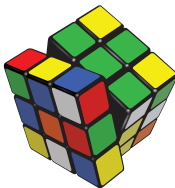


MARKOV CHAINS AND PROCESSES

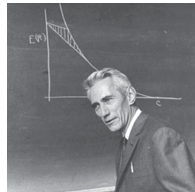
REMARKS AND EXAMPLES

NMAI 060 – 2024

JAROMÍR ANTOCH



January 6, 2025



MAIN GOALS OF THE LECTURE

Main goal of this lecture is study of **Markov chains and Markov processes**, their generalizations and applications.

- Notion of recurrence in theory of probability
 - Random walks
 - Markov chains with discrete states and discrete time
 - Markov chains and beginning of theory of information
 - Ising model and denoising of pictures
 - Markov Chain Monte Carlo simulations
 - Durbin – Watson process
-
- Markov processes with discrete states and continuous time
 - Birth and death process
 - Basics of queuing models, connection to birth and death process
 - Poisson process (homogeneous and non homogeneous)

RECOMMENDED LITERATURE

- Chung, K.L. Markov Chains, Springer
- Durrett, R. Essentials of Stochastic Processes, Springer
- Feller, W. An Introduction to Probability Theory and its Applications, Wiley
- Graham, R., Knuth, D., Patashnik, O. Concrete Mathematics: A Foundation for Computer Science, Addison-Wesley
- Grimmett, G., Stirzaker, D. Probability and Random Processes & Thousand Exercises in Probability, Oxford Univ. Press
- Kemeny, J.G., Snell, L. Finite Markov Chains: With a New Appendix Generalization of a Fundamental Matrix, Springer
- Robert, P.R., Casella, G. Monte Carlo Statistical Methods, Springer
- Ross, S.M. Introduction to Probability Models, 12th ed. Academic Press, Elsevier

EVENTS TO BE CONSIDERED

- Assume a sequence of repeated (not necessarily independent)¹ trials, each of them having the same **finite** or **countable** set of possible outcomes $\{E_j\}_{j \in \mathcal{J}}$, where usually $\mathcal{J} \subseteq \mathbb{Z}$, and let

$$\{E_{j_1}, E_{j_2}, \dots, E_{j_n}\} \quad (1)$$

denotes event that first trial finished with the result E_{j_1} , second trial finished with the result E_{j_2} , ..., n -th trial finished with the result E_{j_n} .

- Let for all finite sequences it holds that

$$P(E_{j_1}, \dots, E_{j_{n-1}}) = \sum_{k=0}^{\infty} P(E_{j_1}, \dots, E_{j_{n-1}}, E_k), \quad 1 < n < \infty,$$

i.e. probability is defined consistently for all finite sequences.

¹Recall that if the events $\{E_{j_1}, E_{j_2}, \dots, E_{j_n}\}$ are mutually independent, then

$$P(\{E_{j_1}, E_{j_2}, \dots, E_{j_n}\}) = \prod_{k=1}^n P(E_{j_k})$$

RECURRENT EVENTS (DEFINITION)

Def. 1: The attribute ξ defines a **recurrent event** if:

- 1** In order that ξ occurred at the n -th and $n + m$ -th place of the sequence $\{E_{j_1}, E_{j_2}, \dots, E_{j_{n+m}}\}$ it is necessary and sufficient that ξ occurs at the last place of each of the subsequences $\{E_{j_1}, E_{j_2}, \dots, E_{j_n}\}$ and $\{E_{j_{n+1}}, E_{j_2}, \dots, E_{j_{n+m}}\}$.

By the other words, both the sequence $\{E_{j_1}, \dots, E_{j_n}\}$ and the sequence $\{E_{j_{n+1}}, \dots, E_{j_{n+m}}\}$ have property ξ .

- 2** If ξ occurs at the n th place then identically

$$P\left(\underbrace{E_{j_1}, \dots, E_{j_n}}_A, \underbrace{E_{j_{n+1}}, \dots, E_{j_{n+m}}}_B\right) = P\left(\underbrace{E_{j_1}, \dots, E_{j_n}}_A\right) \cdot P\left(\underbrace{E_{j_{n+1}}, \dots, E_{j_{n+m}}}_B\right)$$

ξ ξ ξ ξ

RECURRENT EVENTS (CONT.)

- Let ξ be an attribute of finite sequences; that is, we suppose that it is uniquely determined whether a sequence $\{E_{j_1}, E_{j_2}, \dots, E_{j_n}\}$ has, or has not, a characteristic ξ .
- We agree that the expression ξ occurs at the n -th place of (finite or infinite) sequence $\{E_{j_1}, E_{j_2}, \dots\}$ means that the sequence $\{E_{j_1}, E_{j_2}, \dots, E_{j_n}\}$ has the property ξ is an abbreviation for the statement that the sequence $\{E_{j_1}, E_{j_2}, \dots, E_{j_n}\}$ has the attribute ξ .
- This convention also implies, that the occurrence of ξ at n th place depends solely on the outcome of the first n trials.
- It is also understood that when speaking of a recurrent event ξ , we are really referring to a class of events defined by the property that ξ occurs.

EXAMPLES OF RECURRENT EVENTS

Ex. 1: Assume a sequence of independent trials with dichotomous (alternative) response as, e.g., flips of the coin with $P(\text{success}) = p$.

We say that recurrent event ξ occurred in time n , if number of successes and failures after n trials are equal.

Ex. 2: Consider particle moving on grid points of the plane. In each step particle moves up, down, left or right randomly and independently from previous steps. Assume, moreover, that different directions do not arrive necessarily with the same probability.

We say that recurrent event ξ occurred in time n if we are back in the starting point after n steps.

Ex. 3: Consider a particle moving on grid points of the space. In each step particle moves up, down, left, right, backwards or forward randomly and independently from previous steps. Assume, moreover, that different directions do not arrive necessarily with the same probability.

We say that recurrent event ξ occurred in time n if we are back in the starting point after n steps.

PROBABILITIES $\{u_n\}$, $\{f_n\}$ AND THEIR INTERRELATION

Def. 2: To each recurrent event ξ we assign two sequences of numbers

$$u_n = P(\xi \text{ occurred in the } n^{\text{th}} \text{ trial}) \quad 1 \leq n < \infty$$

$$f_n = P(\xi \text{ occurred in the } n^{\text{th}} \text{ trial for the first time}) \quad 1 \leq n < \infty$$

We define formally $u_0 = 1$, $f_0 = 0$ and introduce generating functions $F(x) = \sum_{n=0}^{\infty} f_n x^n$ and $U(x) = \sum_{n=0}^{\infty} u_n x^n$.

Thm. 1: Between probabilities $\{u_n\}$ and $\{f_n\}$, respectively between corresponding generating functions $F(x)$ and $U(x)$, following relations hold

$$u_n = f_0 u_n + f_1 u_{n-1} + \dots + f_n u_0 \quad \forall n \geq 1$$

$$U(x) - 1 = F(x)U(x) \quad -1 < x < 1$$

EXAMPLES OF RECURRENT EVENTS (CONT.)

Ex. 4: Assume Rubik's cube, i.e. mechanical puzzle composed of smaller sub-cubes. An internal pivot mechanism enables each face to turn independently, thus mixing up the colors. Most typical model is $3 \times 3 \times 3$, for which we have $43\,252\,003\,274\,489\,856\,000 \approx 43.25 \times 10^{18}$ possible combinations (https://en.wikipedia.org/wiki/Rubik's_Cube). Generally, for the puzzle to be solved, each face must be returned to have only one color.

We say that recurrent event ξ occurred in time n if we are back in starting position after n steps.

Ex. 5: Assume a sequence of independent trials with dichotomous response (T, F) . We say that recurrent event ξ occurred in time n if some prescribed pattern as, e.g., TTT or $TTFTF$, occurred at the end of n trials. In the case of an event TTT it means that n^{th} , $(n-1)^{\text{st}}$ and $(n-2)^{\text{nd}}$ trial ended with a result T . Moreover, after observing prescribed sequence we start to count from the beginning.

EXAMPLES OF RECURRENT EVENTS (CONT.)

Assume a sequence of independent trials with dichotomous response (T, F) . We say that recurrent event ξ occurred in time n if some prescribed pattern as, e.g., TTT or $TTFTF$, occurred at the end of n trials and we start from scratch every time a pattern is completed.

- The term “success run of length r ” has been defined in the literature in several different ways. It is largely matter of convention and convenience whether a sequence of three consecutive successes is said to contain 0, 1, or 2 runs of length 2, and for different purposes different definitions have been adapted. **However, if we are to use the theory of recurrent events, then the notion of runs of length r must be defined so that we start from scratch every time a run is completed. This means adapting a following definition:**

- *A sequence of n letters T and F has as many runs of length r as there are nonoverlapping uninterrupted successions of exactly r letters T . In a sequence of Bernoulli trials a run of length r occurs at the n^{th} trial, if the n^{th} trial adds a new run to the sequence.*

- Thus in a series $SSSSFSSSSSS$ we have three runs of length 3, and they occur at trials 3, 8, 11; there are five runs of the length 2, and they occur at trials 2, 4, 7, 9, 11.

EXAMPLES OF RECURRENT EVENTS (CONT.)

Ex. 6: “Random walks” on vertexes of graphs, vertexes of multidimensional cubes, etc., can be often described using the theory of recurrent events.

REPEATED OCCURRENCES OF RECURRENT EVENTS

Rem. 1:

- If $f = \sum_n f_n = 1$, then $\{f_n\}$ correspond to a random variable T_1 describing waiting time of the first occurrence of ξ .
- If $f < 1$, then waiting time T_1 is so called **improper random variable**, which with the positive probability ($= 1 - f$) attains improper value ∞ , being interpreted as the recurrent event that ξ did not came up.

REPEATED OCCURRENCES OF RECURRENT EVENTS

Rem. 2:

- If $f = \sum_n f_n = 1$, then $\{f_n\}$ correspond to a random variable T_1 describing waiting time of the first occurrence of ξ .
- If $f < 1$, then waiting time T_1 is so called **improper random variable**, which with the positive probability ($= 1 - f$) attains improper value ∞ , being interpreted as the recurrent event that ξ did not came up.

Thm. 2: Denote by $f_n^{(r)}$, $1 \leq n < \infty$, probability of the event that ξ occurred for the r^{th} time in time n , and denote $f_0^{(r)} = 0$. Then it holds

$$\{f_n^{(r)}\} = \{f_n\}^{r*},$$

where $\{f_n^{(r)}\}$ denote r -th convolution of $\{f_n\}$.

Thm. 3: Probability that event ξ will occurred in infinitely long sequence of trials at least r -times is equal to f^r , where $f = \sum_n f_n$.

ANOTHER APPROACH TO INTRODUCING RECURRENT EVENTS

Let T_i , $1 \leq i \leq r$, are independent integer valued rv's with the same distribution $\{f_n\}$, where T_i is interpreted as the time between $(i-1)$ -st and i -th occurrence of ξ (so called return time). Then

$$T^{(r)} = T_1 + \dots + T_r$$

can be interpreted as the **waiting time to the r -th occurrence of ξ** .

Notion of recurrent event can be introduced in the following way.

Def. 3: Let T_1, T_2, \dots be independent integer valued random variables with the same distribution $\{f_n\}$. Then:

- We match the statement **recurrent event ξ occurred in time n** with the statement **there exists r such that $T_1 + T_2 + \dots + T_r = n$** .
- We match the statement **recurrent event occurred in time n for the r -th time** with the statement **$T_1 + T_2 + \dots + T_r = n$** .

CLASSIFICATION OF RECURRENT EVENTS

Def. 4: Event ξ is called **persistent** if $f = 1$, respectively **transient** if $f < 1$, where $f = \sum_n f_n$.

Thm. 4: Probability, that an event ξ will occurred infinitely many times in infinitely long series of trials is one for persistent events and zero for a transient events.

Thm. 5: An event ξ is transient if and only if $\sum_{n=0}^{\infty} u_n < +\infty$. In such a case $f = (u - 1)/u$, where $u = \sum_{n=0}^{\infty} u_n$.

Def. 5: If $f = 1$, then we denote $\mu = E T_1 = \sum_{n=0}^{\infty} n f_n$ and interpret is as the **mean renewal (return) time** of ξ .

Def. 6:

- An event ξ is called **positive recurrent** if $\mu < +\infty$.
- An event ξ is called **null recurrent** if $\mu = +\infty$.

Def. 7: An event ξ is called **periodical** if there exists natural $\lambda > 1$ such that $u_n = 0 \forall n$ which are not divisible by λ . Largest λ with this property is called a **period** of ξ .

