



WIKIPEDIA
The Free Encyclopedia

Markov chain

A **Markov chain** or **Markov process** is a stochastic model describing a sequence of possible events in which the probability of each event depends only on the state attained in the previous event.^{[1][2][3]} Informally, this may be thought of as, "What happens next depends only on the state of affairs *now*." A countably infinite sequence, in which the chain moves state at discrete time steps, gives a discrete-time Markov chain (DTMC). A continuous-time process is called a continuous-time Markov chain (CTMC). It is named after the Russian mathematician Andrey Markov.

Markov chains have many applications as statistical models of real-world processes,^{[1][4][5][6]} such as studying cruise control systems in motor vehicles, queues or lines of customers arriving at an airport, currency exchange rates and animal population dynamics.^[7]

Markov processes are the basis for general stochastic simulation methods known as Markov chain Monte Carlo, which are used for simulating sampling from complex probability distributions, and have found application in Bayesian statistics, thermodynamics, statistical mechanics, physics, chemistry, economics, finance, signal processing, information theory and speech processing.^{[7][8][9]}

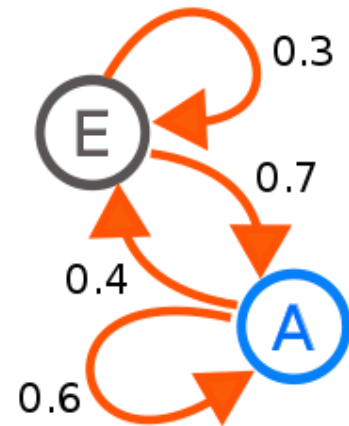
The adjectives *Markovian* and *Markov* are used to describe something that is related to a Markov process.^{[1][10][11]}

Principles

Definition

A Markov process is a stochastic process that satisfies the Markov property^[1] (sometimes characterized as "memorylessness"). In simpler terms, it is a process for which predictions can be made regarding future outcomes based solely on its present state and—most importantly—such predictions are just as good as the ones that could be made knowing the process's full history.^[12] In other words, conditional on the present state of the system, its future and past states are independent.

A Markov chain is a type of Markov process that has either a discrete state space or a discrete index set (often representing time), but the precise definition of a Markov chain varies.^[13] For example, it is common to define a Markov chain as a Markov process in either discrete or continuous time with a countable state space (thus regardless of the



A diagram representing a two-state Markov process. The numbers are the probability of changing from one state to another state.



Russian mathematician Andrey Markov

nature of time),^{[14][15][16][17]} but it is also common to define a Markov chain as having discrete time in either countable or continuous state space (thus regardless of the state space).^[13]

Types of Markov chains

The system's state space and time parameter index need to be specified. The following table gives an overview of the different instances of Markov processes for different levels of state space generality and for discrete time v. continuous time:

	Countable state space	Continuous or general state space
Discrete-time	(discrete-time) Markov chain on a countable or finite state space	Markov chain on a measurable state space (for example, Harris chain)
Continuous-time	Continuous-time Markov process or Markov jump process	Any continuous stochastic process with the Markov property (for example, the Wiener process)

Note that there is no definitive agreement in the literature on the use of some of the terms that signify special cases of Markov processes. Usually the term "Markov chain" is reserved for a process with a discrete set of times, that is, a **discrete-time Markov chain (DTMC)**,^{[1][18]} but a few authors use the term "Markov process" to refer to a **continuous-time Markov chain (CTMC)** without explicit mention.^{[19][20][21]} In addition, there are other extensions of Markov processes that are referred to as such but do not necessarily fall within any of these four categories (see Markov model). Moreover, the time index need not necessarily be real-valued; like with the state space, there are conceivable processes that move through index sets with other mathematical constructs. Notice that the general state space continuous-time Markov chain is general to such a degree that it has no designated term.

While the time parameter is usually discrete, the state space of a Markov chain does not have any generally agreed-on restrictions: the term may refer to a process on an arbitrary state space.^[22] However, many applications of Markov chains employ finite or countably infinite state spaces, which have a more straightforward statistical analysis. Besides time-index and state-space parameters, there are many other variations, extensions and generalizations (see Variations). For simplicity, most of this article concentrates on the discrete-time, discrete state-space case, unless mentioned otherwise.

Transitions

The changes of state of the system are called transitions.^[1] The probabilities associated with various state changes are called transition probabilities. The process is characterized by a state space, a transition matrix describing the probabilities of particular transitions, and an initial state (or initial distribution) across the state space. By convention, we assume all possible states and transitions have been included in the definition of the process, so there is always a next state, and the process does not terminate.

A discrete-time random process involves a system which is in a certain state at each step, with the state changing randomly between steps.^[1] The steps are often thought of as moments in time, but they can equally well refer to physical distance or any other discrete measurement. Formally, the steps are the integers or natural numbers, and the random process is a mapping of these to states.^[23] The Markov property states that the conditional probability distribution for the system at the next step (and in fact at all future steps) depends only on the current state of the system, and not additionally on the state of the system at previous steps.

Since the system changes randomly, it is generally impossible to predict with certainty the state of a Markov chain at a given point in the future.^[23] However, the statistical properties of the system's future can be predicted.^[23] In many applications, it is these statistical properties that are important.

History

Andrey Markov studied Markov processes in the early 20th century, publishing his first paper on the topic in 1906.^{[24][25][26][27]} Markov Processes in continuous time were discovered long before his work in the early 20th century^[1] in the form of the Poisson process.^{[28][29][30]} Markov was interested in studying an extension of independent random sequences, motivated by a disagreement with Pavel Nekrasov who claimed independence was necessary for the weak law of large numbers to hold.^{[1][31]} In his first paper on Markov chains, published in 1906, Markov showed that under certain conditions the average outcomes of the Markov chain would converge to a fixed vector of values, so proving a weak law of large numbers without the independence assumption,^{[1][25][26][27]} which had been commonly regarded as a requirement for such mathematical laws to hold.^[27] Markov later used Markov chains to study the distribution of vowels in Eugene Onegin, written by Alexander Pushkin, and proved a central limit theorem for such chains.^{[1][25]}

In 1912 Henri Poincaré studied Markov chains on finite groups with an aim to study card shuffling. Other early uses of Markov chains include a diffusion model, introduced by Paul and Tatyana Ehrenfest in 1907, and a branching process, introduced by Francis Galton and Henry William Watson in 1873, preceding the work of Markov.^{[25][26]} After the work of Galton and Watson, it was later revealed that their branching process had been independently discovered and studied around three decades earlier by Irénée-Jules Bienaymé.^[32] Starting in 1928, Maurice Fréchet became interested in Markov chains, eventually resulting in him publishing in 1938 a detailed study on Markov chains.^{[25][33]}

Andrey Kolmogorov developed in a 1931 paper a large part of the early theory of continuous-time Markov processes.^{[34][35]} Kolmogorov was partly inspired by Louis Bachelier's 1900 work on fluctuations in the stock market as well as Norbert Wiener's work on Einstein's model of Brownian movement.^{[34][36]} He introduced and studied a particular set of Markov processes known as diffusion processes, where he derived a set of differential equations describing the processes.^{[34][37]} Independent of Kolmogorov's work, Sydney Chapman derived in a 1928 paper an equation, now called the Chapman–Kolmogorov equation, in a less mathematically rigorous way than Kolmogorov, while studying Brownian movement.^[38] The differential equations are now called the Kolmogorov equations^[39] or the Kolmogorov–Chapman equations.^[40] Other mathematicians who contributed significantly to the foundations of Markov processes include William Feller, starting in 1930s, and then later Eugene Dynkin, starting in the 1950s.^[35]

Examples

- Random walks based on integers and the gambler's ruin problem are examples of Markov processes.^{[41][42]} Some variations of these processes were studied hundreds of years earlier in the context of independent variables.^{[43][44][45]} Two important examples of Markov processes are the Wiener process, also known as the Brownian motion process, and the Poisson process,^[28] which are considered the most important and central stochastic processes in the theory of stochastic processes.^{[46][47][48]} These two processes are Markov processes in continuous time, while random walks on the integers and the gambler's ruin problem are examples of Markov processes in discrete time.^{[41][42]}
- A famous Markov chain is the so-called "drunkard's walk", a random walk on the number line where, at each step, the position may change by +1 or −1 with equal probability. From any position there are two possible transitions, to the next or previous integer. The transition probabilities depend only on the current position, not on the manner in which the position was reached. For example, the transition probabilities from 5 to 4 and 5 to 6 are both 0.5, and all other transition probabilities from 5 are 0. These probabilities are independent of whether the system was previously in 4 or 6.

- A series of independent states (for example, a series of coin flips) satisfies the formal definition of a Markov chain. However, the theory is usually applied only when the probability distribution of the next state depends on the current one.

A non-Markov example

Suppose that there is a coin purse containing five quarters (each worth 25¢), five dimes (each worth 10¢), and five nickels (each worth 5¢), and one by one, coins are randomly drawn from the purse and are set on a table. If \mathbf{X}_n represents the total value of the coins set on the table after n draws, with $\mathbf{X}_0 = \mathbf{0}$, then the sequence $\{\mathbf{X}_n : n \in \mathbb{N}\}$ is *not* a Markov process.

To see why this is the case, suppose that in the first six draws, all five nickels and a quarter are drawn. Thus $\mathbf{X}_6 = \$0.50$. If we know not just \mathbf{X}_6 , but the earlier values as well, then we can determine which coins have been drawn, and we know that the next coin will not be a nickel; so we can determine that $\mathbf{X}_7 \geq \$0.60$ with probability 1. But if we do not know the earlier values, then based only on the value \mathbf{X}_6 we might guess that we had drawn four dimes and two nickels, in which case it would certainly be possible to draw another nickel next. Thus, our guesses about \mathbf{X}_7 are impacted by our knowledge of values prior to \mathbf{X}_6 .

However, it is possible to model this scenario as a Markov process. Instead of defining \mathbf{X}_n to represent the *total value* of the coins on the table, we could define \mathbf{X}_n to represent the *count* of the various coin types on the table. For instance, $\mathbf{X}_6 = \mathbf{1}, \mathbf{0}, \mathbf{5}$ could be defined to represent the state where there is one quarter, zero dimes, and five nickels on the table after 6 one-by-one draws. This new model could be represented by $6 \times 6 \times 6 = 216$ possible states, where each state represents the number of coins of each type (from 0 to 5) that are on the table. (Not all of these states are reachable within 6 draws.) Suppose that the first draw results in state $\mathbf{X}_1 = \mathbf{0}, \mathbf{1}, \mathbf{0}$. The probability of achieving \mathbf{X}_2 now depends on \mathbf{X}_1 ; for example, the state $\mathbf{X}_2 = \mathbf{1}, \mathbf{0}, \mathbf{1}$ is not possible. After the second draw, the third draw depends on which coins have so far been drawn, but no longer only on the coins that were drawn for the first state (since probabilistically important information has since been added to the scenario). In this way, the likelihood of the $\mathbf{X}_n = i, j, k$ state depends exclusively on the outcome of the $\mathbf{X}_{n-1} = \ell, m, p$ state.

Formal definition

Discrete-time Markov chain

A discrete-time Markov chain is a sequence of random variables X_1, X_2, X_3, \dots with the Markov property, namely that the probability of moving to the next state depends only on the present state and not on the previous states:

$$\Pr(X_{n+1} = x \mid X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = \Pr(X_{n+1} = x \mid X_n = x_n), \text{ if both conditional probabilities are well defined, that is, if } \Pr(X_1 = x_1, \dots, X_n = x_n) > 0.$$

The possible values of X_i form a countable set S called the state space of the chain.

Variations

- Time-homogeneous Markov chains are processes where

$$\Pr(X_{n+1} = x \mid X_n = y) = \Pr(X_n = x \mid X_{n-1} = y)$$

for all n . The probability of the transition is independent of n .

