theorem – examples

- how many students should be accepted?
- aggregate errors
- confidence intervals

Markov chains

Example: 3 states, transition probabilities



$$\begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}$$

Markov chains: definition

A sequence of random variables $(X_n)_{n=0}^{\infty}$, *taking on values in a finite set* E *is a* **Markov Chain**, *if for any* $n = 1, 2, 3, \ldots$ *and any sequence* $x_0, x_1, \ldots, x_n$ *of elements of the set* E, *we have*

$$\mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \ldots, X_0 = x_0) =$$
$$= \mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1}),$$

provided that
$\mathbb{P}(X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \ldots, X_0 = x_0) > 0.$
If, for any $i, j \in E$, $\mathbb{P}(X_n = j | X_{n-1} = i)$ *does not depend on* n, *the chain is called* **time-homogenous**. *In this case, we can define the* **transition matrix** $P = (p_{ij})_{ij \in E}$, *by the formula* $p_{ij} = \mathbb{P}(X_1 = j | X_0 = i)$.

Markov chains: properties

- sum of elements in a row of the transition matrix = 1, not necessarily so for a column
- more generally: a transition matrix in *n* steps
- more generally: Markov chains for infinite state spaces
- modelling dependence on more than the present

Vector representation of distributions

Theorem:

Let $X_0, X_1, X_2, \ldots$ be a Markov chain with an initial state $\mathbf{q}$ and a transition matrix P. Then, the variable X_n has a distribution equal to $\mathbf{q} \cdot P^n$, and the (so-called) matrix of transition in n steps, whose elements are denoted by $p_{ij}(n)$, is equal to P^n. In other words, for any $j \in E$ we have $\mathbb{P}(X_n = j) = \sum_{i_0 \in E} \sum_{i_1 \in E} \cdots \sum_{i_{n-1} \in E} \mathbf{q}_{i_0} p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} j}$.

Markov chains: characteristics

A Markov chain is **irreducible**, *if for any* $i, j \in E$ *there exists an* $n > 0$, *such that* $p_{ij}(n) > 0$; *in other words, it is possible to go from any state to any state (not necessarily in one step) – each two states* **communicate**.

A state i *has a* **period** *equal to* k, *if* $o(i) = GCD(n : p_{ii}(n) > 0) = k$. *A state is* **aperiodic** *if* $o(i) = 1$, *and* **periodic** *if* $o(i) > 1$.

Markov chains: characteristocs – cont.

Theorem: *If a Markov chain is irreducible, all chains have the same period.*

More definitions:

An irreducible Markov chain is **periodic**, *if all the states are periodic; it is* **aperiodic**, *if all the states are aperiodic.*

A distribution (vector) π *is a* **stationary distribution** *(or state) of a Markov chain of a transition matrix* P, *if* $\pi \cdot P = \pi$.

stationary distribution is calculated based on a system of equations, does not always have a unique solution

Markov chains: Ergodic Theorem

Let $(X_n)_{n \geqslant 0}$ *be an aperiodic irreducible Markov chain over a finite set of states. Then, this Markov chain has a single stationary distribution* π*, which also satisfies the following property: for any* $i, j \in E$*, we have* $\lim_{n \to \infty} p_{ij}(n) = \pi_j > 0$.

Consequences:

- limit distribution does not depend on initial state
- stationary state describes behavior in the far future

Markov chains: still more definitions

The **mean first passage time** *from state i to j for an irreducible Markov chain is the expected number of steps to reach state j from i for the first time, denoted by m_{ij}.*

The **mean recurrence time** *for state i for an irreducible Markov chain is the expected number of steps to return to state i for the first time, denoted by m_i.*

Calculation: systems of equations:

$$m_{ij} = 1 + \sum_{k \neq j} p_{ik} m_{kj} \qquad m_i = 1 + \sum_{k} p_{ik} m_{ki}$$

By convention, $m_{ii} = 0$

a finite value may be calculated for Markov chains which are not irreducible, but not for all states

Markov chains: more properties

Theorem:

Let $(X_n)_{n \geqslant 0}$ *be an aperiodic irreducible Markov chain over a finite set of states. Then, we have that the stationary distribution satisfies* $\pi_j = \frac{1}{m_j}$*, where* m_j *is the mean recurrence time for state* j.

Markov chains: more properties (2)

Definition:

A state i is **absorbing** *if it is impossible to leave the state; in other words, $p_{ii} = 1$ (while $p_{ij} = 0$ for $j \neq i$).*

Typical problems:

- calculate the probability of reaching an absorbing state
- calculate the average time until reaching an absorbing state