EXAMPLES OF RECURRENT EVENTS

Ex. 7: Assume a sequence of independent trials with dichotomous (yes/no) response, where $P(\text{yes}) = p$. We say that recurrent event ξ occurred in time n if the number of positive trials is equal to the number of negative trials. Show that it is a periodical recurrent event for which it holds:

- if $p = 1/2$ then ξ is null recurrent
- if $p \neq 1/2$ then ξ is transient
- calculate probabilities u_n and f_n and their approximations
- $U(x) = \sum_{n=0}^{\infty} \binom{2n}{n} (pqx^2)^n = \frac{1}{\sqrt{1-4pqx^2}}$
- $F(x) = 1 - \sqrt{1-4pqx^2}$
- $f_{2n-1} = 0$, $f_{2n} = \frac{2}{n} \binom{2n-2}{n-1} p^n q^n$, $n = 1, 2, \dots$
- if $p = 1/2$ then $u_{2n} \approx 1/\sqrt{\pi n}$
- simulate couple of random walks of a length at least 10^5 for different values of p (including $p = 1/2$)
- plot corresponding graphs of simulated random walks

EXAMPLES OF RECURRENT EVENTS

Ex. 8: Return to the Examples 2 and 3. Calculate probabilities u_n and their approximations.

Hint. Recall multinomial distribution and Law of total probability.

LIMIT THEOREM

Thm. 6: Let an event ξ is recurrent and non periodical. Then it holds:

$$\lim_{n \rightarrow \infty} u_n = \begin{cases} \frac{1}{\mu} & \mu < \infty \\ 0 & \mu = \infty \end{cases}$$

Thm. 7: Let an event ξ is recurrent and periodical with period λ . Then it holds:

$$\lim_{n \rightarrow \infty} u_{n\lambda} = \begin{cases} \frac{\lambda}{\mu} & \mu < \infty \\ 0 & \mu = \infty \end{cases}$$

Proof: Follows from Thm. 55 and Rmk 57.

ASYMPTOTIC DISTRIBUTION OF FREQUENCIES OF RECURRENT EVENTS

Thm. 8: Let ξ be a positive recurrent event. Denote N_n number of occurrences of ξ up to the time n and $T^{(r)}$ waiting time till the r^{th} occurrence of ξ . Then the events $[N_n \geq r]$ and $[T^{(r)} \leq n]$, $1 \leq r \leq n < \infty$, are equivalent. Assume, moreover, that $E T_1 = \mu$ and $\text{var } T_1 = \sigma^2 < +\infty$. Then it holds:

$$N_n \stackrel{\mathcal{D}}{\sim} \mathcal{N}\left(\frac{n}{\mu}, \frac{n\sigma^2}{\mu^3}\right) \text{ and } T^{(r)} \stackrel{\mathcal{D}}{\sim} \mathcal{N}(r\mu, r\sigma^2), \text{ i.e.}$$

$$\lim_{n \rightarrow \infty} P\left(\frac{N_n - n/\mu}{\sqrt{n\sigma^2/\mu^3}} \leq x\right) = \Phi(x) \quad \forall x \in \mathbb{R}_1$$

$$\lim_{r \rightarrow \infty} P\left(\frac{T^{(r)} - r\mu}{\sqrt{r\sigma^2}} \leq x\right) = \Phi(x) \quad \forall x \in \mathbb{R}_1$$

where

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

ASYMPTOTIC DISTRIBUTION OF FREQUENCIES OF RECURRENT EVENTS

Thm. 9: Let ξ be a positive recurrent event. Then $E N_n \approx n/\mu$ as $n \rightarrow \infty$, where $\mu = E T_1$ is the **mean renewal (return) time of ξ** .

Rem. 3: Let ξ be a null recurrent. Then $E N_n$ is not generally of the order n^1 . An example has been described in Ex. 7. Show that in such a case $E N_{2n} \approx 2 \sqrt{n/\pi}$.

RECURRENT EVENTS WITH DELAY

Recurrent events with delay can be introduced in the same way as indicated above.

Def. 8: Assume independent integer valued random variables T_1, T_2, \dots , where T_1 has distribution $\{b_n\}$, while T_2, T_3, \dots have distribution $\{f_n\}$. We say that recurrent event with delay ξ occurred in time n for the r -th time, if

$$T_1 + T_2 + \dots + T_r = n \tag{2}$$

Analogously, recurrent event with delay ξ occurred in time n if there exists r such that (2) holds.

Rem. 4: In Def. 8 is random variable T_1 interpreted as **time to the first arrival** of ξ , while T_2, T_3, \dots as **renewal (return) times**.

Ex. 9: Assume again Rubik's cube from Example 4, which is at the beginning arbitrarily scrambled. Then delay T_1 describes time to the first arrival when each face have only one color (initial position). Return times T_2, T_3, \dots correspond to the returns to the initial position.

RECURRENT EVENTS WITH DELAY (CONT.)

Thm. 10: Let u_n denotes probability of the event that ξ occurred at time n . Let $u_0 = f_0 = b_0 = 0$. Then

$$u_n = b_n + f_0 u_n + \dots + f_n u_0, \quad \text{i.e.} \quad \{u_n\} = \{b_n\} + \{f_n\} \star \{u_n\}.$$

Rem. 5: Recall equivalence between following events

$$\begin{aligned} & \left[\xi \text{ occurred at time } n \right] \equiv \\ & \bigcup_{k=1}^{n-1} \left[\xi \text{ occurred at time } n \text{ and the last occurrence} \right. \\ & \quad \left. \text{before that event at time } k \right] \\ & \cup \left[\xi \text{ occurred at time } n \text{ for the first time} \right] \end{aligned}$$

RENEWAL EQUATION

Rem. 6: Limit theorems of the previous paragraphs can be considered as a special case of a general theorem, which can be formulated analytically without the use of probability. However, under appropriate assumptions about the sequences which enter, it can have also a probability meaning.

Def. 9: Let a_0, a_1, a_2, \dots and b_0, b_1, b_2, \dots be two sequences of real numbers such that

$a_0 = 0, 0 \leq a_n \leq 1, b_n \geq 0, n = 0, 1, 2, \dots, \sum_{i=n}^{\infty} b_n < \infty$. Put $u_n = b_n + a_0 u_n + a_1 u_{n-1} \dots + a_n u_0, n = 0, 1, 2, \dots$, i.e.

$$\{u_n\} = \{b_n\} + \{a_n\} \star \{u_n\} \quad (3)$$

Relation (3) is in the literature usually called *renewal equation*.

Rem. 7: For generating functions of sequences introduced in Def. 9 it holds

$$U(x) = B(x) + A(x)U(x) \quad \equiv \quad U(x) = \frac{B(x)}{1 - A(x)}$$

RENEWAL EQUATION (CONT.)

Def. 10: We call sequence $\{a_n\}$ periodic if there exist $\lambda > 1$ such that $a_n = 0 \forall n$ not divisible by λ . Largest λ with this property is called **period**.

Thm. 11: Let sequence $\{a_n\}$ is aperiodic. Then it holds:

- 1 If $\sum_{n=1}^{\infty} a_n < 1$ then $\sum_{n=1}^{\infty} u_n < \infty$.
- 2 If $\sum_{n=1}^{\infty} a_n = 1$, i.e. we can consider $\{a_n\}$ to be distribution of some rv describing return time of some aperiodic recurrent event ξ , then

$$\lim_{n \rightarrow \infty} u_n = \begin{cases} \sum_{n=0}^{\infty} b_n / \sum_{n=1}^{\infty} n a_n, & \sum_{n=1}^{\infty} n a_n < \infty, \\ 0, & \sum_{n=1}^{\infty} n a_n = \infty. \end{cases}$$

- 3 If $\sum_{n=1}^{\infty} a_n > 1$ then it holds for $n \rightarrow \infty$

$$u_n \approx \frac{B(x)}{x^{n+1} A'(x)} \Big|_{x=1},$$

where $x < 1$ is the only root of equation $A(x) = 1$.

RENEWAL EQUATION (CONT.)

Thm. 12: Let sequence $\{a_n\}$ is periodical with period λ . Then it holds:

- 1 If $\sum_{n=1}^{\infty} a_n < 1$ then $\sum_{n=1}^{\infty} u_n < \infty$.
- 2 If $\mu = \infty$ then $\lim_{n \rightarrow \infty} u_n = 0$.
- 3 If $\mu < \infty$ and $\sum_{n=1}^{\infty} a_n = 1$, i.e. if $\{a_n\}$ describes distribution of return time of some periodic recurrent event ξ , then it holds for any $0 \leq j < \lambda$:

$$\lim_{n \rightarrow \infty} u_{n\lambda+j} = \frac{\lambda \sum_{k=0}^{\infty} b_{k\lambda+j}}{\mu} \quad \& \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n u_j = \frac{\sum_{k=0}^{\infty} b_k}{\mu}$$

MARKOV CHAINS

Def. 11: A sequence of trials, each of them having the same finite or infinite set of possible outcomes, will be called a **Markov chain (MC)**, if probability of each finite sequence of results is given by

$$P(E_{j_0}, E_{j_1}, \dots, E_{j_n}) = a_{j_0} p_{j_0 j_1} \cdots p_{j_{n-1} j_n}, \quad (4)$$

where a_k , $k = 1, 2, \dots$ are probabilities of the starting outcome (zero's trial) and p_{jk} , $1 \leq j, k < +\infty$, is (for all trials the same) conditional probability of the outcome E_k given the outcome E_j in previous trial.

Rem. 8: A sequence $\{a_k\}$ is called an **initial distribution** and probabilities p_{jk} are called **transition probabilities**. Recall that for independent events it is enough to know just probabilities p_i , while for a description of MC we need to know $\mathbf{a} \equiv \{a_k\}$ and $\mathbf{P} \equiv \{p_{jk}\}$. Notice that $\sum_j p_{ij} = 1 \quad \forall i \in \mathbb{N}$.

MORE GENERAL DEFINITION OF MARKOV CHAIN

Assume MC with states S_1, S_2, \dots , initial distribution $\{a_j\}$ and matrix of transition probabilities $\mathbf{P} \equiv \{p_{ij}\}$. Introduce integer valued random variables (ivrv's) X_n , $0 \leq n < \infty$, using the equivalence

$$X_n \text{ has value } j \Leftrightarrow \text{MC is in time } n \text{ in state } S_j$$

Then it holds $\forall i_0, i_1, \dots, i_{n-1}, i, j$ and $\forall n \in \mathbb{N}_0$

$$P(X_0 = j) = a_j$$

$$P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_1 = x_1, X_0 = i_0) = p_{ij}$$

Rem. 9: MC is then usually identified with a sequence of ivrv's constructed above. If we require fulfillment of the second equation only, we get:

Def. 12: Sequence of ivrv's X_n , $n \in \mathbb{N}_0$, is called **Markov chain**, if $\forall i_0, i_1, \dots, i_{n-1}, i, j$ and $\forall n \in \mathbb{N}_0$ it holds

$$\begin{aligned} &P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_1 = x_1, X_0 = i_0) \\ &= P(X_{n+1} = j | X_n = i) \end{aligned}$$

MORE GENERAL DEFINITION OF MARKOV CHAIN (CONT.)

Rem. 10: Markov chains considered up to now had the property that p_{ij} do not depend on n . Such MC are called **homogeneous**.

If $P(X_{n+1} = j | X_n = i)$ depends on n , we will denote it $p_{ij}(n, n+1)$ and to call them **non homogeneous** Markov chains.

Rem. 11: Probabilities

$$P(X_0 = j_0, X_1 = j_1, \dots, X_n = j_n) = a_{j_0} p_{j_0 j_1} \dots p_{j_{n-1} j_n}$$

change

$$P(X_0 = j_0, X_1 = j_1, \dots, X_n = j_n) = a_{j_0}(0, 1) p_{j_0 j_1}(1, 2) \dots p_{j_{n-1} j_n}(n-1, n)$$

Ex. 10: Let Y_k , $1 \leq k < \infty$, are independent ivrv's and $X_n = \sum_{k=1}^n Y_k$. Then sequence $\{X_n\}$, $1 \leq n < \infty$ forms Markov chain. Prove it.

Ex. 11: Let Y_k , $1 \leq k < \infty$, are independent ivrv's and assume a sequence of moving sums $X_n^* = \sum_{k=1}^r Y_{n+k}$, r being fixed. Then a sequence $\{X_n^*, 1 \leq n < \infty\}$ **generally does not form** Markov chain.

EXAMPLES OF MARKOV CHAINS

- 1 Random walk on line.
- 2 Random walk on line with reflecting barriers.
- 3 Random walk on line with absorbing barriers.
- 4 Ehrenfest's imaginary model. Let a distinguishable molecules are randomly placed into two containers denoted A and B . In each step we select randomly one molecule with the probability $1/a$ and move it to the other container. State of the system is number of molecules in container A .
- 5 Modified Ehrenfest's imaginary model. Let a non distinguishable molecules are randomly placed into two containers denoted A and B . In each step we select randomly one container with the probability $1/2$ and move one randomly selected molecule from this container to the other one. State of the system is number of molecules in container A .
- 6 Sequence of independent random trials.
- 7 Gambler's ruin problem. Rubick's cube. Etc.