- Stationary Markov chains are processes where

$$\Pr(X_0 = x_0, X_1 = x_1, \dots, X_k = x_k) = \Pr(X_n = x_0, X_{n+1} = x_1, \dots, X_{n+k} = x_k)$$

for all n and k . Every stationary chain can be proved to be time-homogeneous by Bayes' rule.

A necessary and sufficient condition for a time-homogeneous Markov chain to be stationary is that the distribution of \mathbf{X}_0 is a stationary distribution of the Markov chain.

- A Markov chain with memory (or a Markov chain of order m) where m is finite, is a process satisfying

$$\begin{aligned} &\Pr(X_n = x_n \mid X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \dots, X_1 = x_1) \\ &= \Pr(X_n = x_n \mid X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \dots, X_{n-m} = x_{n-m}) \text{ for } n > m \end{aligned}$$

In other words, the future state depends on the past m states. It is possible to construct a chain (Y_n) from (X_n) which has the 'classical' Markov property by taking as state space the ordered m -tuples of X values, i.e., $Y_n = (X_n, X_{n-1}, \dots, X_{n-m+1})$.

Continuous-time Markov chain

A continuous-time Markov chain $(X_t)_{t \geq 0}$ is defined by a finite or countable state space S , a transition rate matrix Q with dimensions equal to that of the state space and initial probability distribution defined on the state space. For $i \neq j$, the elements q_{ij} are non-negative and describe the rate of the process transitions from state i to state j . The elements q_{ii} are chosen such that each row of the transition rate matrix sums to zero, while the row-sums of a probability transition matrix in a (discrete) Markov chain are all equal to one.

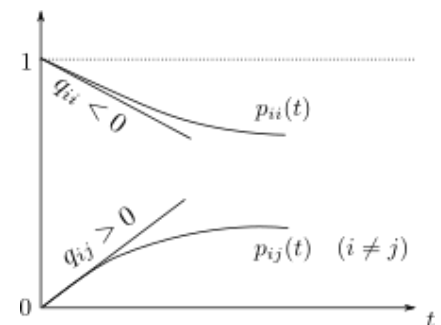
There are three equivalent definitions of the process.^[49]

Infinitesimal definition

Let \mathbf{X}_t be the random variable describing the state of the process at time t , and assume the process is in a state i at time t . Then, knowing $\mathbf{X}_t = i$, $\mathbf{X}_{t+h} = j$ is independent of previous values $(\mathbf{X}_s : s < t)$, and as $h \rightarrow 0$ for all j and for all t ,

$$\Pr(X(t+h) = j \mid X(t) = i) = \delta_{ij} + q_{ij}h + o(h),$$

where δ_{ij} is the Kronecker delta, using the little-o notation. The q_{ij} can be seen as measuring how quickly the transition from i to j happens.



The continuous time Markov chain is characterized by the transition rates, the derivatives with respect to time of the transition probabilities between states i and j .

Jump chain/holding time definition

Define a discrete-time Markov chain Y_n to describe the n th jump of the process and variables S_1, S_2, S_3, \dots to describe holding times in each of the states where S_i follows the exponential distribution with rate parameter $-q_{Y_i Y_i}$.

Transition probability definition

For any value $n = 0, 1, 2, 3, \dots$ and times indexed up to this value of n : t_0, t_1, t_2, \dots and all states recorded at these times $i_0, i_1, i_2, i_3, \dots$ it holds that

$$\Pr(X_{t_{n+1}} = i_{n+1} \mid X_{t_0} = i_0, X_{t_1} = i_1, \dots, X_{t_n} = i_n) = p_{i_n i_{n+1}}(t_{n+1} - t_n)$$

where p_{ij} is the solution of the forward equation (a first-order differential equation)

$$P'(t) = P(t)Q$$

with initial condition $P(0)$ is the identity matrix.

Finite state space

If the state space is finite, the transition probability distribution can be represented by a matrix, called the transition matrix, with the (i, j) th element of \mathbf{P} equal to

$$p_{ij} = \Pr(X_{n+1} = j \mid X_n = i).$$

Since each row of \mathbf{P} sums to one and all elements are non-negative, \mathbf{P} is a right stochastic matrix.

Stationary distribution relation to eigenvectors and simplices

A stationary distribution π is a (row) vector, whose entries are non-negative and sum to 1, is unchanged by the operation of transition matrix \mathbf{P} on it and so is defined by

$$\pi \mathbf{P} = \pi.$$

By comparing this definition with that of an eigenvector we see that the two concepts are related and that

$$\pi = \frac{\mathbf{e}}{\sum_i e_i}$$

is a normalized ($\sum_i \pi_i = 1$) multiple of a left eigenvector \mathbf{e} of the transition matrix \mathbf{P} with an eigenvalue of 1. If there is more than one unit eigenvector then a weighted sum of the corresponding stationary states is also a stationary state. But for a Markov chain one is usually more interested in a stationary state that is the limit of the sequence of distributions for some initial distribution.

The values of a stationary distribution π_i are associated with the state space of \mathbf{P} and its eigenvectors have their relative proportions preserved. Since the components of π are positive and the constraint that their sum is unity can be rewritten as $\sum_i \mathbf{1} \cdot \pi_i = 1$ we see that the dot product of π with a vector whose components are all 1 is unity and that π lies on a simplex.

Time-homogeneous Markov chain with a finite state space

If the Markov chain is time-homogeneous, then the transition matrix \mathbf{P} is the same after each step, so the k -step transition probability can be computed as the k -th power of the transition matrix, \mathbf{P}^k .

If the Markov chain is irreducible and aperiodic, then there is a unique stationary distribution π .^[50] Additionally, in this case \mathbf{P}^k converges to a rank-one matrix in which each row is the stationary distribution π :

$$\lim_{k \rightarrow \infty} \mathbf{P}^k = \mathbf{1}\pi$$

where $\mathbf{1}$ is the column vector with all entries equal to 1. This is stated by the Perron–Frobenius theorem. If, by whatever means, $\lim_{k \rightarrow \infty} \mathbf{P}^k$ is found, then the stationary distribution of the Markov chain in question can be easily determined for any starting distribution, as will be explained below.

For some stochastic matrices \mathbf{P} , the limit $\lim_{k \rightarrow \infty} \mathbf{P}^k$ does not exist while the stationary distribution does, as shown by this example:

$$\mathbf{P} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \mathbf{P}^{2k} = \mathbf{I} \quad \mathbf{P}^{2k+1} = \mathbf{P}$$

$$\left(\frac{1}{2} \quad \frac{1}{2}\right) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \left(\frac{1}{2} \quad \frac{1}{2}\right)$$

(This example illustrates a periodic Markov chain.)

Because there are a number of different special cases to consider, the process of finding this limit if it exists can be a lengthy task. However, there are many techniques that can assist in finding this limit. Let \mathbf{P} be an $n \times n$ matrix, and define $\mathbf{Q} = \lim_{k \rightarrow \infty} \mathbf{P}^k$.

It is always true that

$$\mathbf{Q}\mathbf{P} = \mathbf{Q}.$$

Subtracting \mathbf{Q} from both sides and factoring then yields

$$\mathbf{Q}(\mathbf{P} - \mathbf{I}_n) = \mathbf{0}_{n,n},$$

where \mathbf{I}_n is the identity matrix of size n , and $\mathbf{0}_{n,n}$ is the zero matrix of size $n \times n$. Multiplying together stochastic matrices always yields another stochastic matrix, so \mathbf{Q} must be a stochastic matrix (see the definition above). It is sometimes sufficient to use the matrix equation above and the fact that \mathbf{Q} is a stochastic matrix to solve for \mathbf{Q} . Including the fact that the sum of each the rows in \mathbf{P} is 1, there are $n+1$ equations for determining n unknowns, so it is computationally easier if on the one hand one selects one row in \mathbf{Q} and substitutes each of its elements by one, and on the other one substitutes the corresponding element (the one in the same column) in the vector \mathbf{o} , and next left-multiplies this latter vector by the inverse of transformed former matrix to find \mathbf{Q} .

Here is one method for doing so: first, define the function $f(\mathbf{A})$ to return the matrix \mathbf{A} with its right-most column replaced with all 1's. If $[f(\mathbf{P} - \mathbf{I}_n)]^{-1}$ exists then^{[51][50]}

$$\mathbf{Q} = f(\mathbf{0}_{n,n})[f(\mathbf{P} - \mathbf{I}_n)]^{-1}.$$

Explain: The original matrix equation is equivalent to a system of $n \times n$ linear equations in $n \times n$ variables. And there are n more linear equations from the fact that \mathbf{Q} is a right stochastic matrix whose each row sums to 1. So it needs any $n \times n$ independent linear equations of the $(n \times n + n)$ equations to solve for the $n \times n$ variables. In this example, the n equations from " \mathbf{Q} multiplied by the right-most column of $(\mathbf{P} - \mathbf{I}_n)$ " have been replaced by the n stochastic ones.

One thing to notice is that if \mathbf{P} has an element $\mathbf{P}_{i,i}$ on its main diagonal that is equal to 1 and the i th row or column is otherwise filled with 0's, then that row or column will remain unchanged in all of the subsequent powers \mathbf{P}^k . Hence, the i th row or column of \mathbf{Q} will have the 1 and the 0's in the same positions as in \mathbf{P} .

Convergence speed to the stationary distribution

As stated earlier, from the equation $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}$, (if exists) the stationary (or steady state) distribution $\boldsymbol{\pi}$ is a left eigenvector of row stochastic matrix \mathbf{P} . Then assuming that \mathbf{P} is diagonalizable or equivalently that \mathbf{P} has n linearly independent eigenvectors, speed of convergence is elaborated as follows. (For non-diagonalizable, that is, defective matrices, one may start with the Jordan normal form of \mathbf{P} and proceed with a bit more involved set of arguments in a similar way.^[52])

Let \mathbf{U} be the matrix of eigenvectors (each normalized to having an L2 norm equal to 1) where each column is a left eigenvector of \mathbf{P} and let $\mathbf{\Sigma}$ be the diagonal matrix of left eigenvalues of \mathbf{P} , that is, $\mathbf{\Sigma} = \text{diag}(\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n)$. Then by eigendecomposition

$$\mathbf{P} = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^{-1}.$$

Let the eigenvalues be enumerated such that:

$$1 = |\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|.$$

Since \mathbf{P} is a row stochastic matrix, its largest left eigenvalue is 1. If there is a unique stationary distribution, then the largest eigenvalue and the corresponding eigenvector is unique too (because there is no other π which solves the stationary distribution equation above). Let \mathbf{u}_i be the i -th column of \mathbf{U} matrix, that is, \mathbf{u}_i is the left eigenvector of \mathbf{P} corresponding to λ_i . Also let \mathbf{x} be a length n row vector that represents a valid probability distribution; since the eigenvectors \mathbf{u}_i span \mathbb{R}^n , we can write

$$\mathbf{x}^T = \sum_{i=1}^n a_i \mathbf{u}_i, \quad a_i \in \mathbb{R}.$$

If we multiply \mathbf{x} with \mathbf{P} from right and continue this operation with the results, in the end we get the stationary distribution π . In other words, $\pi = \mathbf{a}_1 \mathbf{u}_1 \leftarrow \mathbf{xPP} \dots \mathbf{P} = \mathbf{xP}^k$ as $k \rightarrow \infty$. That means

$$\begin{aligned} \pi^{(k)} &= \mathbf{x} (\mathbf{U}\mathbf{\Sigma}\mathbf{U}^{-1}) (\mathbf{U}\mathbf{\Sigma}\mathbf{U}^{-1}) \dots (\mathbf{U}\mathbf{\Sigma}\mathbf{U}^{-1}) \\ &= \mathbf{x}\mathbf{U}\mathbf{\Sigma}^k\mathbf{U}^{-1} \\ &= (a_1 \mathbf{u}_1^T + a_2 \mathbf{u}_2^T + \dots + a_n \mathbf{u}_n^T) \mathbf{U}\mathbf{\Sigma}^k\mathbf{U}^{-1} \\ &= a_1 \lambda_1^k \mathbf{u}_1^T + a_2 \lambda_2^k \mathbf{u}_2^T + \dots + a_n \lambda_n^k \mathbf{u}_n^T \\ &= \lambda_1^k \left\{ a_1 \mathbf{u}_1^T + a_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k \mathbf{u}_2^T + a_3 \left(\frac{\lambda_3}{\lambda_1} \right)^k \mathbf{u}_3^T + \dots + a_n \left(\frac{\lambda_n}{\lambda_1} \right)^k \mathbf{u}_n^T \right\} \end{aligned} \quad u_i \perp u_j \text{ for } i \neq j$$

Since π is parallel to \mathbf{u}_1 (normalized by L2 norm) and $\pi^{(k)}$ is a probability vector, $\pi^{(k)}$ approaches to $\mathbf{a}_1 \mathbf{u}_1 = \pi$ as $k \rightarrow \infty$ with a speed in the order of λ_2/λ_1 exponentially. This follows because $|\lambda_2| \geq \dots \geq |\lambda_n|$, hence λ_2/λ_1 is the dominant term. The smaller the ratio is, the faster the convergence is.^[53] Random noise in the state distribution π can also speed up this convergence to the stationary distribution.^[54]

General state space

Harris chains

Many results for Markov chains with finite state space can be generalized to chains with uncountable state space through Harris chains.

The use of Markov chains in Markov chain Monte Carlo methods covers cases where the process follows a continuous state space.

Locally interacting Markov chains

Considering a collection of Markov chains whose evolution takes in account the state of other Markov chains, is related to the notion of locally interacting Markov chains. This corresponds to the situation when the state space has a (Cartesian-) product form. See interacting particle system and stochastic cellular automata (probabilistic cellular automata). See for instance *Interaction of Markov Processes*^[55] or.^[56]

Properties

Two states are said to *communicate* with each other if both are reachable from one another by a sequence of transitions that have positive probability. This is an equivalence relation which yields a set of communicating classes. A class is *closed* if the probability of leaving the class is zero. A Markov chain is *irreducible* if there is one communicating class, the state space.

A state i has period k if k is the greatest common divisor of the number of transitions by which i can be reached, starting from i . That is:

$$k = \gcd\{n > 0 : \Pr(X_n = i \mid X_0 = i) > 0\}$$

A state i is said to be *transient* if, starting from i , there is a non-zero probability that the chain will never return to i . It is called *recurrent* (or *persistent*) otherwise.^[57] For a recurrent state i , the mean *hitting time* is defined as:

$$M_i = E[T_i] = \sum_{n=1}^{\infty} n \cdot f_{ii}^{(n)}.$$

State i is *positive recurrent* if M_i is finite and *null recurrent* otherwise. Periodicity, transience, recurrence and positive and null recurrence are class properties — that is, if one state has the property then all states in its communicating class have the property.

A state i is called *absorbing* if there are no outgoing transitions from the state.

Ergodicity

A state i is said to be *ergodic* if it is aperiodic and positive recurrent. In other words, a state i is ergodic if it is recurrent, has a period of 1, and has finite mean recurrence time. If all states in an irreducible Markov chain are ergodic, then the chain is said to be ergodic. Some authors call any irreducible, positive recurrent Markov chains ergodic, even periodic ones.^[58]

It can be shown that a finite state irreducible Markov chain is ergodic if it has an aperiodic state. More generally, a Markov chain is ergodic if there is a number N such that any state can be reached from any other state in any number of steps less or equal to a number N . In case of a fully connected transition matrix, where all transitions have a non-zero probability, this condition is fulfilled with $N = 1$.

A Markov chain with more than one state and just one out-going transition per state is either not irreducible or not aperiodic, hence cannot be ergodic.

Markovian representations

In some cases, apparently non-Markovian processes may still have Markovian representations, constructed by expanding the concept of the "current" and "future" states. For example, let X be a non-Markovian process. Then define a process Y , such that each state of Y represents a time-interval of states of X . Mathematically, this takes the form:

$$Y(t) = \{X(s) : s \in [a(t), b(t)]\}.$$

If Y has the Markov property, then it is a Markovian representation of X .

An example of a non-Markovian process with a Markovian representation is an autoregressive time series of order greater than one.^[59]

Hitting times

The *hitting time* is the time, starting in a given set of states until the chain arrives in a given state or set of states. The distribution of such a time period has a phase type distribution. The simplest such distribution is that of a single exponentially distributed transition.

Expected hitting times

For a subset of states $A \subseteq S$, the vector k^A of hitting times (where element k_i^A represents the expected value, starting in state i that the chain enters one of the states in the set A) is the minimal non-negative solution to^[60]

$$\begin{aligned} k_i^A &= 0 \text{ for } i \in A \\ - \sum_{j \in S} q_{ij} k_j^A &= 1 \text{ for } i \notin A. \end{aligned}$$

Time reversal

For a CTMC X_t , the time-reversed process is defined to be $\hat{X}_t = X_{T-t}$. By Kelly's lemma this process has the same stationary distribution as the forward process.

A chain is said to be *reversible* if the reversed process is the same as the forward process. Kolmogorov's criterion states that the necessary and sufficient condition for a process to be reversible is that the product of transition rates around a closed loop must be the same in both directions.

Embedded Markov chain

One method of finding the stationary probability distribution, π , of an ergodic continuous-time Markov chain, Q , is by first finding its **embedded Markov chain (EMC)**. Strictly speaking, the EMC is a regular discrete-time Markov chain, sometimes referred to as a **jump process**. Each element of the one-step transition probability matrix of the EMC, S , is denoted by s_{ij} , and represents the conditional probability of transitioning from state i into state j . These conditional probabilities may be found by

$$s_{ij} = \begin{cases} \frac{q_{ij}}{\sum_{k \neq i} q_{ik}} & \text{if } i \neq j \\ 0 & \text{otherwise.} \end{cases}$$

From this, S may be written as

$$S = I - (\text{diag}(Q))^{-1}Q$$

where I is the identity matrix and $\text{diag}(Q)$ is the diagonal matrix formed by selecting the main diagonal from the matrix Q and setting all other elements to zero.

To find the stationary probability distribution vector, we must next find φ such that

$$\varphi S = \varphi,$$

with φ being a row vector, such that all elements in φ are greater than 0 and $\|\varphi\|_1 = 1$. From this, π may be found as

$$\pi = \frac{-\varphi(\text{diag}(Q))^{-1}}{\|\varphi(\text{diag}(Q))^{-1}\|_1}.$$

(S may be periodic, even if Q is not. Once π is found, it must be normalized to a unit vector.)

Another discrete-time process that may be derived from a continuous-time Markov chain is a δ -skeleton—the (discrete-time) Markov chain formed by observing $X(t)$ at intervals of δ units of time. The random variables $X(0), X(\delta), X(2\delta), \dots$ give the sequence of states visited by the δ -skeleton.

Special types of Markov chains

Markov model

Markov models are used to model changing systems. There are 4 main types of models, that generalize Markov chains depending on whether every sequential state is observable or not, and whether the system is to be adjusted on the basis of observations made:

	System state is fully observable	System state is partially observable
System is autonomous	Markov chain	<u>Hidden Markov model</u>
System is controlled	<u>Markov decision process</u>	<u>Partially observable Markov decision process</u>

Bernoulli scheme

A Bernoulli scheme is a special case of a Markov chain where the transition probability matrix has identical rows, which means that the next state is independent of even the current state (in addition to being independent of the past states). A Bernoulli scheme with only two possible states is known as a Bernoulli process.

Note, however, by the Ornstein isomorphism theorem, that every aperiodic and irreducible Markov chain is isomorphic to a Bernoulli scheme;^[61] thus, one might equally claim that Markov chains are a "special case" of Bernoulli schemes. The isomorphism generally requires a complicated recoding. The isomorphism theorem is even a bit stronger: it states that any stationary stochastic process is isomorphic to a Bernoulli scheme; the Markov chain is just one such example.

Subshift of finite type

When the Markov matrix is replaced by the adjacency matrix of a finite graph, the resulting shift is termed a **topological Markov chain** or a **subshift of finite type**.^[61] A Markov matrix that is compatible with the adjacency matrix can then provide a measure on the subshift. Many chaotic dynamical systems are

isomorphic to topological Markov chains; examples include diffeomorphisms of closed manifolds, the Prouhet–Thue–Morse system, the Chacon system, sofic systems, context-free systems and block-coding systems.^[61]

Applications

Research has reported the application and usefulness of Markov chains in a wide range of topics such as physics, chemistry, biology, medicine, music, game theory and sports.

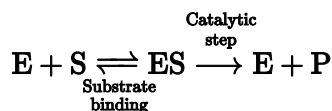
Physics

Markovian systems appear extensively in thermodynamics and statistical mechanics, whenever probabilities are used to represent unknown or unmodelled details of the system, if it can be assumed that the dynamics are time-invariant, and that no relevant history need be considered which is not already included in the state description.^{[62][63]} For example, a thermodynamic state operates under a probability distribution that is difficult or expensive to acquire. Therefore, Markov Chain Monte Carlo method can be used to draw samples randomly from a black-box to approximate the probability distribution of attributes over a range of objects.^[63]

The paths, in the path integral formulation of quantum mechanics, are Markov chains.^[64]

Markov chains are used in lattice QCD simulations.^[65]

Chemistry



Michaelis-Menten kinetics. The enzyme (E) binds a substrate (S) and produces a product (P). Each reaction is a state transition in a Markov chain.

A reaction network is a chemical system involving multiple reactions and chemical species. The simplest stochastic models of such networks treat the system as a continuous time Markov chain with the state being the number of molecules of each species and with reactions modeled as possible transitions of the chain.^[66] Markov chains and continuous-time Markov processes are useful in chemistry when physical systems closely approximate the Markov property. For example, imagine a large number n of molecules in solution in state A, each of which can undergo a chemical reaction to state B with a certain average rate. Perhaps the molecule is an enzyme, and the states refer to how it is folded. The state of any single

enzyme follows a Markov chain, and since the molecules are essentially independent of each other, the number of molecules in state A or B at a time is n times the probability a given molecule is in that state.

The classical model of enzyme activity, Michaelis–Menten kinetics, can be viewed as a Markov chain, where at each time step the reaction proceeds in some direction. While Michaelis-Menten is fairly straightforward, far more complicated reaction networks can also be modeled with Markov chains.^[67]

An algorithm based on a Markov chain was also used to focus the fragment-based growth of chemicals in silico towards a desired class of compounds such as drugs or natural products.^[68] As a molecule is grown, a fragment is selected from the nascent molecule as the "current" state. It is not aware of its past (that is, it is not aware of what is already bonded to it). It then transitions to the next state when a fragment is attached to it. The transition probabilities are trained on databases of authentic classes of compounds.^[69]

Also, the growth (and composition) of copolymers may be modeled using Markov chains. Based on the reactivity ratios of the monomers that make up the growing polymer chain, the chain's composition may be calculated (for example, whether monomers tend to add in alternating fashion or in long runs of the same

monomer). Due to steric effects, second-order Markov effects may also play a role in the growth of some polymer chains.