HIGHER ORDER PROBABILITIES

Thm. 13: Transition probability from state S_j to the state S_k after n -steps, denoted $p_{jk}^{(n)}$, is (j, k) -th element of matrix \mathbf{P}^n . We define $\mathbf{P}^0 = \mathbf{I}$.

Rem. 12: Matrix \mathbf{P}^n can be calculated using, e.g.

- “Sequential” power raising of transition matrix \mathbf{P} .
- Directly from principle.
- Using Perron’s formula that uses eigenvalues of \mathbf{P} . Etc.

Def. 13: Aside conditional probabilities $p_{jk}^{(n)}$ we introduce **non conditional (absolute) probability** $a_k^{(n)}$ as probability of event describing that system is in time n at state S_k .

Rem. 13: Evidently it holds:

$$a_k^{(0)} = a_k, \quad a_k^{(n)} = \sum_j a_j p_{jk}^{(n)} \quad \text{and} \quad a_k^{(n+m)} = \sum_j a_j^{(m)} p_{jk}^{(n)}$$

If there exists $\lim_{n \rightarrow \infty} p_{jk}^{(n)}$ independent on j , then there exists also $\lim_{n \rightarrow \infty} a_k^{(n)}$ and are equal each to other.

NOTATION

- $f_{jj}^{(n)}$... probability of the first return to state S_j in time n , provided in time 0 we were in state S_j
- $f_{ij}^{(n)}$... probability of the first return to state S_j in time n , provided in time 0 we were in state S_i
- $p_{jj}^{(n)}$... probability of the event that system is in time n at state S_j , provided at time 0 has been at state S_j
- $p_{ij}^{(n)}$... probability of the event that system is in time n at state S_j , provided at time 0 has been at state S_i

Thm. 14: Put $f_{jj}^{(0)} = 0$, $f_{ij}^{(0)} = 0$, $p_{jj}^{(0)} = 1$, $p_{ij}^{(0)} = 0$, $p_{jj}^{(1)} = p_{jj}$. Then it holds

$$p_{jj}^{(n)} = f_{jj}^{(0)} p_{jj}^{(n)} + f_{jj}^{(1)} p_{jj}^{(n-1)} + \dots + f_{jj}^{(n-1)} p_{jj} + f_{jj}^{(n)} p_{jj}^{(0)}, \quad 1 \leq n < \infty$$

$$\{p_{ij}^{(n)}\} = \{f_{ij}^{(n)}\} + \{f_{ij}^{(n)}\} \star \{p_{ij}^{(n)}\}$$

CLASSIFICATION OF STATES OF MC

Thm. 15: Let us fix in a MC a state S_j .

- a** If the system is at the beginning at state S_j , then each visit to state S_j is recurrent event.
- b** If the system is at the beginning at state S_j , then each visit to state S_j is recurrent event with delay.

CLASSIFICATION OF STATES OF MC

Thm. 15: Let us fix in a MC a state S_j .

- a** If the system is at the beginning at state S_j , then each visit to state S_j is recurrent event.
- b** If the system is at the beginning at state S_j , then each visit to state S_j is recurrent event with delay.

**Theory of Markov chains is in principle theory of recurrent events.
New is the fact that we study many recurrent events in parallel !!!**

CLASSIFICATION OF STATES OF MC

Thm. 15: Let us fix in a MC a state S_j .

- a** If the system is at the beginning at state S_j , then each visit to state S_j is recurrent event.
- b** If the system is at the beginning at state S_j , then each visit to state S_j is recurrent event with delay.

**Theory of Markov chains is in principle theory of recurrent events.
New is the fact that we study many recurrent events in parallel !!!**

Notions concerning classification of recurrent events naturally transfer to the states of MC.

Thm. 16: Let us fix in a Markov chain state S_j .

- State S_j is transient $\Leftrightarrow \sum_{n=1}^{\infty} p_{jj}^{(n)} < \infty$. In such a case $\sum_{n=1}^{\infty} p_{ij}^{(n)} < \infty \forall i$.
- State S_j is null recurrent $\Leftrightarrow \sum_{n=1}^{\infty} p_{jj}^{(n)} = \infty$ and $\lim_{n \rightarrow \infty} p_{jj}^{(n)} = 0$. In such a case $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0 \forall i$.

CLASSIFICATION OF STATES OF MC (CONT.)

Thm. 17: Let us fix in a Markov chain state S_j .

- If state S_j is positive recurrent, then

$$\lim_{n \rightarrow \infty} p_{jj}^{(n)} = \frac{1}{\mu_j}, \quad \lim_{n \rightarrow \infty} p_{ij}^{(n)} = \frac{f_{ij}}{\mu_j}, \quad i \neq j, \quad \text{where} \quad f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)}$$

- If state S_j is positive recurrent with period λ , then

$$\lim_{n \rightarrow \infty} p_{jj}^{(n\lambda)} = \frac{\lambda}{\mu_j}$$

and for all $i \neq j$ and $0 \leq \nu \leq \lambda - 1$

$$\lim_{n \rightarrow \infty} p_{ij}^{(n\lambda + \nu)} = \frac{\lambda \sum_{k=0}^{\infty} f_{ij}^{(k\lambda + \nu)}}{\mu_j}$$

Further, it holds:

$$\lim_{n \rightarrow \infty} \bar{p}_{ij}^{(n)} = \frac{f_{ij}}{\mu_j}, \quad \text{where} \quad \bar{p}_{ij}^{(n)} = \frac{1}{n} \sum_{k=1}^{\infty} p_{ij}^{(k)}$$

REDUCIBLE AND IRREDUCIBLE MC

Def. 14: We say that state S_k is accessible from state S_j , if there exists $n \geq 0$ such that $p_{jk}^{(n)} > 0$.

Rem. 14: In the sense of Def. 14 is each state accessible from itself, because $p_{jj}^0 = 1$.

Def. 15: Nonempty set of events C is called closed, if no state outside of C is accessible from any state inside C . Smallest closed set containing given set of states is called its **closure**.

Thm. 18: Set of states C is closed $\Leftrightarrow p_{jk} = 0$ for all $S_j \in C$ and $S_k \notin C$.

Def. 16: If a set with one point $\{S_j\}$ is closed, i.e. if $p_{jj} = 1$, then state S_j is called **absorbing state**.

Rem. 15: If we omit in a matrix of transitional probabilities P of given Markov chain all rows and columns corresponding to the states outside closed set C , we obtain again stochastic matrix. Thus, C correspond also to some markov chains, usually called **subchain** of the original Markov chain.

REDUCIBLE AND IRREDUCIBLE MC (CONT.)

Def. 17: MC is called **irreducible** if it does not contain aside the set of all states another closed set of states. Otherwise it is called reducible.

Thm. 19: MC is irreducible \Leftrightarrow each of its states is accessible from any other state.

Thm. 20: Markov chain with **finitely** many states is reducible \Leftrightarrow corresponding matrix of transitional probabilities P can be, after eventual renumeration of states, written in the form

$$P = \begin{pmatrix} P_1 & 0 \\ A & B \end{pmatrix}$$

where on diagonal we have square matrices.

Rem. 16: We say that states S_j and S_k are of the same type, if both are transient, or both are null recurrent or positive recurrent, and in parallel are both either periodical or non periodical with the same period λ .

REDUCIBLE AND IRREDUCIBLE MC (CONT.)

Thm. 21: If state S_k is accessible from state S_j and state S_j is accessible from state S_k , then they are of the same type.

Thm. 22: In the irreducible MC all the states are of the same type.

Thm. 23: In MC with **finitely** many states there does not exist null states and it is not possible, that all states are transient ones.

STATIONARY DISTRIBUTION

Def. 18: Assume irreducible MC described by a matrix of transitional probabilities \mathbf{P} . A distribution $\{v_j\}$ is called **stationary distribution** of this chain if $\forall j$

$$v_j = \sum_i v_i p_{ij} \quad (5)$$

This relation can be written in the matrix form as $\mathbf{v} = \mathbf{P}^\top \mathbf{v}$, where \mathbf{P}^\top denotes transposed matrix \mathbf{P} .

Thm. 24: In irreducible MC there exists a stationary distribution \Leftrightarrow all states are positive recurrent. This stationary distribution \mathbf{v} is unique and $\forall i, j$ it holds:

$$v_j = \lim_{n \rightarrow \infty} p_{ij}^{(n)} > 0 \quad \text{in non periodical case}$$

$$v_j = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n p_{ij}^{(k)} > 0 \quad \text{in periodical case}$$

STATIONARY DISTRIBUTION (CONT.)

Rem. 17: In an irreducible MC with **finitely** many states stationary distribution exists, compare Thm. **23**.

Def. 19: Matrix with non negative elements such that all row's and column's sums are equal to one is called **doubly stochastic**.

Thm. 25: Assume irreducible MC with doubly stochastic matrix **P** . If number of states n is **finite**, then stationary distribution $\{v_j\}$ is discrete uniform, i.e. $v_i = 1/n$ for $1 \leq i \leq n$. If number of states is infinite, then stationary distribution does not exist.

Ex. 12:

- Find stationary distribution for a random walk with two reflecting barriers.
- Decide for which values of p there exists a stationary distribution for a random walk with one reflecting barrier in zero.
- Find stationary distribution for Ehrenfest's imaginary models.

REVERSIBLE MARKOV CHAINS

Def. 20: Assume irreducible MC (\mathbf{P}, \mathbf{a}) . If there exist positive numbers π_i such that $\pi_i p_{ij} = \pi_j p_{ji} \quad \forall i, j$, we say that this MC is reversible.

Rem. 18: Assume irreducible reversible MC (\mathbf{P}, \mathbf{a}) with $\sum \pi_i = 1$. Then it holds:

$$P(X_0 = i, X_1 = j, X_2 = k) = P(X_0 = k, X_1 = j, X_2 = i) \quad \forall i, j, k$$

Thm. 26: If MC (\mathbf{P}, \mathbf{A}) is reversible, then corresponding vector π is its stationary distribution if $\sum \pi_i = 1$.

Rem. 19: Thm. 26 does not hold in opposite direction, i.e., existence of the stationary distribution does not imply that corresponding MC is reversible.

REDUCIBLE MARKOV CHAINS

Rem. 20: Recall that nonempty set of states C is called closed, if no one of states outside C is accessible from any state in C . Smallest closed set containing a given set of states is called its **closure**.

Thm. 27: Assume reducible MC with **finitely** many states. Then probability that chain will be absorbed in any closed subset of states converges to 1 for $n \rightarrow \infty$, and it does not matter in which state we started.

Rem. 21: Assume irreducible MC with **finitely** many states described by matrix of transitional probabilities P , and construct sequentially matrices P^2, P^3, \dots

$$\underbrace{\begin{pmatrix} S & 0 \\ R & Q \end{pmatrix}}_P \quad \underbrace{\begin{pmatrix} S^2 & 0 \\ RS + QR & Q^2 \end{pmatrix}}_{P^2} \quad \underbrace{\begin{pmatrix} S^3 & 0 \\ \dots & Q^3 \end{pmatrix}}_{P^3} \dots$$

Then $\lim_{n \rightarrow \infty} Q^n = \mathbf{0}$ and $p_{ij}^n \rightarrow 0$ exponentially fast for each two states S_i and S_j which are transient.

REDUCIBLE MARKOV CHAINS (CONT.)

Rem. 22: For each reducible MC described by matrix of transitional probabilities $\mathbf{P} = \begin{pmatrix} \mathbf{S} & \mathbf{0} \\ \mathbf{R} & \mathbf{Q} \end{pmatrix}$ corresponding matrix $\mathbf{I} - \mathbf{Q}$ has inversion and it holds

$$\mathbf{I} + \mathbf{Q} + \mathbf{Q}^2 + \dots = \sum_{k=0}^{\infty} \mathbf{Q}^k = (\mathbf{I} - \mathbf{Q})^{-1}$$

Matrix $(\mathbf{I} - \mathbf{Q})^{-1}$ is called **fundamental matrix**.

Thm. 28: Assume reducible MC with finitely many states and denote by \mathcal{T} subset of temporary states. Let u_{ij} be expected number of time periods MC will spend in state $S_j \in \mathcal{T}$ given that it starts in state $S_i \in \mathcal{T}$. Then it holds:

$$u_{ij} = I_{[i=j]} + \sum_{k \in \mathcal{T}} p_{ik} u_{kj} + \sum_{k \notin \mathcal{T}} p_{ik} u_{kj} \quad \equiv \quad \mathbf{U} = \mathbf{I} + \mathbf{QU}$$

so that

$$\mathbf{U} = (\mathbf{I} - \mathbf{Q})^{-1} = \mathbf{N}$$

REDUCIBLE MARKOV CHAINS (CONT.)

Rem. 23: Analogously, let w_{ij} be variance of time periods MC will spend in state $S_j \in \mathcal{T}$ given that it starts in state $S_i \in \mathcal{T}$, and $\mathbf{W} = (w_{i,j})_{i,j}$. It holds

$$\mathbf{W} = \mathbf{N}(2\text{diag}(\mathbf{N}) - \mathbf{I}) - \mathbf{N}^2$$

TRANSIENT STATES

Assume MC containing both transient and recurrent states. Let T denotes a set of all transient states and C is any irreducible closed set of recurrent states.

Let us fix state $S_j \in T$ and denote:

- $x_j = P(S_j \rightarrow C)$ probability of absorption in C provided we started in transient state S_j
- $1 - x_j$ is probability of event that system, which is at the beginning at state $S_j \in T$, either stay forever in T or will be absorbed in some other closed set of states
- $x_j^{(1)} = \sum_{k \in C} p_{jk}$ is probability of absorption (in C) in the first step
- y_j is probability of the event that system, which is at the beginning at state $S_j \in T$, will stay in T forever

TRANSIENT STATES (CONT.)

Thm. 29: Probabilities x_j , $j \in T$, satisfy a set of equations

$$\xi_j - \sum_{v \in T} p_{jv} \xi_v = x_j^{(1)} \quad (6)$$

Thm. 30: Probabilities y_j , $j \in T$, satisfy a set of equations

$$\eta_j = \sum_{v \in T} p_{jv} \eta_v \quad (7)$$

Thm. 31: Set of equations (6) has unique bounded solution \Leftrightarrow set of equations (7) does not have other bounded solution than the trivial one.

Thm. 32: Probabilities y_j are equal to zero $\forall j \in T \Leftrightarrow$ a set of equations (7) does not have other solution than the trivial one.

Thm. 33: In a chain with **finitely** many states all $y_j = 0$ and x_j form unique solution of set of equations (6).

TRANSIENT STATES (CONT.)

Thm. 34: In a chain with states S_0, S_1, S_2, \dots are all states transient \Leftrightarrow set of equations

$$\eta_j = \sum_{\nu=1}^{\infty} p_{j\nu} \eta_{\nu}, \quad 1 \leq j < \infty, \quad (8)$$

has nontrivial bounded solution.

ISING MODEL – NOTATION

- G ... graph
- V ... vertexes of given graph
- $|V| = \text{card}(V)$
- E ... edges of given graph
- $[i \leftrightarrow j]$... indicator of edge between vertices i and j
- Simplest situation : Each vertex i has two states $\sigma_i \in \{-1, +1\}$ (black and white)
- Generally we assume states $\{1, \dots, K\}$ which can describe levels of gray or colors, $|V| = \text{card}(V)$, etc.
- $\sigma = (\sigma_1, \dots, \sigma_{|V|})$ describes states of the system
- State space \mathcal{S} is $\{-1, +1\}^{|V|}$, respectively $\{1, \dots, K\}^{|V|}$, etc.

ISING MODEL

Def. 21: **Ising model** is probability distribution $\pi(\sigma; \beta)$ on state space $\mathcal{S} = \{-1, +1\}^{|V|}$, where

$$\pi(\sigma; \beta) = \frac{e^{-\beta H(\sigma)}}{C_\beta}, \quad (9)$$

$$H(\sigma) = \sum_{[i \leftrightarrow j] \in E} I[\sigma_i \neq \sigma_j] \quad \text{and} \quad C_\beta = \sum_{\sigma^* \in \mathcal{S}} e^{-\beta H(\sigma^*)}$$

Rem. 24: Function $H(\sigma)$ is called (in physics) **Hamiltonian** and represents “energy” of system’s configuration σ .

Rem. 25: For $\beta > 0$ are less probable those configurations σ having $H(\sigma)$ small, being a case that many neighbors have the same value (spin). We often say that they have small energy (small information).

Def. 22: Mean spin (mean energy) of the configuration σ is

$$M(\sigma) = \frac{1}{|V|} \sum_{i \in V} \sigma_i$$

MODIFICATIONS OF ISING MODEL

■ Classical Ising model:

$$\mathcal{S} = \{-1, +1\}^{|V|} \quad \text{a} \quad H(\sigma) = \sum_{[i \leftrightarrow j] \in E} I[\sigma_i \neq \sigma_j].$$

$$\begin{array}{ccc} \begin{pmatrix} 1 & 1 & 1 \\ 0 & \mathbf{1} & 0 \\ 1 & 0 & 1 \end{pmatrix} & \Rightarrow & \begin{pmatrix} 1 & 0 & 1 \\ 3 & 3 & 3 \\ 2 & 3 & 2 \end{pmatrix} & \begin{pmatrix} 1 & 1 & 1 \\ 0 & \mathbf{0} & 0 \\ 1 & 0 & 1 \end{pmatrix} & \Rightarrow & \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 2 \\ 2 & 2 & 2 \end{pmatrix} \\ S & & H(S) & S & & H(S) \end{array}$$

■ Ising model with outside field:

$$\mathcal{S} = \{-1, +1\}^{|V|} \quad \text{a} \quad H(\sigma, h) = \sum_{[i \leftrightarrow j] \in E} I[\sigma_i \neq \sigma_j] - h \sum_{i \in V} \sigma_i.$$

For all $\beta > 0$ and $h > 0$ are the values of $+1$ preferred to the values -1 .

MODIFICATIONS OF ISING MODEL (CONT.)

- **Pot's model** for “random patching” of images:

$$\mathcal{S} = \{1, \dots, K\}^{|V|} \quad \text{a} \quad H(\sigma) = \sum_{[i \leftrightarrow j] \in E} I[\sigma_i \neq \sigma_j].$$

- **Ising model for gray-scale images:**

$$\mathcal{S} = \{1, \dots, K\}^{|V|} \quad \text{a} \quad H(\sigma) = \sum_{[i \leftrightarrow j] \in E} f(\sigma_i, \sigma_j),$$

a $f(\cdot)$ is any suitable distance, e.g.:

$$f(\sigma_i, \sigma_j) = |\sigma_i - \sigma_j|^p, \quad p \geq 1.$$

Vertexes typically represent pixels and, unlike as in Pot's model, desire is that neighboring pixels have similar value of “gray”, not being identical.

APPLICATION IN IMAGE ANALYSIS

Problem:

- Assume image represented by a matrix of pixels of size $L_1 \times L_2$
- Vertices correspond to pixels
- Edges connect neighboring pixels
- States $\{1, \dots, K\}$ represent grey levels
- State space $\mathcal{S} = \{1, \dots, K\}^V$
- Image is represented by configuration $\sigma = (\sigma_1, \dots, \sigma_{|V|}) \in \mathcal{S}$
- We observe image \mathbf{Y} including noise, i.e.
 $\mathbf{Y} = \sigma + \varepsilon$, where $\varepsilon_1, \dots, \varepsilon_{|V|} \sim N(0, \delta^2)$

Problem: To reconstruct true image σ provided we observe \mathbf{Y} and assume that σ has prior distribution $C_\beta^{-1} e^{-\beta H(\sigma)}$.

Basic tool: Bayesian statistics and Markov chains including random walk on graph(s).

APPLICATION IN IMAGE ANALYSIS (CONT.)

Joint distribution of the vector (σ, \mathbf{Y}) is

$$\mathcal{L}(\sigma, \mathbf{Y}) \sim \frac{e^{-\beta H(\sigma)} \cdot \prod_{i \in V} \exp \left\{ - (Y_i - \sigma_i)^2 / 2\delta^2 \right\}}{\text{constant that depends on } (\sigma, \mathbf{Y})}$$

Posterior distribution is

$$\mathcal{L}(\sigma | \mathbf{Y}) \sim \frac{\exp \left[-\beta H(\sigma) + (2\delta^2)^{-1} \sum_{i \in V} (2Y_i \sigma_i - \sigma_i^2) \right]}{\text{function depending on } (\sigma, \beta, \mathbf{Y})}$$

What can we do:

- To generate from a posterior distribution $(\sigma | \mathbf{Y})$. Large enough sample representing configurations that can be considered likely representations of image.
- An alternative is to find the most likely image, i.e. to find the configuration $\hat{\sigma}$ that maximizes $P(\sigma | \mathbf{Y})$.

METROPOLIS ALGORITHM

Goal To generate reversible Markov chain with prescribed stationary distribution π , $\pi_i > 0 \forall i$, corresponding to the countable state space \mathcal{S}

- \mathbf{Q} be any **symmetric** matrix of transition probabilities q_{ij} on \mathcal{S}
- \mathbf{a} be any probability vector ($a_i \geq 0, \sum_i a_i = 1$) on \mathcal{S}

Alg. 1: Metropolis algorithm

- State S_{init} is randomly selected using \mathbf{a} [$Y_0 = S_{init}$]
- In each of the following steps, say step i when we are in state S_{cur}
 - First, select new candidate state S_{can} using \mathbf{Q}
 - Second, accept S_{can} as the next state with probability $\alpha = \min(1, \pi_{can}/\pi_{cur})$ [$Y_i = S_{can}$]; otherwise stay at S_{cur} [$Y_i = S_{cur}$]

Rem. 26: Markov chain constructed in this way, so called **Metropolis Markov chain for π with proposal matrix \mathbf{Q}** , has transition probabilities

$$p_{ij} = \begin{cases} q_{ij} \min(1, \pi_j/\pi_i) & j \neq i \\ 1 - \sum_{k \neq i} q_{ik} \min(1, \pi_k/\pi_i) & j = i \end{cases}$$

METROPOLIS – HASTINS ALGORITHM

Thm. 35: Assume that Q is an irreducible Markov chain on S , and π is strictly positive probability distribution on S . Then the Metropolis chain generated by the Metropolis algorithm is irreducible and reversible with respect to π .

METROPOLIS – HASTINGS ALGORITHM

Thm. 35: Assume that \mathbf{Q} is an irreducible Markov chain on \mathcal{S} , and π is strictly positive probability distribution on \mathcal{S} . Then the Metropolis chain generated by the Metropolis algorithm is irreducible and reversible with respect to π .

- If proposal matrix \mathbf{Q} is not symmetric, we can use, among others, so called **Metropolis – Hastings algorithm**

Alg. 3: Metropolis – Hastings algorithm Let $\pi, \pi_i > 0 \forall i$, be a “target” distribution on the discrete state space \mathcal{S} . Let \mathbf{Q} be any transition probability matrix on \mathcal{S} and define for any $i, j, i \neq j$,

$$t_{ij} = \pi_j q_{ji} / \pi_i q_{ij}$$

Let $A : [0, \infty] \rightarrow [0, 1]$ be any function such that $A(z) = zA(1/z) \forall z \in [0, \infty]$. Finally, define

$$p_{ij} = \begin{cases} q_{ij} A(t_{ij}) & j \neq i \\ 1 - \sum_{j \neq i} q_{ij} A(t_{ij}) & j = i \end{cases}$$

METROPOLIS – HASTINGS ALGORITHM CONT.

Show:

- Assume $q_{ij} > 0 \forall i, j$. Then \mathbf{P} is transition matrix of a Markov chain and is reversible with respect to π .
- If \mathbf{Q} is irreducible then \mathbf{P} is also irreducible.
- Assume that $q_{ij} = 0$ for some (i, j) . Then \mathbf{P} is still well defined and reversible with respect to π . However, \mathbf{P} may not be irreducible in this case even if \mathbf{Q} is irreducible.
- Show that $\min\{1, z\}$ can be used for $A(z)$, and that this leads to the classical Metropolis algorithm when \mathbf{Q} is symmetric. In what sense is this the “maximal” choice of $A(z)$? Intuitively, why this is important?
- What functions $A(z) = z^a / (1 + z^b)$ can be used?

EXPONENTIAL DISTRIBUTION – REPETITION

Def. 23: We say that random variable (rv) X follows exponential distribution ($X \sim \text{Exp}(\lambda)$), if corresponding density has the form

$$f(x; \lambda) = \begin{cases} \lambda e^{-\lambda x} & x > 0, \lambda > 0 \\ 0 & \text{otherwise} \end{cases}$$

Thm. 36: Let $X \sim \text{Exp}(\lambda)$. Then $F(x) = 1 - e^{-\lambda x}$, $E X = \lambda^{-1}$ and $\text{var } X = \lambda^{-2}$.

Thm. 37: No memory property. Let $X \sim \text{Exp}(\lambda)$. We interpret X as description of (random) life time of some unit. Then probability of the event that unit will survive time $y(> 0)$ conditioned by the event that survived time $x(> 0)$, does not depend on x , i.e.

$$P(X > x + y | X > x) = P(X > y), \quad \forall x, y > 0 \quad (10)$$

Rem. 27: Exponential distribution is the only continuous distribution for which (10) holds. Among discrete distributions the only one with the same property is geometric distribution.

INTENSITY FUNCTION

Def. 24: Let rv X has density $f(x)$ and df $F(x)$. Then the function

$$\Lambda(x) = \frac{f(x)}{1 - F(x)}, \quad x \in \mathbb{R}_1.$$

is called intensity function.