Similarly, it has been suggested that the crystallization and growth of some epitaxial superlattice oxide materials can be accurately described by Markov chains.^[70]

Biology

Markov chains are used in various areas of biology. Notable examples include:

- Phylogenetics and bioinformatics, where most models of DNA evolution use continuous-time Markov chains to describe the nucleotide present at a given site in the genome.
- Population dynamics, where Markov chains are in particular a central tool in the theoretical study of matrix population models.
- Neurobiology, where Markov chains have been used, e.g., to simulate the mammalian neocortex.^[71]
- Systems biology, for instance with the modeling of viral infection of single cells.^[72]
- Compartmental models for disease outbreak and epidemic modeling.

Testing

Several theorists have proposed the idea of the Markov chain statistical test (MCST), a method of conjoining Markov chains to form a "Markov blanket", arranging these chains in several recursive layers ("wafering") and producing more efficient test sets—samples—as a replacement for exhaustive testing. MCSTs also have uses in temporal state-based networks; Chilukuri et al.'s paper entitled "Temporal Uncertainty Reasoning Networks for Evidence Fusion with Applications to Object Detection and Tracking" (ScienceDirect) gives a background and case study for applying MCSTs to a wider range of applications.

Solar irradiance variability

Solar irradiance variability assessments are useful for solar power applications. Solar irradiance variability at any location over time is mainly a consequence of the deterministic variability of the sun's path across the sky dome and the variability in cloudiness. The variability of accessible solar irradiance on Earth's surface has been modeled using Markov chains,^{[73][74][75][76]} also including modeling the two states of clear and cloudiness as a two-state Markov chain.^{[77][78]}

Speech recognition

Hidden Markov models are the basis for most modern automatic speech recognition systems.

Information theory

Markov chains are used throughout information processing. Claude Shannon's famous 1948 paper *A Mathematical Theory of Communication*, which in a single step created the field of information theory, opens by introducing the concept of entropy through Markov modeling of the English language. Such idealized models can capture many of the statistical regularities of systems. Even without describing the full structure of the system perfectly, such signal models can make possible very effective data compression through entropy encoding techniques such as arithmetic coding. They also allow effective state estimation and pattern recognition. Markov chains also play an important role in reinforcement learning.

Markov chains are also the basis for hidden Markov models, which are an important tool in such diverse fields as telephone networks (which use the Viterbi algorithm for error correction), speech recognition and bioinformatics (such as in rearrangements detection^[79]).

The LZMA lossless data compression algorithm combines Markov chains with Lempel-Ziv compression to achieve very high compression ratios.

Queueing theory

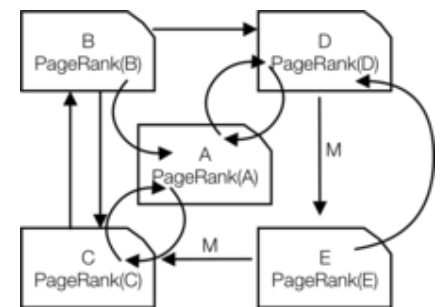
Markov chains are the basis for the analytical treatment of queues (queueing theory). Agner Krarup Erlang initiated the subject in 1917.^[80] This makes them critical for optimizing the performance of telecommunications networks, where messages must often compete for limited resources (such as bandwidth).^[81]

Numerous queueing models use continuous-time Markov chains. For example, an M/M/1 queue is a CTMC on the non-negative integers where upward transitions from i to $i + 1$ occur at rate λ according to a Poisson process and describe job arrivals, while transitions from i to $i - 1$ (for $i > 1$) occur at rate μ (job service times are exponentially distributed) and describe completed services (departures) from the queue.

Internet applications

The PageRank of a webpage as used by Google is defined by a Markov chain.^{[82][83][84]} It is the probability to be at page i in the stationary distribution on the following Markov chain on all (known) webpages. If N is the number of known webpages, and a page i has k_i links to it then it has transition probability $\frac{\alpha}{k_i} + \frac{1 - \alpha}{N}$ for all pages that are linked to and $\frac{1 - \alpha}{N}$ for all pages that are not linked to. The parameter α is taken to be about 0.15.^[85]

Markov models have also been used to analyze web navigation behavior of users. A user's web link transition on a particular website can be modeled using first- or second-order Markov models and can be used to make predictions regarding future navigation and to personalize the web page for an individual user.



A state diagram that represents the PageRank algorithm with a transitional probability of M , or

$$\frac{\alpha}{k_i} + \frac{1 - \alpha}{N}.$$

Statistics

Markov chain methods have also become very important for generating sequences of random numbers to accurately reflect very complicated desired probability distributions, via a process called Markov chain Monte Carlo (MCMC). In recent years this has revolutionized the practicability of Bayesian inference methods, allowing a wide range of posterior distributions to be simulated and their parameters found numerically.

Economics and finance

Markov chains are used in finance and economics to model a variety of different phenomena, including the distribution of income, the size distribution of firms, asset prices and market crashes. D. G. Champernowne built a Markov chain model of the distribution of income in 1953.^[86] Herbert A. Simon and co-author

Charles Bonini used a Markov chain model to derive a stationary Yule distribution of firm sizes.^[87] Louis Bachelier was the first to observe that stock prices followed a random walk.^[88] The random walk was later seen as evidence in favor of the efficient-market hypothesis and random walk models were popular in the literature of the 1960s.^[89] Regime-switching models of business cycles were popularized by James D. Hamilton (1989), who used a Markov chain to model switches between periods of high and low GDP growth (or, alternatively, economic expansions and recessions).^[90] A more recent example is the Markov switching multifractal model of Laurent E. Calvet and Adlai J. Fisher, which builds upon the convenience of earlier regime-switching models.^{[91][92]} It uses an arbitrarily large Markov chain to drive the level of volatility of asset returns.

Dynamic macroeconomics makes heavy use of Markov chains. An example is using Markov chains to exogenously model prices of equity (stock) in a general equilibrium setting.^[93]

Credit rating agencies produce annual tables of the transition probabilities for bonds of different credit ratings.^[94]

Social sciences

Markov chains are generally used in describing path-dependent arguments, where current structural configurations condition future outcomes. An example is the reformulation of the idea, originally due to Karl Marx's *Das Kapital*, tying economic development to the rise of capitalism. In current research, it is common to use a Markov chain to model how once a country reaches a specific level of economic development, the configuration of structural factors, such as size of the middle class, the ratio of urban to rural residence, the rate of political mobilization, etc., will generate a higher probability of transitioning from authoritarian to democratic regime.^[95]

Games

Markov chains can be used to model many games of chance.^[1] The children's games Snakes and Ladders and "Hi Ho! Cherry-O", for example, are represented exactly by Markov chains. At each turn, the player starts in a given state (on a given square) and from there has fixed odds of moving to certain other states (squares).

Music

Markov chains are employed in algorithmic music composition, particularly in software such as Csound, Max, and SuperCollider. In a first-order chain, the states of the system become note or pitch values, and a probability vector for each note is constructed, completing a transition probability matrix (see below). An algorithm is constructed to produce output note values based on the transition matrix weightings, which could be MIDI note values, frequency (Hz), or any other desirable metric.^[96]

1st-order matrix

Note	A	C#	E ♭
A	0.1	0.6	0.3
C#	0.25	0.05	0.7
E ♭	0.7	0.3	0

2nd-order matrix

Notes	A	D	G
AA	0.18	0.6	0.22
AD	0.5	0.5	0
AG	0.15	0.75	0.1
DD	0	0	1
DA	0.25	0	0.75
DG	0.9	0.1	0
GG	0.4	0.4	0.2
GA	0.5	0.25	0.25
GD	1	0	0

A second-order Markov chain can be introduced by considering the current state *and* also the previous state, as indicated in the second table. Higher, *n*th-order chains tend to "group" particular notes together, while 'breaking off' into other patterns and sequences occasionally. These higher-order chains tend to generate results with a sense of phrasal structure, rather than the 'aimless wandering' produced by a first-order system.^[97]

Markov chains can be used structurally, as in Xenakis's Analogique A and B.^[98] Markov chains are also used in systems which use a Markov model to react interactively to music input.^[99]

Usually musical systems need to enforce specific control constraints on the finite-length sequences they generate, but control constraints are not compatible with Markov models, since they induce long-range dependencies that violate the Markov hypothesis of limited memory. In order to overcome this limitation, a new approach has been proposed.^[100]

Baseball

Markov chain models have been used in advanced baseball analysis since 1960, although their use is still rare. Each half-inning of a baseball game fits the Markov chain state when the number of runners and outs are considered. During any at-bat, there are 24 possible combinations of number of outs and position of the runners. Mark Pankin shows that Markov chain models can be used to evaluate runs created for both individual players as well as a team.^[101] He also discusses various kinds of strategies and play conditions: how Markov chain models have been used to analyze statistics for game situations such as bunting and base stealing and differences when playing on grass vs. AstroTurf.^[102]

Markov text generators

Markov processes can also be used to generate superficially real-looking text given a sample document. Markov processes are used in a variety of recreational "parody generator" software (see dissociated press, Jeff Harrison,^[103] Mark V. Shaney,^{[104][105]} and Academias Neutronium). Several open-source text generation libraries using Markov chains exist.

Probabilistic forecasting

Markov chains have been used for forecasting in several areas: for example, price trends,^[106] wind power,^[107] and solar irradiance.^[108] The Markov chain forecasting models utilize a variety of settings, from discretizing the time series,^[107] to hidden Markov models combined with wavelets,^[106] and the Markov chain mixture distribution model (MCM).^[108]

See also

- [Dynamics of Markovian particles](#)
- [Gauss–Markov process](#)
- [Markov chain approximation method](#)
- [Markov chain geostatistics](#)
- [Markov chain mixing time](#)
- [Markov chain tree theorem](#)
- [Markov decision process](#)
- [Markov information source](#)
- [Markov odometer](#)
- [Markov operator](#)
- [Markov random field](#)
- [Master equation](#)
- [Quantum Markov chain](#)
- [Semi-Markov process](#)
- [Stochastic cellular automaton](#)
- [Telescoping Markov chain](#)
- [Variable-order Markov model](#)

Notes

1. Gagniuc, Paul A. (2017). *Markov Chains: From Theory to Implementation and Experimentation*. USA, NJ: John Wiley & Sons. pp. 1–235. ISBN 978-1-119-38755-8.
2. "Markov chain | Definition of Markov chain in US English by Oxford Dictionaries" (https://web.archive.org/web/20171215001435/https://en.oxforddictionaries.com/definition/us/markov_chain). *Oxford Dictionaries*. Archived from the original (https://en.oxforddictionaries.com/definition/us/markov_chain) on December 15, 2017. Retrieved 2017-12-14.
3. Definition at Brilliant.org "Brilliant Math and Science Wiki" (<https://brilliant.org/wiki/markov-chains/>). Retrieved on 12 May 2019
4. Samuel Karlin; Howard E. Taylor (2 December 2012). *A First Course in Stochastic Processes* (<https://books.google.com/books?id=dSDxjX9nmmMC>). Academic Press. p. 47. ISBN 978-0-08-057041-9. Archived (<https://web.archive.org/web/20170323132708/https://books.google.com/books?id=dSDxjX9nmmMC>) from the original on 23 March 2017.
5. Bruce Hajek (12 March 2015). *Random Processes for Engineers* (<https://books.google.com/books?id=Owy0BgAAQBAJ>). Cambridge University Press. ISBN 978-1-316-24124-0. Archived (<https://web.archive.org/web/20170323134957/https://books.google.com/books?id=Owy0BgAAQBAJ>) from the original on 23 March 2017.
6. G. Latouche; V. Ramaswami (1 January 1999). *Introduction to Matrix Analytic Methods in Stochastic Modeling* (<https://books.google.com/books?id=Kan2ki8jqzgC&pg=PR4>). SIAM. pp. 4–. ISBN 978-0-89871-425-8. Archived (<https://web.archive.org/web/20170323180008/https://books.google.com/books?id=Kan2ki8jqzgC&pg=PR4>) from the original on 23 March 2017.
7. Sean Meyn; Richard L. Tweedie (2 April 2009). *Markov Chains and Stochastic Stability* (<https://books.google.com/books?id=Md7RnYEPkJwC>). Cambridge University Press. p. 3. ISBN 978-0-521-73182-9. Archived (<https://web.archive.org/web/20170323131959/https://books.google.com/books?id=Md7RnYEPkJwC>) from the original on 23 March 2017.
8. Reuven Y. Rubinstein; Dirk P. Kroese (20 September 2011). *Simulation and the Monte Carlo Method* (<https://books.google.com/books?id=yWcvT80gQK4C>). John Wiley & Sons. p. 225. ISBN 978-1-118-21052-9. Archived (<https://web.archive.org/web/20170323154112/https://books.google.com/books?id=yWcvT80gQK4C>) from the original on 23 March 2017.
9. Dani Gamerman; Hedibert F. Lopes (10 May 2006). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference, Second Edition* (https://books.google.com/books?id=yPvECi_L3bwC). CRC Press. ISBN 978-1-58488-587-0. Archived (https://web.archive.org/web/20170323161034/https://books.google.com/books?id=yPvECi_L3bwC) from the original on 23 March 2017.

10. "Markovian" (<https://www.oed.com/search/dictionary/?q=Markovian>). *Oxford English Dictionary* (Online ed.). Oxford University Press. (Subscription or participating institution membership (<https://www.oed.com/public/login/loggingin#withyourlibrary>) required.)
11. *Model-Based Signal Processing* (<https://books.google.com/books?id=mNbw1a9WZYUC&dq=%22say+that+a+process+is+markov%22&pg=PA41>). John Wiley & Sons. 27 October 2005. ISBN 9780471732662.
12. Øksendal, B. K. (Bernt Karsten) (2003). *Stochastic differential equations : an introduction with applications* (6th ed.). Berlin: Springer. ISBN 3540047581. OCLC 52203046 (<https://www.worldcat.org/oclc/52203046>).
13. Søren Asmussen (15 May 2003). *Applied Probability and Queues* (<https://books.google.com/books?id=BeYaTxesKy0C>). Springer Science & Business Media. p. 7. ISBN 978-0-387-00211-8. Archived (<https://web.archive.org/web/20170323151525/https://books.google.com/books?id=BeYaTxesKy0C>) from the original on 23 March 2017.
14. Emanuel Parzen (17 June 2015). *Stochastic Processes* (<https://books.google.com/books?id=0mB2CQAAQBAJ>). Courier Dover Publications. p. 188. ISBN 978-0-486-79688-8. Archived (<https://web.archive.org/web/20171120031137/https://books.google.com/books?id=0mB2CQAAQBAJ>) from the original on 20 November 2017.
15. Samuel Karlin; Howard E. Taylor (2 December 2012). *A First Course in Stochastic Processes* (<https://books.google.com/books?id=dSDxjX9nmmMC>). Academic Press. pp. 29 and 30. ISBN 978-0-08-057041-9. Archived (<https://web.archive.org/web/20170323132708/https://books.google.com/books?id=dSDxjX9nmmMC>) from the original on 23 March 2017.
16. John Lamperti (1977). *Stochastic processes: a survey of the mathematical theory* (<https://books.google.com/books?id=Pd4cvgAACAAJ>). Springer-Verlag. pp. 106–121. ISBN 978-3-540-90275-1. Archived (<https://web.archive.org/web/20170323133602/https://books.google.com/books?id=Pd4cvgAACAAJ>) from the original on 2017-03-23.
17. Sheldon M. Ross (1996). *Stochastic processes* (<https://books.google.com/books?id=ImUPAQAAMAAJ>). Wiley. pp. 174 and 231. ISBN 978-0-471-12062-9. Archived (<https://web.archive.org/web/20170323134723/https://books.google.com/books?id=ImUPAQAAMAAJ>) from the original on 2017-03-23.
18. Everitt, B.S. (2002) *The Cambridge Dictionary of Statistics*. CUP. ISBN 0-521-81099-X
19. Parzen, E. (1962) *Stochastic Processes*, Holden-Day. ISBN 0-8162-6664-6 (Table 6.1)
20. Dodge, Y. (2003) *The Oxford Dictionary of Statistical Terms*, OUP. ISBN 0-19-920613-9 (entry for "Markov chain")
21. Dodge, Y. *The Oxford Dictionary of Statistical Terms*, OUP. ISBN 0-19-920613-9
22. Meyn, S. Sean P., and Richard L. Tweedie. (2009) *Markov chains and stochastic stability*. Cambridge University Press. (Preface, p. iii)
23. Gagniuc, Paul A. (2017). *Markov Chains: From Theory to Implementation and Experimentation*. USA, NJ: John Wiley & Sons. pp. 159–163. ISBN 978-1-119-38755-8.
24. Gagniuc, Paul A. (2017). *Markov Chains: From Theory to Implementation and Experimentation*. USA, NJ: John Wiley & Sons. pp. 2–8. ISBN 978-1-119-38755-8.
25. Charles Miller Grinstead; James Laurie Snell (1997). *Introduction to Probability* (<https://archive.org/details/flooved3489>). American Mathematical Soc. pp. 464 (<https://archive.org/details/flooved3489/page/n473>)–466. ISBN 978-0-8218-0749-1.
26. Pierre Bremaud (9 March 2013). *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues* (<https://books.google.com/books?id=jrPVBwAAQBAJ>). Springer Science & Business Media. p. ix. ISBN 978-1-4757-3124-8. Archived (<https://web.archive.org/web/20170323160437/https://books.google.com/books?id=jrPVBwAAQBAJ>) from the original on 23 March 2017.
27. Hayes, Brian (2013). "First links in the Markov chain". *American Scientist*. **101** (2): 92–96. doi:10.1511/2013.101.92 (<https://doi.org/10.1511%2F2013.101.92>).
28. Sheldon M. Ross (1996). *Stochastic processes* (<https://books.google.com/books?id=ImUPAQAAMAAJ>). Wiley. pp. 235 and 358. ISBN 978-0-471-12062-9. Archived (<https://web.archive.org/web/20170323134723/https://books.google.com/books?id=ImUPAQAAMAAJ>) from the original on 2017-03-23.

29. Jarrow, Robert; Protter, Philip (2004). "A short history of stochastic integration and mathematical finance: The early years, 1880–1970". *A Festschrift for Herman Rubin*. Institute of Mathematical Statistics Lecture Notes – Monograph Series. pp. 75–91. CiteSeerX [10.1.1.114.632](https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.114.632) (<https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.114.632>). doi:[10.1214/lnms/1196285381](https://doi.org/10.1214/lnms/1196285381) (<https://doi.org/10.1214/lnms/1196285381>). ISBN 978-0-940600-61-4. ISSN 0749-2170 (<https://www.worldcat.org/issn/0749-2170>).
30. Guttorp, Peter; Thorarinsdottir, Thordis L. (2012). "What Happened to Discrete Chaos, the Quenouille Process, and the Sharp Markov Property? Some History of Stochastic Point Processes". *International Statistical Review*. **80** (2): 253–268. doi:[10.1111/j.1751-5823.2012.00181.x](https://doi.org/10.1111/j.1751-5823.2012.00181.x) (<https://doi.org/10.1111/j.1751-5823.2012.00181.x>). ISSN 0306-7734 (<https://www.worldcat.org/issn/0306-7734>). S2CID [80836](https://api.semanticscholar.org/CorpusID:80836) (<https://api.semanticscholar.org/CorpusID:80836>).
31. Seneta, E. (1996). "Markov and the Birth of Chain Dependence Theory". *International Statistical Review / Revue Internationale de Statistique*. **64** (3): 255–257. doi:[10.2307/1403785](https://doi.org/10.2307/1403785) (<https://doi.org/10.2307/1403785>). ISSN 0306-7734 (<https://www.worldcat.org/issn/0306-7734>). JSTOR [1403785](https://www.jstor.org/stable/1403785) (<https://www.jstor.org/stable/1403785>).
32. Seneta, E. (1998). "I.J. Bienaymé [1796–1878]: Criticality, Inequality, and Internationalization". *International Statistical Review / Revue Internationale de Statistique*. **66** (3): 291–292. doi:[10.2307/1403518](https://doi.org/10.2307/1403518) (<https://doi.org/10.2307/1403518>). ISSN 0306-7734 (<https://www.worldcat.org/issn/0306-7734>). JSTOR [1403518](https://www.jstor.org/stable/1403518) (<https://www.jstor.org/stable/1403518>).
33. Bru B, Hertz S (2001). "Maurice Fréchet". In Heyde CC, Seneta E, Crépel P, Fienberg SE, Gani J (eds.). *Statisticians of the Centuries*. New York, NY: Springer. pp. 331–334. doi:[10.1007/978-1-4613-0179-0_71](https://doi.org/10.1007/978-1-4613-0179-0_71) (https://doi.org/10.1007/978-1-4613-0179-0_71). ISBN 978-0-387-95283-3.
34. Kendall, D. G.; Batchelor, G. K.; Bingham, N. H.; Hayman, W. K.; Hyland, J. M. E.; Lorentz, G. G.; Moffatt, H. K.; Parry, W.; Razborov, A. A.; Robinson, C. A.; Whittle, P. (1990). "Andrei Nikolaevich Kolmogorov (1903–1987)". *Bulletin of the London Mathematical Society*. **22** (1): 33. doi:[10.1112/blms/22.1.31](https://doi.org/10.1112/blms/22.1.31) (<https://doi.org/10.1112/blms/22.1.31>). ISSN 0024-6093 (<https://www.worldcat.org/issn/0024-6093>).
35. Cramér, Harald (1976). "Half a Century with Probability Theory: Some Personal Recollections" (<https://doi.org/10.1214/aop/1176996025>). *The Annals of Probability*. **4** (4): 509–546. doi:[10.1214/aop/1176996025](https://doi.org/10.1214/aop/1176996025) (<https://doi.org/10.1214/aop/1176996025>). ISSN 0091-1798 (<https://www.worldcat.org/issn/0091-1798>).
36. Marc Barbut; Bernard Locker; Laurent Mazliak (23 August 2016). *Paul Lévy and Maurice Fréchet: 50 Years of Correspondence in 107 Letters* (https://books.google.com/books?id=ISz_vQAACAAJ). Springer London. p. 5. ISBN 978-1-4471-7262-8. Archived (https://web.archive.org/web/20170323153347/https://books.google.com/books?id=ISz_vQAACAAJ) from the original on 23 March 2017.
37. Valeriy Skorokhod (5 December 2005). *Basic Principles and Applications of Probability Theory* (<https://books.google.com/books?id=dQkYMjRK3fYC>). Springer Science & Business Media. p. 146. ISBN 978-3-540-26312-8. Archived (<https://web.archive.org/web/20170323172037/https://books.google.com/books?id=dQkYMjRK3fYC>) from the original on 23 March 2017.
38. Bernstein, Jeremy (2005). "Bachelier". *American Journal of Physics*. **73** (5): 395–398. Bibcode:2005AmJPh..73..395B (<https://ui.adsabs.harvard.edu/abs/2005AmJPh..73..395B>). doi:[10.1119/1.1848117](https://doi.org/10.1119/1.1848117) (<https://doi.org/10.1119/1.1848117>). ISSN 0002-9505 (<https://www.worldcat.org/issn/0002-9505>).
39. William J. Anderson (6 December 2012). *Continuous-Time Markov Chains: An Applications-Oriented Approach* (<https://books.google.com/books?id=YpHfBwAAQBAJ&pg=PR8>). Springer Science & Business Media. p. vii. ISBN 978-1-4612-3038-0. Archived (<https://web.archive.org/web/20170323151730/https://books.google.com/books?id=YpHfBwAAQBAJ&pg=PR8>) from the original on 23 March 2017.
40. Kendall, D. G.; Batchelor, G. K.; Bingham, N. H.; Hayman, W. K.; Hyland, J. M. E.; Lorentz, G. G.; Moffatt, H. K.; Parry, W.; Razborov, A. A.; Robinson, C. A.; Whittle, P. (1990). "Andrei Nikolaevich Kolmogorov (1903–1987)". *Bulletin of the London Mathematical Society*. **22** (1): 57. doi:[10.1112/blms/22.1.31](https://doi.org/10.1112/blms/22.1.31) (<https://doi.org/10.1112/blms/22.1.31>). ISSN 0024-6093 (<https://www.worldcat.org/issn/0024-6093>).