Rem. 28: Let rv X , which we interpret as a survival time of some process, has density $f(x)$ and distribution function $F(x)$. Then it holds:

$$\begin{aligned} P(x < X \leq x + \Delta | X > x) &= \frac{P(x < X \leq x + \Delta)}{P(X > x)} = \frac{F(x + \Delta) - F(x)}{1 - F(x)} \\ &= \frac{F(x + \Delta) - F(x)}{1 - F(x)} \frac{\Delta}{\Delta} \stackrel{\Delta \rightarrow 0}{\approx} \Delta \frac{f(x)}{1 - F(x)} \end{aligned}$$

Rem. 29: Let rv $X \sim \text{Exp}(\lambda)$. Then $\Lambda(x) = \lambda$. Exponential distribution is the only continuous distribution with constant intensity.

LINEAR PROCESS OF BIRTH AND DEATH

Assume a system with finitely or countably many states. Assume, moreover, that from a given state S_n we can move with non negligible probability only to the neighboring states, i.e.

- $S_n \rightarrow S_{n+1} \dots$ **birth**
- $S_n \rightarrow S_{n-1} \dots$ **death**
- As concerns other neighbors, we can move to them only with probability infinitesimally small.

Let the transitional probabilities in a small time interval $(t, t + h)$ are

- $P(S_n \rightarrow S_{n+1}) = \lambda_n h + o(h)$
- $P(S_n \rightarrow S_{n-1}) = \mu_n h + o(h)$
- $P(S_n \rightarrow S_{n \pm j}, j > 1) = o(h)$

LINEAR PROCESS OF BIRTH AND DEATH (CONT.)

Denote $P_n(t)$ probability of the event that system is in time t in state S_n .

Goal is to determine $P_n(t+h)$ and to find $p_n = \lim_{t \rightarrow \infty} P_n(t)$.

- Probabilities $P_n(t)$ satisfy following system of differential equations:

$$P'_0(t) = -\lambda_0 P_0(t) + \mu_1 P_1(t) \quad (11)$$

$$P'_n(t) = -(\lambda_n + \mu_n)P_n(t) + \mu_{n+1}P_{n+1}(t) + \lambda_{n-1}P_{n-1}(t), \quad n \geq 1$$

- If system is at time 0 at state S_i , then following initial conditions hold: $P_i(0) = 1$ and $P_n(0) = 0$ for $n \neq i$.
- Probabilities p_n exist, do not depend on initial conditions and satisfy system of linear equations

$$0 = -\lambda_0 p_0 + \mu_1 p_1 \quad (12)$$

$$0 = -(\lambda_n + \mu_n)p_n + \mu_{n+1}p_{n+1} + \lambda_{n-1}p_{n-1} \quad n \geq 1$$

which we receive if we set in (11) $P'_n(t) = 0$, $n \geq 0$, and replace $P_n(\infty)$ by p_n

LINEAR PROCESS OF GROWTH AND DEATH (CONT.)

Model

$$P(S_n \rightarrow S_{n+1}) = \lambda_n h + o(h)$$

$$P(S_n \rightarrow S_{n-1}) = \mu_n h + o(h)$$

$$P(S_n \rightarrow S_{n \pm j}, j > 1) = o(h)$$

Q	S_0	S_1	S_2	S_3	\dots
S_0	·	λ_0	·	·	·
S_1	μ_1	·	λ_1	·	·
S_2	·	μ_2	·	λ_2	·
S_3	·	·	μ_3	·	λ_3

Add diagonal element to Q so that **rows** sums to zero

Q^*	S_0	S_1	S_2	S_3	S_4	\dots	Σ
S_0	$-\lambda_0$	λ_0	0	0	\dots	\dots	0
S_1	μ_1	$-(\lambda_1 + \mu_1)$	λ_1	0	0	\dots	0
S_2	0	μ_2	$-(\lambda_2 + \mu_2)$	λ_2	0	\dots	0
S_3	0	0	μ_3	$-(\lambda_3 + \mu_3)$	λ_3	\dots	0

LINEAR PROCESS OF GROWTH AND DEATH (CONT.)

- $P_n(t)$ probability of the event that the system is at time t at state S_n
- $P'_0(t) = -\lambda_0 P_0(t) + \mu_1 P_1(t)$
- $P'_n(t) = -(\lambda_n + \mu_n) P_n(t) + \mu_{n+1} P_{n+1}(t) + \lambda_{n-1} P_{n-1}(t)$

$Q^{*\top}$	S_0	S_1	S_2	S_3	S_4	
S_0	$-\lambda_0$	μ_1	0	0	0	...
S_1	λ_0	$-(\lambda_1 + \mu_1)$	μ_2	0	0	...
S_2	0	λ_1	$-(\lambda_2 + \mu_2)$	μ_3	0	...
S_3	0	0	λ_2	$-(\lambda_3 + \mu_3)$	μ_3	...
...
Σ	0	0	0	0	0	...

System of retrospective Kolmogorov's differential equations has (matrix) form

$$P'(t) = Q^{*\top} P(t)$$

kde $P'(t) = (P'_0(t), P'_1(t), \dots)^\top$ a $P(t) = (P_0(t), P_1(t), \dots)^\top$

LINEAR PROCESS OF BIRTH AND DEATH (CONT.)

Model. Assume system composed of elements which can both split and vanish. Assume that in a small time interval of the length h probability of the event, that one element will split into two is equal to $\lambda h + o(h)$, and probability that will vanish (die) is equal to $\mu h + o(h)$, where λ and μ are constants characterizing behavior of the elements of considered system.

Rem. 30: When behavior of individuals (elements) of system is independent each from other, we have model of birth and death with parameters $\lambda_n = n\lambda$, $\mu_n = n\mu$.

Thm. 38: System of equations (11) for considered process has solution

$$P_0(t) = A(t)$$

$$P_n(t) = (1 - A(t))(1 - B(t))(B(t))^{n-1}, \quad n \geq 1$$

$$A(t) = \frac{\mu(e^{(\lambda-\mu)t} - 1)}{\lambda e^{(\lambda-\mu)t} - \mu} \quad \text{and} \quad B(t) = \frac{\lambda(e^{(\lambda-\mu)t} - 1)}{\lambda e^{(\lambda-\mu)t} - \mu}$$

TELEPHONE CENTRAL WITH INFINITELY MANY LINES

Ex. 13: Assume telephone central with infinitely many lines. We say that system is in state S_n if exactly n lines are occupied.

Assume moreover that:

- Probability of event that one telephone call will terminate during the interval $(t, t + h)$ is equal to $\mu h + o(h)$.
- Lengths of calls are mutually independent.
- Probability of event that in time interval $(t, t + h)$ new line will be occupied is equal to $\lambda h + o(h)$.

Tasks:

- Form system of differential equations for probabilities $P_n(t)$.
- Show that limit probabilities p_n follow Poisson distribution with parameter λ/μ .

TELEPHONE DOUBLE-BOOTH WITH UNLIMITED QUEUE

Ex. 14: Assume system that can serve at one moment at most two clients as, e.g., telephone double-booth. Clients that cannot be served form one unlimited queue. We say that the system is in state S_n if number of clients being served and in a queue is exactly n .

Moreover, we assume that:

- Probability of event that client which is served at time t will terminate call in interval $(t, t + h)$ is equal to $\mu h + o(h)$.
- Lengths of service times are independent.
- Probability that during interval $(t, t + h)$ will arrive new customer is equal to $\lambda h + o(h)$.

Tasks:

- Form system of differential equations for probabilities $P_n(t)$.
- Provided $\lambda < 2\mu$, show that for limit probabilities p_n holds

$$p_n = p_0 \left(\frac{\lambda}{\mu} \right)^n \frac{1}{2^{n-1}}, \quad \text{kde} \quad p_0 = \frac{2\mu - \lambda}{2\mu + \lambda}.$$

MODEL OF CAR PARK WITHOUT A QUEUE

Ex. 15: Assume car park with finite capacity N . State of the system is number of cars in car park. Queue is not formed.

Moreover, we assume that:

- Probability that car, which in time t parks, will depart in interval $(t, t + h)$ is equal to $\mu h + o(h)$.
- Lengths of staying at car park are independent.
- Probability that during interval $(t, t + h)$ will arrive new car is equal to $\lambda h + o(h)$.

Tasks:

- Form system of differential equations for probabilities $P_n(t)$.
- Show that limit probabilities p_n follow truncated Poisson distribution with parameter λ/μ .

TELEPHONE BOOTH WITH LIMITED QUEUE

Ex. 16: Assume system that can serve at one moment at most one client as, e.g., telephone booth. Clients that cannot be served form one queue of limited length N . We say that the system is in state S_n if number of clients being served and in a queue is exactly n .

Moreover, we assume that:

- Probability that client which is calling at time t will terminate in interval $(t, t + h)$ is equal to $\mu h + o(h)$.
- Lengths of calls are independent.
- Probability that during interval $(t, t + h)$ will arrive new customer is equal to $\lambda h + o(h)$.

Tasks:

- Form system of differential equations for probabilities $P_n(t)$.
- Show that for limit probabilities p_n it holds

$$p_n = \left(\frac{\lambda}{\mu}\right)^n p_0, \quad \text{kde} \quad p_0 = \mu^{N+1} \frac{\lambda - \mu}{\lambda^{N+2} - \mu^{N+2}}.$$

PROBLEM OF ONE REPAIRMAN AND MANY MACHINES

Ex. 17: Assume M machines serviced by one repairman. We say that system is in state S_n if exactly n machines are not working.

Moreover, we assume that:

- Probability that machine which is in time t repaired will start to work in interval $(t, t + h)$ is equal to $\mu h + o(h)$.
- Lengths of service times (repairs) are independent.
- Probability that machine, which is in time t working will break in interval $(t, t + h)$, is equal to $\lambda h + o(h)$.

Tasks:

- Form system of differential equations for probabilities $P_n(t)$.
- Show that limit probabilities p_n follow truncated Poisson distribution with parameter μ/λ , i.e. for $k = 1, \dots, M$ it holds

$$p_{M-k} = \frac{1}{k!} \left(\frac{\mu}{\lambda}\right)^k p_M, \quad \text{kde} \quad p_M = \left[1 + \sum_{k=1}^M \frac{1}{k!} \left(\frac{\mu}{\lambda}\right)^k\right]^{-1}$$

MODEL DESCRIBING WORK OF SEVERAL WELDERS

Ex. 18: Assume N welders which, independently each of other, take a current in random time intervals. We say that system is in state S_n if exactly n welders are working.

Moreover, we assume that:

- Probability that welder, which is in time t working, will stop welding in interval $(t, t + h)$ is equal to $\mu h + o(h)$.
- Lengths of service times (weldings) are independent.
- Probability that welder, which in time t does not work, will start working in interval $(t, t + h)$ is equal to $\lambda h + o(h)$.

Tasks:

- Form system of differential equations for probabilities $P_n(t)$.
- Show that limit probabilities p_n follow binomial distribution $Bi(N, \mu/(\mu + \lambda))$, i.e.

$$p_n = \binom{N}{n} \left(\frac{\mu}{\mu + \lambda} \right)^{N-n} \left(\frac{\lambda}{\mu + \lambda} \right)^n, \quad n = 0, 1, \dots, N.$$

MODEL OF KINETICS OF IRREVERSIBLE CHEMICAL REACTION

Ex. 19: Assume reagent A , molecules of which irreversibly change in molecules of product B (product). Speed of reaction is described by constant $\kappa > 0$. Let concentration of reagent A in time t is described by rv $X(t)$ and $X(0) = n_0$.

From physical principle assume that:

- Probability that one molecule will change in during $(t, t + h)$ provided $n_0 - n$ molecules changed up to time t , i.e. during interval $(0, t]$, is equal to $n\kappa h + o(h)$
- Probability of change of more than one molecule during interval $(t, t + h)$ is equal to zero.
- Reagent A and product B are statistically independent.
- Inverse reaction $B \rightarrow A$ arise with probability zero.

Tasks:

- Form system of differential equations for probabilities $P_n(t)$.
- Show that limit probabilities p_n follow binomial distribution $Bi(n_0, e^{-\kappa t})$.

COUNTING PROCESS

Def. 25: Stochastic process $N(t)$, $t \geq 0$ is called **counting process** if it represents overall number of “events” that occurred up to time t .

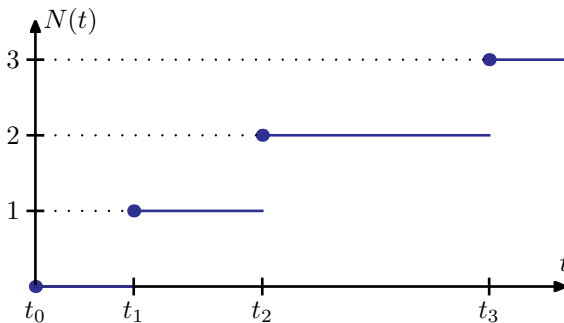
Rem. 31: For counting process must always hold that

- $N(t) \geq 0$.
- $N(t)$ is integer valued.
- If $s < t$, then $N(s) \leq N(t)$.
- For $s < t$ $N(t) - N(s)$ equals to the number of events that occur in the interval $(s, t]$.