41. Ionut Florescu (7 November 2014). *Probability and Stochastic Processes* (<https://books.google.com/books?id=Z5xEBQAAQBAJ&pg=PR22>). John Wiley & Sons. pp. 373 and 374. ISBN 978-1-118-59320-2. Archived (<https://web.archive.org/web/20170323153417/https://books.google.com/books?id=Z5xEBQAAQBAJ&pg=PR22>) from the original on 23 March 2017.
42. Samuel Karlin; Howard E. Taylor (2 December 2012). *A First Course in Stochastic Processes* (<https://books.google.com/books?id=dSDxjX9nmmMC>). Academic Press. p. 49. ISBN 978-0-08-057041-9. Archived (<https://web.archive.org/web/20170323132708/https://books.google.com/books?id=dSDxjX9nmmMC>) from the original on 23 March 2017.
43. Gagniuc, Paul A. (2017). *Markov Chains: From Theory to Implementation and Experimentation*. USA, NJ: John Wiley & Sons. pp. 1–2. ISBN 978-1-119-38755-8.
44. Weiss, George H. (2006). "Random Walks". *Encyclopedia of Statistical Sciences*. p. 1. doi:10.1002/0471667196.ess2180.pub2 (<https://doi.org/10.1002%2F0471667196.ess2180.pub2>). ISBN 978-0471667193.
45. Michael F. Shlesinger (1985). *The Wonderful world of stochastics: a tribute to Elliott W. Montroll* (<https://books.google.com/books?id=p6fvAAAAMAAJ>). North-Holland. pp. 8–10. ISBN 978-0-444-86937-1. Archived (<https://web.archive.org/web/20170323175853/https://books.google.com/books?id=p6fvAAAAMAAJ>) from the original on 2017-03-23.
46. Emanuel Parzen (17 June 2015). *Stochastic Processes* (<https://books.google.com/books?id=0mB2CQAQBAJ>). Courier Dover Publications. p. 7 and 8. ISBN 978-0-486-79688-8. Archived (<https://web.archive.org/web/20171120031137/https://books.google.com/books?id=0mB2CQAQBAJ>) from the original on 20 November 2017.
47. Joseph L. Doob (1990). *Stochastic processes* (<https://books.google.com/books?id=7Bu8jgEACAAJ>). Wiley. p. 46 and 47. Archived (<https://web.archive.org/web/20171120031137/https://books.google.com/books?id=7Bu8jgEACAAJ>) from the original on 2017-11-20.
48. Donald L. Snyder; Michael I. Miller (6 December 2012). *Random Point Processes in Time and Space* (https://books.google.com/books?id=c_3UBwAAQBAJ). Springer Science & Business Media. p. 32. ISBN 978-1-4612-3166-0. Archived (https://web.archive.org/web/20171120031137/https://books.google.com/books?id=c_3UBwAAQBAJ) from the original on 20 November 2017.
49. Norris, J. R. (1997). "Continuous-time Markov chains I". *Markov Chains*. pp. 60–107. doi:10.1017/CBO9780511810633.004 (<https://doi.org/10.1017%2FCBO9780511810633.004>). ISBN 9780511810633.
50. Serfozo, Richard (2009). "Basics of Applied Stochastic Processes" (https://archive.org/details/basicsappliedsto00serf_111). *Probability and Its Applications*. doi:10.1007/978-3-540-89332-5 (<https://doi.org/10.1007%2F978-3-540-89332-5>). ISBN 978-3-540-89331-8. ISSN 1431-7028 (<https://www.worldcat.org/issn/1431-7028>).
51. "Chapter 11 "Markov Chains" " (https://www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/Chapter11.pdf) (PDF). Archived (https://web.archive.org/web/20170215053623/http://www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/Chapter11.pdf) (PDF) from the original on 2017-02-15. Retrieved 2017-06-02.
52. Schmitt, Florian; Rothlauf, Franz (2001). "On the Importance of the Second Largest Eigenvalue on the Convergence Rate of Genetic Algorithms". *Proceedings of the 14th Symposium on Reliable Distributed Systems*. CiteSeerX 10.1.1.28.6191 (<https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.28.6191>).
53. Rosenthal, Jeffrey S. (1995). "Convergence Rates for Markov Chains" (<https://www.jstor.org/stable/2132659>). *SIAM Review*. **37** (3): 387–405. doi:10.1137/1037083 (<https://doi.org/10.1137%2F1037083>). ISSN 0036-1445 (<https://www.worldcat.org/issn/0036-1445>). JSTOR 2132659 (<https://www.jstor.org/stable/2132659>). Retrieved 2021-05-31.
54. Franzke, Brandon; Kosko, Bart (1 October 2011). "Noise can speed convergence in Markov chains". *Physical Review E*. **84** (4): 041112. Bibcode:2011PhRvE..84d1112F (<https://ui.adsabs.harvard.edu/abs/2011PhRvE..84d1112F>). doi:10.1103/PhysRevE.84.041112 (<https://doi.org/10.1103%2FPhysRevE.84.041112>). PMID 22181092 (<https://pubmed.ncbi.nlm.nih.gov/22181092>).

55. Spitzer, Frank (1970). "Interaction of Markov Processes" (<https://doi.org/10.1016%2F0001-8708%2870%2990034-4>). *Advances in Mathematics*. **5** (2): 246–290. doi:10.1016/0001-8708(70)90034-4 (<https://doi.org/10.1016%2F0001-8708%2870%2990034-4>).
56. Dobrushin, R. L.; Kryukov, V. I.; Toom, A. L. (1978). *Stochastic Cellular Systems: Ergodicity, Memory, Morphogenesis* (<https://books.google.com/books?id=0Wa7AAAAIAAJ&q=locally+interacting+markov+chains+toom+Dobrushin&pg=PA181>). ISBN 9780719022067. Archived (<https://web.archive.org/web/20170205211328/https://books.google.com/books?id=0Wa7AAAAIAAJ&pg=PA181&lpg=PA181&dq=locally+interacting+markov+chains+toom+Dobrushin#v=onepage&q=locally%20interacting%20markov%20chains%20toom%20Dobrushin&f=false>) from the original on 2017-02-05. Retrieved 2016-03-04.
57. Heyman, Daniel P.; Sobel, Mathew J. (1982). *Stochastic Models in Operations Research, Volume 1*. New York: McGraw-Hill. p. 230. ISBN 0-07-028631-0.
58. Parzen, Emanuel (1962). *Stochastic Processes*. San Francisco: Holden-Day. p. 145. ISBN 0-8162-6664-6.
59. Doblinger, G. (September 1998). "Smoothing of noisy AR signals using an adaptive Kalman filter" (https://publik.tuwien.ac.at/files/pub-et_3285.pdf) (PDF). *9th European Signal Processing Conference (EUSIPCO 1998)*: 781–784.
60. Norris, J. R. (1997). "Continuous-time Markov chains II". *Markov Chains*. pp. 108–127. doi:10.1017/CBO9780511810633.005 (<https://doi.org/10.1017%2FCBO9780511810633.005>). ISBN 9780511810633.
61. Matthew Nicol and Karl Petersen, (2009) "Ergodic Theory: Basic Examples and Constructions (<https://www.math.uh.edu/~nicol/pdf/petersen.pdf>)", *Encyclopedia of Complexity and Systems Science*, Springer https://doi.org/10.1007/978-0-387-30440-3_177
62. Fitzpatrick, Richard. "Thermodynamics and Statistical Mechanics" (<https://farside.ph.utexas.edu/teaching/sm1/Thermal.pdf>) (PDF). Archived (<https://web.archive.org/web/20161130172856/http://farside.ph.utexas.edu/teaching/sm1/Thermal.pdf>) (PDF) from the original on 2016-11-30. Retrieved 2017-06-02.
63. van Ravenzwaaij, Don; Cassey, Pete; Brown, Scott D. (2016-03-11). "A simple introduction to Markov Chain Monte–Carlo sampling" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5862921>). *Psychonomic Bulletin & Review*. **25** (1): 143–154. doi:10.3758/s13423-016-1015-8 (<https://doi.org/10.3758%2Fs13423-016-1015-8>). ISSN 1069-9384 (<https://www.worldcat.org/issn/1069-9384>). PMC 5862921 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5862921>). PMID 26968853 (<https://pubmed.ncbi.nlm.nih.gov/26968853>).
64. Ryder, Lewis H. (1985). *Quantum field theory* (<https://archive.org/details/quantumfieldtheo0000ryde>). Cambridge [Cambridgeshire]: Cambridge University Press. pp. 160 (<https://archive.org/details/quantumfieldtheo0000ryde/page/160>). ISBN 978-0521338592. OCLC 10533049 (<https://www.worldcat.org/oclc/10533049>).
65. Gattringer, Christof; Lang, Christian B (2010). *Quantum Chromodynamics on the Lattice* (<https://www.springer.com/gb/book/9783642018497>). Lecture Notes in Physics. Vol. 788. Springer-Verlag Berlin Heidelberg. doi:10.1007/978-3-642-01850-3 (<https://doi.org/10.1007%2F978-3-642-01850-3>). ISBN 978-3-642-01849-7. Archived (<https://web.archive.org/web/20170801154439/http://www.springer.com/gb/book/9783642018497>) from the original on 2017-08-01.
66. Anderson, David F.; Kurtz, Thomas G. (2011), "Continuous Time Markov Chain Models for Chemical Reaction Networks", *Design and Analysis of Biomolecular Circuits*, Springer New York, pp. 3–42, doi:10.1007/978-1-4419-6766-4_1 (https://doi.org/10.1007%2F978-1-4419-6766-4_1), ISBN 9781441967657
67. Du, Chao; Kou, S. C. (September 2012). "Correlation analysis of enzymatic reaction of a single protein molecule" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3568780>). *The Annals of Applied Statistics*. **6** (3): 950–976. arXiv:1209.6210 (<https://arxiv.org/abs/1209.6210>). Bibcode:2012arXiv1209.6210D (<https://ui.adsabs.harvard.edu/abs/2012arXiv1209.6210D>). doi:10.1214/12-aos541 (<https://doi.org/10.1214%2F12-aos541>). ISSN 1932-6157 (<https://www.worldcat.org/issn/1932-6157>). PMC 3568780 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3568780>). PMID 23408514 (<https://pubmed.ncbi.nlm.nih.gov/23408514>).

68. Kutchukian, Peter; Lou, David; Shakhnovich, Eugene (2009). "FOG: Fragment Optimized Growth Algorithm for the de Novo Generation of Molecules occupying Druglike Chemical". *Journal of Chemical Information and Modeling*. **49** (7): 1630–1642. doi:10.1021/ci9000458 (https://doi.org/10.1021%2Fci9000458). PMID 19527020 (https://pubmed.ncbi.nlm.nih.gov/19527020).
69. Kutchukian, Peter S.; Lou, David; Shakhnovich, Eugene I. (2009-06-15). "FOG: Fragment Optimized Growth Algorithm for the de Novo Generation of Molecules Occupying Druglike Chemical Space". *Journal of Chemical Information and Modeling*. **49** (7): 1630–1642. doi:10.1021/ci9000458 (https://doi.org/10.1021%2Fci9000458). ISSN 1549-9596 (https://www.worldcat.org/issn/1549-9596). PMID 19527020 (https://pubmed.ncbi.nlm.nih.gov/19527020).
70. Kopp, V. S.; Kaganer, V. M.; Schwarzkopf, J.; Waidick, F.; Remmele, T.; Kwasniewski, A.; Schmidbauer, M. (2011). "X-ray diffraction from nonperiodic layered structures with correlations: Analytical calculation and experiment on mixed Aurivillius films". *Acta Crystallographica Section A*. **68** (Pt 1): 148–155. Bibcode:2012AcCrA..68..148K (https://ui.adsabs.harvard.edu/abs/2012AcCrA..68..148K). doi:10.1107/S0108767311044874 (https://doi.org/10.1107%2FS0108767311044874). PMID 22186291 (https://pubmed.ncbi.nlm.nih.gov/22186291).
71. George, Dileep; Hawkins, Jeff (2009). Friston, Karl J. (ed.). "Towards a Mathematical Theory of Cortical Micro-circuits" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2749218). *PLOS Comput Biol*. **5** (10): e1000532. Bibcode:2009PLSCB...5E0532G (https://ui.adsabs.harvard.edu/abs/2009PLSCB...5E0532G). doi:10.1371/journal.pcbi.1000532 (https://doi.org/10.1371%2Fjournal.pcbi.1000532). PMC 2749218 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2749218). PMID 19816557 (https://pubmed.ncbi.nlm.nih.gov/19816557).
72. Gupta, Ankur; Rawlings, James B. (April 2014). "Comparison of Parameter Estimation Methods in Stochastic Chemical Kinetic Models: Examples in Systems Biology" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4946376). *AIChE Journal*. **60** (4): 1253–1268. doi:10.1002/aic.14409 (https://doi.org/10.1002%2Faic.14409). PMC 4946376 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4946376). PMID 27429455 (https://pubmed.ncbi.nlm.nih.gov/27429455).
73. Aguiar, R. J.; Collares-Pereira, M.; Conde, J. P. (1988). "Simple procedure for generating sequences of daily radiation values using a library of Markov transition matrices". *Solar Energy*. **40** (3): 269–279. Bibcode:1988SoEn...40..269A (https://ui.adsabs.harvard.edu/abs/1988SoEn...40..269A). doi:10.1016/0038-092X(88)90049-7 (https://doi.org/10.1016%2F0038-092X%2888%2990049-7).
74. Ngoko, B. O.; Sugihara, H.; Funaki, T. (2014). "Synthetic generation of high temporal resolution solar radiation data using Markov models". *Solar Energy*. **103**: 160–170. Bibcode:2014SoEn..103..160N (https://ui.adsabs.harvard.edu/abs/2014SoEn..103..160N). doi:10.1016/j.solener.2014.02.026 (https://doi.org/10.1016%2Fj.solener.2014.02.026).
75. Bright, J. M.; Smith, C. I.; Taylor, P. G.; Crook, R. (2015). "Stochastic generation of synthetic minutely irradiance time series derived from mean hourly weather observation data" (https://doi.org/10.1016%2Fj.solener.2015.02.032). *Solar Energy*. **115**: 229–242. Bibcode:2015SoEn..115..229B (https://ui.adsabs.harvard.edu/abs/2015SoEn..115..229B). doi:10.1016/j.solener.2015.02.032 (https://doi.org/10.1016%2Fj.solener.2015.02.032).
76. Munkhammar, J.; Widén, J. (2018). "An N-state Markov-chain mixture distribution model of the clear-sky index". *Solar Energy*. **173**: 487–495. Bibcode:2018SoEn..173..487M (https://ui.adsabs.harvard.edu/abs/2018SoEn..173..487M). doi:10.1016/j.solener.2018.07.056 (https://doi.org/10.1016%2Fj.solener.2018.07.056). S2CID 125538244 (https://api.semanticscholar.org/CorpusID:125538244).
77. Morf, H. (1998). "The stochastic two-state solar irradiance model (STSIM)". *Solar Energy*. **62** (2): 101–112. Bibcode:1998SoEn...62..101M (https://ui.adsabs.harvard.edu/abs/1998SoEn...62..101M). doi:10.1016/S0038-092X(98)00004-8 (https://doi.org/10.1016%2FS0038-092X%2898%2900004-8).
78. Munkhammar, J.; Widén, J. (2018). "A Markov-chain probability distribution mixture approach to the clear-sky index". *Solar Energy*. **170**: 174–183. Bibcode:2018SoEn..170..174M (https://ui.adsabs.harvard.edu/abs/2018SoEn..170..174M). doi:10.1016/j.solener.2018.05.055 (https://doi.org/10.1016%2Fj.solener.2018.05.055). S2CID 125867684 (https://api.semanticscholar.org/CorpusID:125867684).

79. Pratas, D; Silva, R; Pinho, A; Ferreira, P (May 18, 2015). "An alignment-free method to find and visualise rearrangements between pairs of DNA sequences" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4434998>). *Scientific Reports*. **5** (10203): 10203. Bibcode:2015NatSR...510203P (https://ui.adsabs.harvard.edu/abs/2015NatSR...510203P). doi:10.1038/srep10203 (https://doi.org/10.1038%2Fsrep10203). PMC 4434998 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4434998). PMID 25984837 (https://pubmed.ncbi.nlm.nih.gov/25984837).
80. O'Connor, John J.; Robertson, Edmund F., "Markov chain" (<https://mathshistory.st-andrews.ac.uk/Biographies/Erlang.html>), *MacTutor History of Mathematics Archive*, University of St Andrews
81. S. P. Meyn, 2007. *Control Techniques for Complex Networks* (http://www.meyn.ece.ufl.edu/archive/spm_files/CTCN/MonographTocBib.pdf) Archived (https://web.archive.org/web/20150513155013/http://www.meyn.ece.ufl.edu/archive/spm_files/CTCN/MonographTocBib.pdf) 2015-05-13 at the *Wayback Machine*, Cambridge University Press, 2007.
82. U.S. Patent 6,285,999 (<https://patents.google.com/patent/US6285999>)
83. Gupta, Brij; Agrawal, Dharma P.; Yamaguchi, Shingo (16 May 2016). *Handbook of Research on Modern Cryptographic Solutions for Computer and Cyber Security* (<https://books.google.com/books?id=Ctk6DAAQBAJ&pg=PA448>). IGI Global. pp. 448–. ISBN 978-1-5225-0106-0. Archived (<https://web.archive.org/web/20170323150736/https://books.google.com/books?id=Ctk6DAAQBAJ&pg=PA448>) from the original on 23 March 2017.
84. Langville, Amy N.; Meyer, Carl D. (2006). "A Reordering for the PageRank Problem" (http://meyer.math.ncsu.edu/Meyer/PS_Files/ReorderingPageRank.pdf) (PDF). *SIAM Journal on Scientific Computing*. **27** (6): 2112–2113. Bibcode:2006SJSC...27.2112L (https://ui.adsabs.harvard.edu/abs/2006SJSC...27.2112L). CiteSeerX 10.1.1.58.8652 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.58.8652). doi:10.1137/040607551 (https://doi.org/10.1137%2F040607551). ISSN 1064-8275 (https://www.worldcat.org/issn/1064-8275). Archived (https://web.archive.org/web/20170921211826/http://meyer.math.ncsu.edu/Meyer/PS_Files/ReorderingPageRank.pdf) (PDF) from the original on 2017-09-21.
85. Page, Lawrence; Brin, Sergey; Motwani, Rajeev; Winograd, Terry (1999). *The PageRank Citation Ranking: Bringing Order to the Web* (Technical report). CiteSeerX 10.1.1.31.1768 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.31.1768).
86. Champernowne, D (1953). "A model of income distribution". *The Economic Journal*. **63** (250): 318–51. doi:10.2307/2227127 (https://doi.org/10.2307%2F2227127). JSTOR 2227127 (https://www.jstor.org/stable/2227127).
87. Simon, Herbert; C Bonini (1958). "The size distribution of business firms". *Am. Econ. Rev.* **42**: 425–40.
88. Bachelier, Louis (1900). "Théorie de la spéculation". *Annales Scientifiques de l'École Normale Supérieure*. **3**: 21–86. doi:10.24033/asens.476 (https://doi.org/10.24033%2Fasens.476). hdl:2027/coo.31924001082803 (https://hdl.handle.net/2027%2Fcoo.31924001082803).
89. e.g.Fama, E (1965). "The behavior of stock market prices". *Journal of Business*. **38**.
90. Hamilton, James (1989). "A new approach to the economic analysis of nonstationary time series and the business cycle". *Econometrica*. **57** (2): 357–84. CiteSeerX 10.1.1.397.3582 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.397.3582). doi:10.2307/1912559 (https://doi.org/10.2307%2F1912559). JSTOR 1912559 (https://www.jstor.org/stable/1912559).
91. Calvet, Laurent E.; Fisher, Adlai J. (2001). "Forecasting Multifractal Volatility" (<http://archive.nyu.edu/handle/2451/26894>). *Journal of Econometrics*. **105** (1): 27–58. doi:10.1016/S0304-4076(01)00069-0 (https://doi.org/10.1016%2FS0304-4076%2801%2900069-0). S2CID 119394176 (https://api.semanticscholar.org/CorpusID:119394176).
92. Calvet, Laurent; Adlai Fisher (2004). "How to Forecast long-run volatility: regime-switching and the estimation of multifractal processes". *Journal of Financial Econometrics*. **2**: 49–83. CiteSeerX 10.1.1.536.8334 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.536.8334). doi:10.1093/jfinec/nbh003 (https://doi.org/10.1093%2Fjfinec%2Fnbh003). S2CID 6020401 (https://api.semanticscholar.org/CorpusID:6020401).