Def. 26: Counting process $N(t)$, $t \geq 0$, is called **process with independent increments** if number of events observed in nonintersecting intervals are independent random variables.

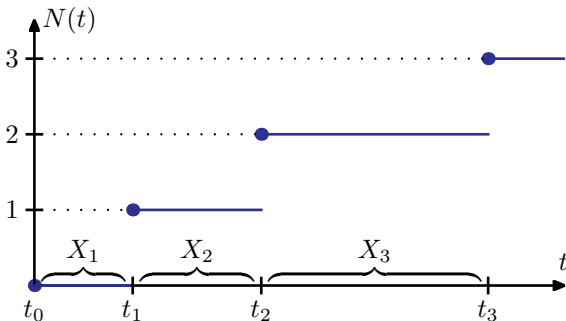
Def. 27: Counting process $N(t)$, $t \geq 0$, is called **process with stationary increments** if distribution of number of events in any interval depends only on its length and not on its placement.

COUNTING PROCESS



Course of counting process $N(t)$.

COUNTING PROCESS



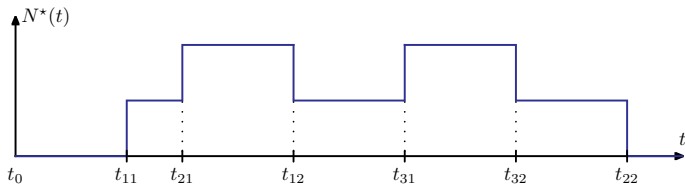
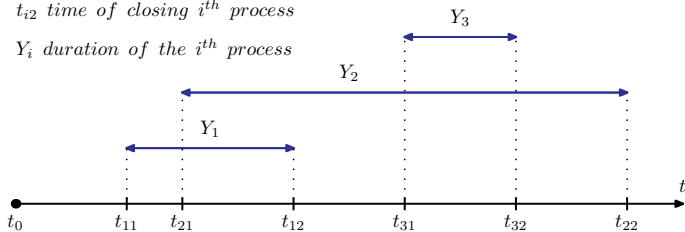
Times between arrivals of clients and corresponding counting process $N(t)$.

COUNTING PROCESS

t_{i1} time of opening i^{th} process

t_{i2} time of closing i^{th} process

Y_i duration of the i^{th} process



service times Y_i and course of the counting process describing occupancy of the system

HOMOGENEOUS POISSON PROCESS

Def. 28: Counting process $N(t)$, $t \geq 0$, is called **homogeneous Poisson process with intensity** λ , $\lambda > 0$ if:

- i $N(0) = 0$
- ii Process has independent increments.
- iii Number of events in any interval of length t follows Poisson distributin with mean λt , i.e.

$$P(N(t+s) - N(s) = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, \quad n = 0, 1, \dots$$

Def. 29: Counting process $N(t)$, $t \geq 0$, is called **homogeneous Poisson process with intensity** λ , $\lambda > 0$, if:

- i $N(0) = 0$
- ii Process has stationary and independent increments.
- iii $P(N(h) = 1) = \lambda h + o(h)$
- iv $P(N(h) \geq 2) = o(h)$

HOMOGENEOUS POISSON PROCESS – PROPERTIES

Thm. 39: Definitions 28 and 29 of Poisson process are equivalent.

Rem. 32: Conditions (ii) – (iv) can be replaced by the following set of equivalent conditions:

iii) Process has independent increments.

iv) $P(N(t+h) - N(t) = 1) = \lambda h + o(h)$

v) $P(N(t+h) - N(t) \geq 2) = o(h)$

Rem. 33: Poisson process has stationary increments and $EN(t) = \lambda t$.

Rem. 34: The fact that $N(t)$ follow Poisson distribution is a consequence of the approximation of binomial distribution by the Poisson one.

HPP – TIMES BETWEEN EVENTS

Def. 30: Assume Poisson process with intensity λ . Denote by T_n , $n = 1, 2, \dots$ time between n -th and $(n - 1)$ -st event, where $T_0 = 0$. Sequence $\{T_n\}$ is called **sequence of times between the events**.

Thm. 40: Sequence of times between the events follow exponential distribution with intensity λ .

Thm. 41: Assume sequence of times between events $\{T_n\}$. let $S_n = \sum_{i=1}^n T_i$. Then S_n follows gamma distribution $\Gamma(n, \lambda)$ with density

$$f_{S_n}(t; n; \lambda) = \frac{\lambda^n}{(n-1)!} t^{n-1} e^{-\lambda t} = \lambda \cdot e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!}$$

Hint. Recall renewal process (9), from where

$$N(t) \geq n \iff S_n \leq t$$

and differentiate corresponding distribution function.

HPP AS PROCESS OF BIRTH AND DEATH

Ex. 20: Assume system which changes according to some random influences as, e.g., telephone calls, radiation, etc. Denote $P_n(t)$ probability of event that during time interval of the length t we observed exactly n changes.

Assume, moreover:

- We observe stationary process, i.e. observed situation depend neither on placement on time horizon nor on its length.
- Not taking into account the number of events in interval $(0, t]$ let probability of one change in interval $(t, t + h)$ be $\lambda h + o(h)$, and probability of more than one event be $o(h)$.

Rem. 35: Notice that changes in intervals $(0, t]$ and $(t, t + h)$ are independent.

Tasks:

- Form differential equations for probabilities $P_n(t)$.
- Show that probabilities $P_n(t)$ follow Poisson distribution with parameter λt .

NONHOMOGENOUS POISSON PROCESS (NHPP)

Def. 31: Counting process $N(t)$, $t \geq 0$ is called **homogeneous Poisson process with intensity** λ , $\lambda > 0$, if:

- i) $N(0) = 0$
- ii) Process has stationary and independent increments
- iii) $P(N(h) = 1) = \lambda h + o(h)$
- iv) $P(N(h) \geq 2) = o(h)$

Def. 32: Counting process $N(t)$, $t \geq 0$ is called **nonhomogeneous Poisson process with intensity function** $\lambda(t) > 0$, $t > 0$, if:

- i) $N(0) = 0$.
- ii) Process $N(t)$, $t \geq 0$ has independent increments
- iii) $P(N(t+h) - N(t) = 1) = \lambda(t)h + o(h)$
- iv) $P(N(t+h) - N(t) \geq 2) = o(h)$

NHPP – PROPERTIES

Thm. 42: Let $N(t)$ and $M(t)$ be independent NHPP with intensity functions $\lambda(t)$ and $\mu(t)$. Then for random process $R(t) = N(t) + M(t)$ holds:

- i) It is NHPP with intensity $\kappa(t) = \lambda(t) + \mu(t)$
- ii) Let in time t appeared an event of the process $R(t)$. Then, independently what happened up to time t , probability that it has been realization of the process $N(t)$ is equal $\lambda(t)/(\lambda(t) + \mu(t))$

Thm. 43: Assume process $M/G/\infty$, i.e. the arrivals of clients are described by Poisson process, infinity many of channels and general distribution of service times. Then output from this system is NHPP with intensity $\lambda(t) = \lambda G(t)$.

COMPOUND POISSON PROCESS

Random process $X(t)$ is called **Compound Poisson Process** if it can be represented as

$$X(t) = \sum_{i=1}^{N(t)} Y_i, \quad t \geq 0,$$

where $N(t)$, $t \geq 0$ is Poisson process and Y_i are iid rv's independent of $N(t)$

- If $Y_i = 1$ then $X(t) \equiv N(t)$
- $E X(t) = \lambda t E Y_1$ and $var X(t) = \lambda t E Y_1^2$
- If customers arrive in cars randomly according to some HPP and numbers of customers in each car follow the same distribution, and are independent on $N(t)$, then $X(t)$ is compound Poisson process describing number of customers that arrived until the time t
- If customers leave supermarket randomly according to some HPP and the amount of money spent in the supermarket follows the same distribution, and are independent on $N(t)$, then $X(t)$ is compound Poisson process describing amount of money spent up to time t

QUEUING SYSTEM

Def. 33: **Queuing system** will be characterized by:

- One or more parallel service station(s), to which arrive customers. When serving is finished, customer leaves system and next one in line, if there is any, enters service.
- Clients that cannot be served (because system is fully occupied) form one queue (line).
- Times between arrivals are iid rv's with distribution A .
- Service times (time in a queue is not included) are iid rv's with distribution B .

Rem. 36: Distribution of arrival/service times is usually one of:

- exponential ... M (Markovian)
- deterministic ... D (Deterministic)
- general ... G (General)
- Erlang $\Gamma(n, \lambda)$... S_n (Erlang)

M/M/1, M/M/c AND M/M/∞

Def. 34: Queuing system **M/M/x** is characterized by the fact that arrivals of clients follow homogeneous Poisson process with intensity λ and service times follow exponential distribution.

Thm. 44: System **M/M/x** can be described by the general process of birth and death.

Rem. 37: For model:

- **M/M/1** : $\lambda_j = \lambda$ & $\mu_j = \mu \forall j \in \mathbb{N}_0$
- **M/M/c** : $\lambda_j = \lambda \forall j \in \mathbb{N}_0$ & $\mu_j = j\mu, 0 \leq j < c, \mu_j = c\mu, c \leq j < \infty$
- **M/M/∞** : $\lambda_j = \lambda$ & $\mu_j = j\mu \forall j \in \mathbb{N}_0$

For examples and details see models describing telephone central, telephone booths, etc.

M/M/1, M/M/c AND M/M/∞

Thm. 45: For systems **M/M/x** it holds:

- **M/M/1** : If $\lambda/\mu < 1$ then limit probabilities p_n follow geometric distribution with parameter $1 - \lambda/\mu$
- **M/M/c** : If $\lambda/(c\mu) < 1$ then limit probabilities p_n follow truncated Poisson distribution with parameters $(c + 1, \lambda/\mu)$
- **M/M/∞** : Limit probabilities p_n follow Poisson distribution with parameter λ/μ

Thm. 46: For system **M/M/c** it holds that departures from the stabilized system with unlimited queue without any departures from it with intensities λ (input) and μ (output) are described by homogeneous Poisson process with parameter $\lambda!$

Rem. 38: Systems **M/M/c** may be “easily” combined and under the assumption of stability can be described by appropriate Markov process.

M/M/1

Rem. 39: Assume model **M/M/1**. We know from Theorem 45 that limit (stationary) probabilities p_n describing number of clients in steady state follows geometric distribution with parameter $1 - \lambda/\mu$. It follows from basic properties of geometric distribution that:

- Mean of rv describing number of clients in system is $\frac{\lambda}{\mu} / \left(1 - \frac{\lambda}{\mu}\right)$
- Variance of rv describing number of clients in system is $\frac{\lambda}{\mu} / \left(1 - \frac{\lambda}{\mu}\right)^2$
- Mean length of the queue is $\sum_j j p_{j+1} = \left(\frac{\lambda}{\mu}\right)^2 / \left(1 - \frac{\lambda}{\mu}\right)$

Rem. 40: Notice that difference between mean number of clients in system and mean length of the queue is λ/μ , not 1. Why?

M/M/1 (CONT.)

Rem. 41: Assume model **M/M/1**.

Then time T_n , which client, who entered as the n -th one, will spend in the system follow gamma distribution $\Gamma(n + 1, \mu)$, because it consists of the remaining time of client who is served and service times of waiting clients including the entering one.

Distribution of the waiting time T of randomly chosen client is, according to the complete probability theorem, mixture of distributions of T_n with weights given by the steady state probabilities p_n . Show that it holds:

$$P(T \leq t) = 1 - e^{-(\mu-\lambda)t}, \quad t \geq 0$$

so that the mean time spent in the system is equal to $\frac{1}{\mu-\lambda}$.

SIMULATION OF M/M/1

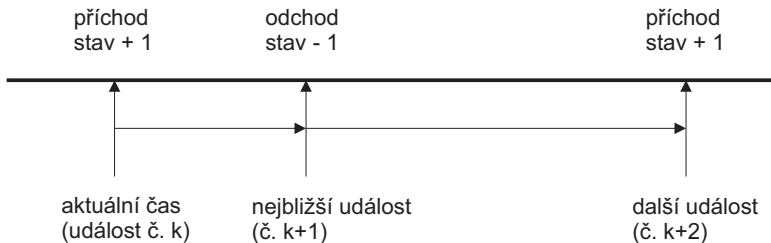
```
for (i in 1:maximální počet zákazníků){
  if (aktcas >= maximální čas){break}
  # simulace končí, pokud je dosaženo maximálního času

  while (min(časy naplánovaných událostí) < čas příchodu
    dalšího zákazníka){
    # někdo skončí obsluhu dřív, než dorazí další zákazník
    aktcas <- min(časy naplánovaných událostí)
    stav <- stav - 1
    záznam události, odebrání zákazníka z`obsluhy

    # začátek obsluhy dalšího zákazníka
    # (uvolnila se obslužná stanice)
    if (fronta není prázdná){
      délka obsluhy <- obsluha(aktcas,stav,stanice)
      začátek obsluhy zákazníka, aktualizace fronty
    }
  }
  # nejbližší událostí je příchod i-tého zákazníka
  aktcas <- čas příchodu dalšího zákazníka
  stav <- stav + 1
  záznam události

  # pokud je volná stanice, bude zákazník rovnou obsluhován
  if (některá stanice je volná){začátek obsluhy zákazníka}
  else{zařazení zákazníka na konec fronty}

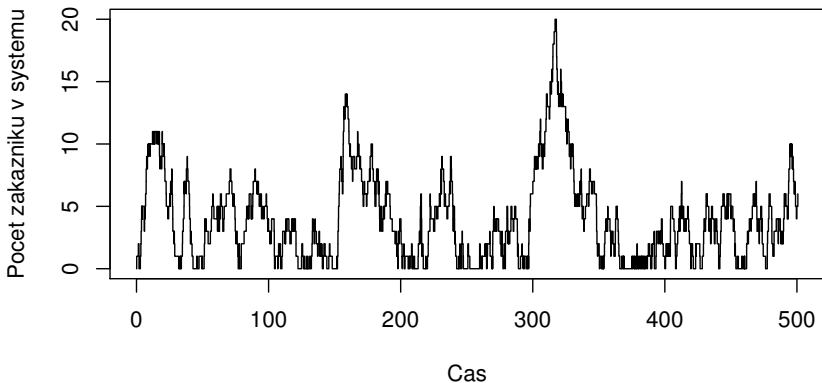
  # určení času příchodu dalšího zákazníka
  příchod dalšího zákazníka <- aktcas + prichod(aktcas,stav)}
```


M/M/1

Scheme of considered simulation algorithm.

M/M/1

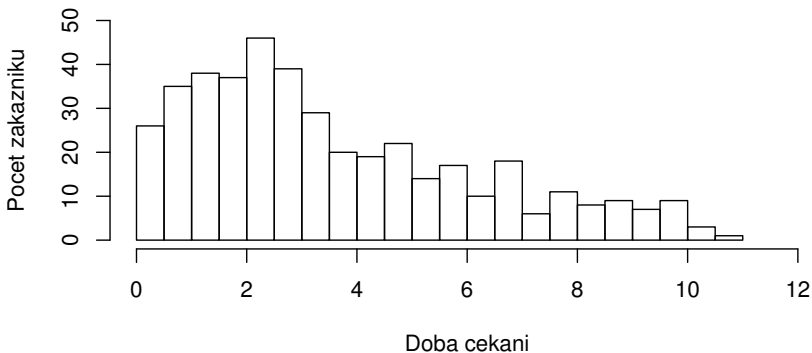
Vyvo systemu v case



Evolution of number of clients in system M/M/1 with parameters $\lambda = 1$ a $\mu = 1.2$

M/M/1

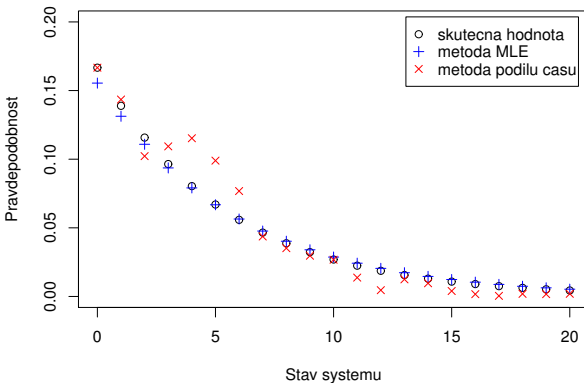
Histogram dob cekani aka niku, kteri museli na acatek obsluhy cekat kladnou dobu



Histogram of waiting times of clients that had to wait positive time in M/M/1 system with parameters $\lambda = 1$ and $\mu = 1.2$.

M/M/1

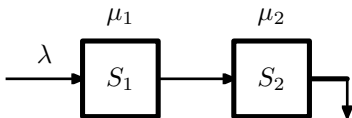
Odhady pravdepodobnosti stacionarniho ro deleni



Comparison of estimators of stationary distribution in system M/M/1 with parameters $\lambda = 1$ a $\mu = 1.2$.

M/M/1 + M/M/1 \equiv TANDEM

Def. 35: Serial connection of two **independent** queuing systems **M/M/1** is called **tandem**



Thm. 47: In tandem connecting two independent queuing systems **M/M/1** such that $\rho_i = \lambda/\mu_i < 1$, $i = 1, 2$, there exists a stationary distribution π and is described by a product of stationary distributions for individual nodes, i.e.

$$\pi(k_1, k_2) = \prod_{i=1}^2 \pi_i(k_i) = \prod_{i=1}^2 \rho_i^{k_i} (1 - \rho_i) \quad k_i \in \mathbb{N}_0$$

Rem. 42: Similar result holds for a serial connection of **J independent** queuing systems obsluhy **M/M/c_i**, $i = 1, \dots, J$, if $\lambda/(c\mu_i) < 1$.

JACKSON(IAN) NETWORKS

A network of m interconnected queues is known as a Jackson(ian) open network if it meets the following conditions:

- 1 Arrivals to the system follows homogeneous Poisson process with intensity $\alpha > 0$; and in the same time each customer is randomly assigned to the node j with probability $p_{0j} \geq 0$, $\sum_{j=1}^J p_{0j} = 1$
- 2 Any external arrivals to node i form a Poisson process
- 3 All service times are exponentially distributed and the service discipline at all queues is first-come, first-served
- 4 Customer completing service at queue i will either move to some new queue j with probability p_{ij} or leave the system with probability $p_{i0} = 1 - \sum_{j=1}^m p_{ij}$, which is non-zero for some subset of the queues (nodes)
- 5 Utilization of all of the queues is less than one (server utilization means the proportion of time the server is busy)

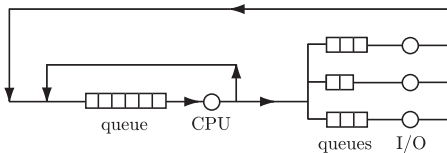
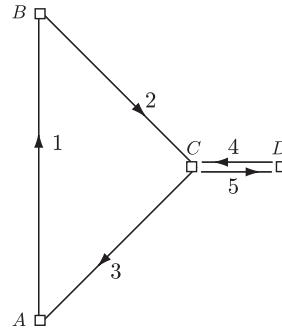
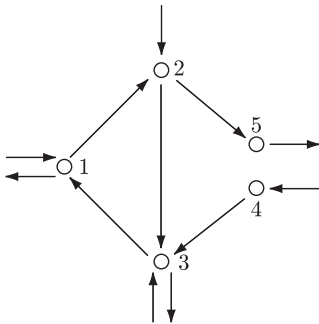
JACKSON(IAN) NETWORKS

A network of m interconnected queues is known as a Jackson(ian) open network if it meets the following conditions:

- 1 Arrivals to the system follows homogeneous Poisson process with intensity $\alpha > 0$; and in the same time each customer is randomly assigned to the node j with probability $p_{0j} \geq 0$, $\sum_{j=1}^J p_{0j} = 1$
- 2 Any external arrivals to node i form a Poisson process
- 3 All service times are exponentially distributed and the service discipline at all queues is first-come, first-served
- 4 Customer completing service at queue i will either move to some new queue j with probability p_{ij} or leave the system with probability $p_{i0} = 1 - \sum_{j=1}^m p_{ij}$, which is non-zero for some subset of the queues (nodes)
- 5 Utilization of all of the queues is less than one (server utilization means the proportion of time the server is busy)

Rem. 44: $\lambda_i = \alpha p_{0i} + \sum_{j=1}^J \lambda_j p_{ji}$, $i = 1, \dots, J$

JACKSON(IAN) NETWORKS



Opened and closed Jackson(ian) networks

JACKSON(IAN) OPEN NETWORK FOR MODEL M/M/C

Rem. 45: Recall, that for a model:

- **M/M/1** : $\lambda_j = \lambda$ & $\mu_j = \mu \quad \forall j \in \mathbb{N}_0$

If $\lambda/\mu < 1$ then stationary distribution p_n follows geometric distribution with parameter $1 - \lambda/\mu$

- **M/M/c** : $\lambda_j = \lambda \quad \forall j \in \mathbb{N}_0$ & $\mu_j = j\mu, \quad 0 \leq j < c, \quad \mu_j = c\mu, \quad c \leq j < \infty$

If $\lambda/(c\mu) < 1$ then stationary distribution p_n follows truncated Poisson distribution with parameters $(c + 1, \lambda/\mu)$

Thm. 48: In opened Jackson(ian) network composed of J nets **M/M/1** such that $\rho_i = \lambda/\mu_i < 1, \quad i = 1, \dots, J$, there exists stationary distribution π and is given by a product of individual stationary distributions for respective nets, i.e.

$$\pi(k_1, \dots, k_J) = \prod_{i=1}^J \pi_i(k_i) = \prod_{i=1}^J \rho_i^{k_i} (1 - \rho_i), \quad k_i \in \mathbb{N}_0$$

Rem. 46: Similar result holds for opened Jackson(ian) network of J queues of the type **M/M/c_i**, $i = 1, \dots, J$, provided $\lambda_i/(c_i\mu_i) < 1$

COST EQUILIBRIUM

Assume stationary system

- $N(t)$... number of clients that entered system up to time t
- λ_a ... average arrival rate of entering customers
- ARE ... average rate at which system earns
- AAP ... average amount an entering customer pays

Cost equilibrium

- $\lambda_a = \lim_{t \rightarrow \infty} \frac{N(t)}{t}$
- $ARE = \lambda_a \cdot AAP$

(Recall $S_N = \sum_{i=1}^N X_i$, where both X_i and N are random)

LITTLE'S LAW (FORMULA)

- ARE ... average rate at which system earns
- AAP ... average amount an entering customer pays
- L ... mean number of clients in system
- L_Q ... mean number of clients in queue
- W ... mean time client spends in system
- W_Q ... mean time client spends in queue

- Assume that each client will pay 1 US for unit of time spent in system. Then from cost equilibrium $ARE = \lambda_a \cdot AAP$ we get

$$L = \lambda_a W$$

- If client pays 1 US for unit of time spent in queue, the

$$L_Q = \lambda_a W_Q$$

- If client pays 1 US for unit of time spent in service, then

$$\text{average number of clients in service} = \lambda_a E(S)$$

where $E(S)$ is mean time client spends in service.

Rem. 47: These relations are independent of distribution of arrival times, service times, number of servers, order of service, etc.

BRANCHING PROCESS ALIAS HOW VIRUSES CAN PROPAGATE

Model. Assume individuals that can give origin to new individuals of the same type with probabilities

$$P(U = j) = p_j, \quad j = 0, 1, 2, \dots \quad (13)$$

and probability generating function (PGF) $P(x) = \sum_{j=0}^{\infty} p_j x^j$.

At the beginning (generation 0) there exists only one individual. Let immediate descendants of n -th generation form $(n + 1)$ -st generation and let descendants behave independently each from other.

Problems.

- Find distribution of members in the n -th generation.
- Find limit ($n \rightarrow \infty$) probability of event that population will die out.

BRANCHING PROCESS (CONT.)

Model. Let $X_0 = 1$, $X_1 (\equiv U)$ follow (13) with PGF $P_1(x) \equiv P(x)$. Let X_n describes number of elements in n -th generation.

Let number of descendants of each of the X_1 elements of first generation is again rv with distribution (13). Let these rv's are mutually independent and independent on X_1 , i.e.

$$X_2 = U_1 + \dots + U_{X_1}$$

Analogously, X_3 are descendants of the second order of the individuals of first generation, i.e. X_1 rv's with distribution as X_2 , resp. descendants of X_2 rv's with distribution as X_1 , i.e. (13) etc.

Thm. 49: Probability generation function $P_n(x)$ of X_n fulfil recurrent relation

$$P_{n+1}(x) = P(P_n(x)) = P_n(P(x))$$

and, moreover, it holds

$$E X_n = (E X_1)^n, \quad \text{var} X_n = \begin{cases} \sigma^2 \mu^{n-1} \left(\frac{1-\mu^n}{1-\mu} \right), & \mu \neq 1, \\ n\sigma^2, & \mu = 1, \end{cases} \quad n = 1, 2, \dots$$

MARKOV PROCESSES WITH DISCRETE STATES AND CONTINUOUS TIME

Def. 36: Markov processes with discrete states and continuous time

will be for us random process which moves from one state to another according to some Markov chain and the times which spends in different states follow exponential distribution. Moreover, (random) times it spends in respective states are independent rv's.