93. Brennan, Michael; Xiab, Yihong. "Stock Price Volatility and the Equity Premium" (<https://web.archive.org/web/20081228200849/http://bbs.cenet.org.cn/uploadImages/200352118122167693.pdf>) (PDF). *Department of Finance, the Anderson School of Management, UCLA*. Archived from the original (<http://bbs.cenet.org.cn/uploadImages/200352118122167693.pdf>) (PDF) on 2008-12-28.
94. "A Markov Chain Example in Credit Risk Modelling Columbia University lectures" (https://web.archive.org/web/20160324112501/http://www.columbia.edu/~ww2040/4106S11/MC_BondRating.pdf) (PDF). *Columbia University*. Archived from the original (http://www.columbia.edu/~ww2040/4106S11/MC_BondRating.pdf) (PDF) on March 24, 2016.
95. Acemoglu, Daron; Georgy Egorov; Konstantin Sonin (2011). "Political model of social evolution" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3271566>). *Proceedings of the National Academy of Sciences*. **108** (Suppl 4): 21292–21296. Bibcode:2011PNAS..10821292A (<https://ui.adsabs.harvard.edu/abs/2011PNAS..10821292A>). CiteSeerX 10.1.1.225.6090 (<https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.225.6090>). doi:10.1073/pnas.1019454108 (<https://doi.org/10.1073/pnas.1019454108>). PMC 3271566 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3271566>). PMID 22198760 (<https://pubmed.ncbi.nlm.nih.gov/22198760>).
96. K McAlpine; E Miranda; S Hoggar (1999). "Making Music with Algorithms: A Case-Study System". *Computer Music Journal*. **23** (2): 19–30. doi:10.1162/014892699559733 (<https://doi.org/10.1162/014892699559733>). S2CID 40057162 (<https://api.semanticscholar.org/CorpusID:40057162>).
97. Curtis Roads, ed. (1996). *The Computer Music Tutorial*. MIT Press. ISBN 978-0-262-18158-7.
98. Xenakis, Iannis; Kanach, Sharon (1992) *Formalized Music: Mathematics and Thought in Composition*, Pendragon Press. ISBN 1576470792
99. "Continuator" (<https://web.archive.org/web/20120713235933/http://www.csl.sony.fr/~pachet/>). Archived from the original (<http://www.csl.sony.fr/~pachet/>) on July 13, 2012.
100. Pachet, F.; Roy, P.; Barbieri, G. (2011) "Finite-Length Markov Processes with Constraints" (<http://www.csl.sony.fr/downloads/papers/2011/pachet-11b.pdf>) Archived (<https://web.archive.org/web/20120414183247/http://www.csl.sony.fr/downloads/papers/2011/pachet-11b.pdf>) 2012-04-14 at the Wayback Machine, *Proceedings of the 22nd International Joint Conference on Artificial Intelligence, IJCAI*, pages 635–642, Barcelona, Spain, July 2011
101. Pankin, Mark D. "MARKOV CHAIN MODELS: THEORETICAL BACKGROUND" (<https://web.archive.org/web/20071209122054/http://www.pankin.com/markov/theory.htm>). Archived from the original (<http://www.pankin.com/markov/theory.htm>) on 2007-12-09. Retrieved 2007-11-26.
102. Pankin, Mark D. "BASEBALL AS A MARKOV CHAIN" (<http://www.pankin.com/markov/intro.htm>). Archived (<https://web.archive.org/web/20090513011104/http://www.pankin.com/markov/intro.htm>) from the original on 2009-05-13. Retrieved 2009-04-24.
103. "Poet's Corner – Fieralingue" (https://web.archive.org/web/20101206043430/http://www.fieralingue.it/modules.php?name=Content&pa=list_pages_categories&cid=111). Archived from the original (http://www.fieralingue.it/modules.php?name=Content&pa=list_pages_categories&cid=111) on December 6, 2010.
104. Kenner, Hugh; O'Rourke, Joseph (November 1984). "A Travesty Generator for Micros". *BYTE*. **9** (12): 129–131, 449–469.
105. Hartman, Charles (1996). *Virtual Muse: Experiments in Computer Poetry* (<https://archive.org/details/virtuallmuseexper00hart>). Hanover, NH: Wesleyan University Press. ISBN 978-0-8195-2239-9.
106. de Souza e Silva, E.G.; Legey, L.F.L.; de Souza e Silva, E.A. (2010). "Forecasting oil price trends using wavelets and hidden Markov models" (<https://www.sciencedirect.com/science/article/pii/S0140988310001271>). *Energy Economics*. **32**.
107. Carpinone, A; Giorgio, M; Langella, R.; Testa, A. (2015). "Markov chain modeling for very-short-term wind power forecasting" (<https://doi.org/10.1016/j.epsr.2014.12.025>). *Electric Power Systems Research*. **122**: 152–158. doi:10.1016/j.epsr.2014.12.025 (<https://doi.org/10.1016/j.epsr.2014.12.025>).
108. Munkhammar, J.; van der Meer, D.W.; Widén, J. (2019). "Probabilistic forecasting of high-resolution clear-sky index time-series using a Markov-chain mixture distribution model". *Solar Energy*. **184**: 688–695. Bibcode:2019SoEn..184..688M (<https://ui.adsabs.harvard.edu/abs/2019SoEn..184..688M>). doi:10.1016/j.solener.2019.04.014 (<https://doi.org/10.1016/j.solener.2019.04.014>). S2CID 146076100 (<https://api.semanticscholar.org/CorpusID:146076100>).

References

- A. A. Markov (1906) "Распространение закона бол'ших чisel на величины, зависящие друг от друга". *Izvestiya Fiziko-matematicheskogo obschestva pri Kazanskom universitete*, 2-ya seriya, tom 15, pp. 135–156.
- A. A. Markov (1971). "Extension of the limit theorems of probability theory to a sum of variables connected in a chain". reprinted in Appendix B of: R. Howard. *Dynamic Probabilistic Systems, volume 1: Markov Chains*. John Wiley and Sons.
- Classical Text in Translation: Markov, A. A. (2006). Translated by Link, David. "An Example of Statistical Investigation of the Text Eugene Onegin Concerning the Connection of Samples in Chains" (<https://semanticscholar.org/paper/343829c84e1f9b3fcf6f0ff2d2a1c680426faaaf>). *Science in Context*. **19** (4): 591–600. doi:10.1017/s0269889706001074 (<https://doi.org/10.1017%2Fs0269889706001074>). S2CID 144854176 (<https://api.semanticscholar.org/CorpusID:144854176>).
- Leo Breiman (1992) [1968] *Probability*. Original edition published by Addison-Wesley; reprinted by Society for Industrial and Applied Mathematics ISBN 0-89871-296-3. (See Chapter 7)
- J. L. Doob (1953) *Stochastic Processes*. New York: John Wiley and Sons ISBN 0-471-52369-0.
- S. P. Meyn and R. L. Tweedie (1993) *Markov Chains and Stochastic Stability*. London: Springer-Verlag ISBN 0-387-19832-6. online: MCSS (https://web.archive.org/web/20100619010320/https://netfiles.uiuc.edu/meyn/www/spm_files/book.html) . Second edition to appear, Cambridge University Press, 2009.
- Dynkin, Eugene Borisovich (1965) [1963-03-10, 1962-03-31]. Written at University of Moscow, Moscow, Russia. *Markov Processes-I* (<https://archive.org/details/markovprocesses0001dynk>). Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen mit besonderer Berücksichtigung der Anwendungsgebiete. Vol. I (121). Translated by Fabius, Jaap [at Wikidata]; Greenberg, Vida Lazarus [at Wikidata]; Maitra, Ashok Prasad [at Wikidata]; Majone, Giandomenico (1 ed.). New York, USA / Berlin, Germany: Springer-Verlag (Academic Press, Inc., Publishers). doi:10.1007/978-3-662-00031-1 (<https://doi.org/10.1007%2F978-3-662-00031-1>). ISBN 978-3-662-00033-5. ISSN 0072-7830 (<https://www.worldcat.org/issn/0072-7830>). LCCN 64-24812 (<https://lccn.loc.gov/64-24812>). Title-No. 5104. Retrieved 2023-09-02. [1] (<https://books.google.com/books?id=vHrpCAAQBAJ>) (xii+365+1 pages); Dynkin, Eugene Borisovich (1965) [1963-03-10, 1962-03-31]. Written at University of Moscow, Moscow, Russia. *Markov Processes-II* (<https://archive.org/details/markovprocesses0002dynk>). Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen mit besonderer Berücksichtigung der Anwendungsgebiete. Vol. II (122). Translated by Fabius, Jaap [at Wikidata]; Greenberg, Vida Lazarus [at Wikidata]; Maitra, Ashok Prasad [at Wikidata]; Majone, Giandomenico (1 ed.). New York, USA / Berlin, Germany: Springer-Verlag. doi:10.1007/978-3-662-25360-1 (<https://doi.org/10.1007%2F978-3-662-25360-1>). ISBN 978-3-662-23320-7. ISSN 0072-7830 (<https://www.worldcat.org/issn/0072-7830>). LCCN 64-24812 (<https://lccn.loc.gov/64-24812>). Title-No. 5105. Retrieved 2023-09-02. (viii+274+2 pages) (NB. This was originally published in Russian as "Марковские процессы" (Марковские процессы) by Fizmatgiz (Физматгиз) in 1963 and translated to English with the assistance of the author.)
- S. P. Meyn. *Control Techniques for Complex Networks*. Cambridge University Press, 2007. ISBN 978-0-521-88441-9. Appendix contains abridged Meyn & Tweedie. online: [2] (https://web.archive.org/web/20100619011046/https://netfiles.uiuc.edu/meyn/www/spm_files/CTCN/CTCN.html)
- Booth, Taylor L. (1967). *Sequential Machines and Automata Theory* (1st ed.). New York, NY: John Wiley and Sons, Inc. Library of Congress Card Catalog Number 67-25924.] Extensive, wide-ranging book meant for specialists, written for both theoretical computer scientists as well as electrical engineers. With detailed explanations of state minimization techniques, FSMs, Turing machines, Markov processes, and undecidability. Excellent treatment of Markov processes pp. 449ff. Discusses Z-transforms, D transforms in their context.
- Kemeny, John G.; Hazleton Mirkil; J. Laurie Snell; Gerald L. Thompson (1959). *Finite Mathematical Structures* (https://archive.org/details/finitemathematic0000keme_h5g0) (1st ed.). Englewood Cliffs, NJ: Prentice-Hall, Inc. Library of Congress Card Catalog Number 59-12841. Classical text. cf Chapter 6 *Finite Markov Chains* pp. 384ff.
- John G. Kemeny & J. Laurie Snell (1960) *Finite Markov Chains*, D. van Nostrand Company ISBN 0-442-04328-7

- E. Nummelin. "General irreducible Markov chains and non-negative operators". Cambridge University Press, 1984, 2004. [ISBN 0-521-60494-X](#)
- Seneta, E. *Non-negative matrices and Markov chains*. 2nd rev. ed., 1981, XVI, 288 p., Softcover Springer Series in Statistics. (Originally published by Allen & Unwin Ltd., London, 1973) [ISBN 978-0-387-29765-1](#)
- Kishor S. Trivedi, *Probability and Statistics with Reliability, Queueing, and Computer Science Applications*, John Wiley & Sons, Inc. New York, 2002. [ISBN 0-471-33341-7](#).
- K. S. Trivedi and R.A.Sahner, *SHARPE at the age of twenty-two*, vol. 36, no. 4, pp. 52–57, ACM SIGMETRICS Performance Evaluation Review, 2009.
- R. A. Sahner, K. S. Trivedi and A. Puliafito, *Performance and reliability analysis of computer systems: an example-based approach using the SHARPE software package*, Kluwer Academic Publishers, 1996. [ISBN 0-7923-9650-2](#).
- G. Bolch, S. Greiner, H. de Meer and K. S. Trivedi, *Queueing Networks and Markov Chains*, John Wiley, 2nd edition, 2006. [ISBN 978-0-7923-9650-5](#).

External links

- Introduction to Markov Chains (https://www.youtube.com/watch?v=o-jdJxXL_W4) on YouTube
 - "Markov chain" (https://www.encyclopediaofmath.org/index.php?title=Markov_chain), *Encyclopedia of Mathematics*, EMS Press, 2001 [1994]
 - Techniques to Understand Computer Simulations: Markov Chain Analysis (<http://jasss.soc.surrey.ac.uk/12/1/6.html>)
 - Markov Chains chapter in American Mathematical Society's introductory probability book (http://www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/Chapter11.pdf) Archived (https://web.archive.org/web/20080522131917/http://www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/Chapter11.pdf) 2008-05-22 at the Wayback Machine(pdf)
 - A beautiful visual explanation of Markov Chains (<http://setosa.io/blog/2014/07/26/markov-chains/index.html>)
 - Making Sense and Nonsense of Markov Chains (<http://rarlindseysmash.com/posts/2009-11-21-making-sense-and-nonsense-of-markov-chains>) Archived (<https://web.archive.org/web/20200204031008/http://rarlindseysmash.com/posts/2009-11-21-making-sense-and-nonsense-of-markov-chains>) 2020-02-04 at the Wayback Machine
 - Original paper by A.A Markov(1913): An Example of Statistical Investigation of the Text Eugene Onegin Concerning the Connection of Samples in Chains (translated from Russian) (http://www.alpha60.de/research/markov/DavidLink_AnExampleOfStatistical_MarkovTrans_2007.pdf)
-

Retrieved from "https://en.wikipedia.org/w/index.php?title=Markov_chain&oldid=1181848303"

▪