Rem. 48: By the other words, it is a random process such that:

- Time spend in state i follow exponential distribution with mean say $1/v_i$.
- Process leaves state i and enters to state j with probability P_{ij} , and it holds

$$P_{ii} = 0 \quad \forall i \quad \& \quad \sum_j P_{ij} = 1 \quad \forall i.$$

BASIC CHARACTERIZATIONS OF RANDOM VARIABLES

Rem. 49: Recall that any random variable (rv) can be unambiguously characterized by the following characteristics:

- density $f(x)$, $x \in \mathbb{R}_1$
- distribution function $F(x) = P(X \leq x)$, $x \in \mathbb{R}_1$
- characteristic function $\varphi(t) = E e^{itX}$, $t \in \mathbb{R}_1$

Some “special” types of random variables (rv’s) can be also characterized by another characteristics as, e.g.

- Integer valued rv’s are unambiguously characterized by corresponding **generating function** $P(x) = \sum_{i=0}^{\infty} p_i x^i$, $x \in \mathbb{R}_1$
- Non-negative rv’s are unambiguously characterized by corresponding **reliability function (survival function)** $R(x) = P(X > x) = 1 - F(x)$, $x \in \mathbb{R}_1$, or **intensity function** $\lambda(x) = \frac{f(x)}{1-F(x)}$, $x > 0$, or **cumulative intensity function** $\Lambda(x) = \int_0^x \lambda(t) dt$

BASIC CHARACTERIZATIONS AND CHARACTERISTICS OF RANDOM VARIABLES

Another useful characterizations can be obtained using

- characteristic function $\varphi(t) = E e^{itX}$, $t \in \mathbb{R}_1$, and its logarithm
- moment generating function $m(t) = E e^{tX}$, $t \in \mathbb{R}_1$, and its logarithm

CONDITIONAL PROBABILITY

Def. 37: Assume probability space $(\Omega, \mathcal{A}, \mathcal{P})$. Let events $A, B \in \mathcal{A}$ and $\mathcal{P}(B) > 0$. Then **conditional probability** of event A under condition B is defined as

$$\mathcal{P}(A|B) = \frac{\mathcal{P}(A \cap B)}{\mathcal{P}(B)}$$

CONDITIONAL CHARACTERISTICS FOR DISCRETE RV'S

Def. 38: Let X and Y are two discrete rv's. Then:

- **conditional density of X** given that $Y = y$ is defined by

$$p_{X|Y}(x|y) = \mathcal{P}(X = x|Y = y) = \frac{\mathcal{P}(X = x, Y = y)}{\mathcal{P}(Y = y)} = \frac{p_{X,Y}(x, y)}{p_Y(y)}$$

- **conditional distribution function of X** given that $Y = y$ is defined by

$$F_{X|Y}(x|y) = \mathcal{P}(X \leq x|Y = y) = \sum_{z \leq x} p_{X|Y}(z|y)$$

- **conditional expectation of X** given that $Y = y$ is defined by

$$E[X|Y = y] = \sum_x x \cdot \mathcal{P}(X = x|Y = y) = \sum_x x \cdot p_{X|Y}(x|y)$$

- **conditional variance of X** given that $Y = y$ is defined by

$$\text{var}[X|Y = y] = E\left[(X - E[X|Y = y])^2 | Y = y\right]$$

CONDITIONAL CHARACTERISTICS FOR CONTINUOUS RV'S

Def. 39: Let X and Y are two continuous rv's. Then:

- **conditional density of X** given that $Y = y$ is defined by

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)}$$

- **conditional distribution function of X** given that $Y = y$ is defined by

$$F_{X|Y}(x|y) = \mathcal{P}(X \leq x | Y = y) = \int_{-\infty}^x p_{X|Y}(z|y) dz$$

- **conditional expectation of X** given that $Y = y$ is defined by

$$E[X|Y = y] = \int_{-\infty}^{\infty} x \cdot f_{X|Y}(x|y) dx$$

- **conditional variance of X** given that $Y = y$ is defined by

$$\text{var}[X|Y = y] = E\left[\left(X - E[X|Y = y]\right)^2 | Y = y\right]$$

GENERATING FUNCTIONS

Def. 40: Let a_0, a_1, \dots be a sequence of real numbers. If a series $\mathcal{A}(x) = \sum_{j=0}^{\infty} a_j x^j$ converge in some neighborhood of zero, we call $\mathcal{A}(x)$ corresponding *generating function*.

Rem. 50: If $\{a_j\}$ is bounded, then $\mathcal{A}(x)$ evidently converge at least in the interval $(-1, 1)$.

Def. 41: If X is integer valued rv for which $P(X = j) = p_j \geq 0$, $j = 0, 1, \dots$, $\sum_j p_j = 1$, then its (*probability*) *generating function* has the form $\mathcal{P}(x) = \sum_{j=0}^{\infty} p_j x^j$.

Rem. 51:

- Generating function $\mathcal{P}(x)$ unambiguously characterize corresponding random variable X .
- Notice that $\mathcal{P}(t) = E t^X$
 [recall that $E X = \sum_j p_j x_j$ and $E g(X) = \sum_j p_j g(x_j)$].
- For integer valued rv X corresponding generating function always converge in point $x = 1$, because $\mathcal{P}(1) = \sum_j p_j = 1$.

CONVOLUTION

Def. 42: Let a_0, a_1, \dots and b_0, b_1, \dots are two sequences of real numbers. Then a sequence c_0, c_1, \dots defined by relation

$$c_n = a_0 b_n + a_1 b_{n-1} + \dots + a_n b_0, \quad n = 0, 1, \dots$$

is called **convolution** of sequences $\{a_j\}$ and $\{b_j\}$. We will write

$$\{c_j\} = \{a_j\} \star \{b_j\}$$

Thm. 50: Let $\{a_j\}$ and $\{b_j\}$ are two sequences with generating functions $\mathcal{A}(x)$ and $\mathcal{B}(x)$. Then for the generating function corresponding to their convolution $\{c_j\}$ it holds

$$C(x) = \mathcal{A}(x)\mathcal{B}(x).$$

Rem. 52: Convolution of a sequence $\{a_j\}$ with itself is called convolution power and is denoted $\{a_j\}^{2\star}$. Analogously, n -th convolution power $\{a_j\} \star \dots \star \{a_j\}$ will be denoted $\{a_j\}^{n\star}$.

CONVOLUTION

Thm. 51: Let X_1, X_2, \dots, X_n are **independent** identically distributed (iid) rv's with integer valued distribution described by probabilities $\{p_j\}$. Denote corresponding generating function by $\mathcal{P}(x)$. Then distribution of their sum, i.e. distribution of $X_1 + X_2 + \dots + X_n$ is described by the n -th convolution power $\{p_j\}^{n\star}$, and corresponding generating function has the form

$$\mathcal{P}(x) \dots \mathcal{P}(x) = \mathcal{P}^n(x)$$

GENERATING FUNCTIONS – EXAMPLES

Ex. 21: Check the form of generating functions for following most important discrete distributions:

- Alternative ... $\mathcal{P}(x) = q + px$
- Binomial ... $\mathcal{P}(x) = (q + px)^n$
- Poisson ... $\mathcal{P}(x) = \exp\{-\lambda + \lambda x\}$
- Geometrical ... $\mathcal{P}(x) = p/(1 - qx)$
resp. = $px/(1 - qx)$
- Negative binomial ... $\mathcal{P}(x) = (p/(1 - qx))^r$
resp. = $(px/(1 - qx))^r$
- Discrete uniform ... $\mathcal{P}(x) = (1 - x^{n+1})/((n + 1)(1 - x))$
resp. = $(x(1 - x^n))/(n(1 - x))$

Using these generating functions calculate corresponding expectations and variances.

Rem. 53: Recall that geometric and negative binomial distributions are the simplest models describing waiting times.

PROPERTIES OF GENERATING FUNCTIONS

Thm. 52: Denote $q_k = P(X > k) = \sum_{j>k} p_j$, $k = 0, 1, 2, \dots$ and corresponding generating function $Q(x) = \sum_{j=0}^{\infty} q_j x^j$. Then for $-1 < x < 1$ it holds that $Q(x) = (1 - \mathcal{P}(x))/(1 - x)$.

Thm. 53: For integer valued random variable X it holds that

$$E X = \sum_{i=0}^{\infty} i p_i = \sum_{j=0}^{\infty} q_j = \mathcal{P}'(1) = Q(1)$$

Thm. 54: Let generating function $\mathcal{P}(x)$ of integer valued rv X has radius of convergence larger than one. Then it holds

$$\text{var } X = \mathcal{P}''(1) + \mathcal{P}'(1) - (\mathcal{P}'(1))^2 = 2Q'(1) + Q(1) - (Q(1))^2$$

Rem. 54: Thm. 54 holds also in the case when radius of convergence is equal to 1, provided $\lim_{x \rightarrow 1-} Q''(x) < \infty$ and derivatives in point $x = 1$ are replaced by their limits for $x \rightarrow 1-$.

PARTIAL FRACTION DECOMPOSITION

Rem. 55: Knowledge of $\mathcal{P}(x)$ is theoretically equivalent to knowledge of $\{p_j\}$, and vice versa. However, the use of the fact that $p_j = \mathcal{P}^{(j)}(0)/j!$ can be quite complicated in practice. In such a case following approximation can be useful.

Thm. 55: Let generating function $\mathcal{P}(x)$ of the sequence $\{p_n\}$ can be written in the form $\mathcal{P}(x) = U(x)/V(x)$, where $U(x)$ and $V(x)$ are polynomials without common roots, order of $U(x)$ is smaller than order of $V(x)$, and the roots of polynomial $V(x)$ are simple. Then

$$p_n = \frac{\rho_1}{x_1^{n+1}} + \cdots + \frac{\rho_m}{x_m^{n+1}}, \quad 0 \leq n < \infty,$$

where m is order of polynomial $V(x)$, x_1, \dots, x_m are its roots and $\rho_k = -U(x_k)/V'(x_k)$, $1 \leq k \leq m$.

Rem. 56: For calculation of ρ_k , $1 \leq k \leq m$, can be used decomposition into partial fractions, embedded in programs *Maple* or *Mathematica*, e.g.

PARTIAL FRACTION DECOMPOSITION

Rem. 57: Assume that x_1 is that root of $V(x)$ for which $|x_1| < x_k$, $2 \leq k \leq m$. Then

$$p_n = \frac{\rho_1}{x_1^{n+1}} \left(1 + \frac{\rho_2}{\rho_1} \left(\frac{x_1}{x_2} \right)^{n+1} + \dots + \frac{\rho_m}{\rho_1} \left(\frac{x_1}{x_m} \right)^{n+1} \right), \quad (14)$$

so that for $n \rightarrow \infty$ it holds that $p_n \approx \rho_1/x_1^{n+1}$, where $\rho_1 = -U(x_1)/V'(x_1)$.

Rem. 58: For validity of the assertion $p_n \approx \rho_1/x_1^{n+1}$ it is possible to omit assumption that degree of $U(x)$ is smaller than degree of $V(x)$. Instead it is sufficient to assume only that the root x_1 is unique. Moreover, recall that practical experience shows that the approximation (14) is satisfactory even for small values of n .

Ex. 22: Let q_n be probability that in the sequence of n trials with dichotomous response (T, F) a subsequence FFF will not occur. Find corresponding generating function and calculate corresponding probabilities q_n both precisely and using the above approximations.

Solution: $Q(x) = (8 + 4x + 2x^2)/(8 - 4x - 2x^2 - x^3)$.

COMPOUND DISTRIBUTIONS

Thm. 56: Let X_1, X_2, \dots and N are independent integer valued rv's, X_i 's have the same distribution $\{f_j\}$ and N has distribution $\{g_j\}$.

Then $S_N = X_1 + \dots + X_N$ is also integer valued random variable with the distribution $\{h_j\}$, where

$$h_j = P(S_N = j) = \sum_{n=0}^{\infty} g_n \cdot \{f_j\}^{n*}.$$

If $\mathcal{A}(x), \mathcal{B}(x)$ and $\mathcal{C}(x)$ are generating functions corresponding to the sequences $\{f_j\}, \{g_j\}$ and $\{h_j\}$, then $\mathcal{C}(x) = \mathcal{B}(\mathcal{A}(x))$ and $E S_N = E X_1 \cdot E N$. Corresponding variance σ can be calculated using Thm. 54.

Rem. 59: Notice that random variable $S_N = X_1 + \dots + X_N$ is nothing else than **random sum of random variables**.

Ex. 23: Let number N of laid eggs follow Poisson distribution $Po(\lambda)$ and probability of arrival of individual from an egg is p , i.e. X_i follow alternative distribution. Show that in such a case S_N follows Poisson distribution $Po(\lambda \cdot p)$.