MAI 060
REMARKS AND EXAMPLES
(text will be steadily completed)

JAROMÍR ANTOCH

January 1, 2019

Main goals of the lecture

Main goal of this lecture is study of Markov chains and Markov processes, their generalizations and applications.
We will concentrate (in more or less details) on:

- Notion of recurrence in theory of probability
- Random walks
- Markov chains with discrete states and discrete time
- Markov processes with discrete states and continuous time
- Birth and death models
- Poisson process
- Durbin – Watson process
- Basics of queuing models

EVENTS TO BE CONSIDERED

Assume a sequence of repeated (not necessarily independent) trials, each of them having the same finite or countable set of possible outcomes $\{E_j\}_{j\in\mathcal{J}}$, where usually $\mathcal{J} \equiv \mathbb{N}_0$, $\mathbb{N}$ or $\mathbb{Z}$. Let

$$\left\{E_{j_1}, E_{j_2}, \dots, E_{j_n}\right\} \tag{1}$$

denotes an 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 (1):

- $P\Big(E_{j_1}, \dots, E_{j_{n-1}}\Big) = \sum_{k=0}^{\infty} P\Big(E_{j_1}, \dots, E_{j_{n-1}}, E_k\Big), \ 1 < n < \infty$
- About each sequence (1) it can be decided whether it has or does not have a "property ξ"

Rem. 1: Recall that if the events $\left\{E_{j_1}, E_{j_2}, \dots, E_{j_n}\right\}$ are independent, then

$$P\Big(\Big\{E_{j_1}, E_{j_2}, \dots, E_{j_n}\Big\}\Big) = \prod_{k=1}^{n} P\Big(E_{j_k}\Big) \tag{2}$$

Recurrent events

Def. 1: Statement ξ occurred on the n-th place of (finite or infinite) sequence $E_{j_1}, E_{j_2}, \ldots$ means that the sequence $E_{j_1}, E_{j_2}, \ldots, E_{j_n}$ has the property ξ.

Def. 2: We will call property ξ a recurrent event, if:

1. ξ occurred on the n-th and $n+m$-th place place of the sequence $E_{j_1}, E_{j_2}, \ldots, E_{j_{n+m}}$ if and only both the sequence $E_{j_1}, \ldots, E_{j_n}$ and the sequence $E_{j_{n+1}}, \ldots, E_{j_{n+m}}$ have property ξ
2. In such a case it holds:

$$P\Big(\underbrace{\overbrace{E_{j_1}, \ldots, E_{j_n}}^{A}}_{\xi}, \underbrace{\overbrace{E_{j_{n+1}}, \ldots, E_{j_{n+m}}}^{B}}_{\xi}\Big) = P\Big(\underbrace{\overbrace{E_{j_1}, \ldots, E_{j_n}}^{A}}_{\xi}\Big) \cdot P\Big(\underbrace{\overbrace{E_{j_{n+1}}, \ldots, E_{j_{n+m}}}^{B}}_{\xi}\Big)$$

Examples of recurrent events

Ex. 1: Assume a sequence of independent trials with dichotomous (alternative) response as, e.g., flips of the coin with $P(success) = p$. 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. 3: To each recurrent event ξ we assign two sequences of numbers

$$u_n = P\Big(\xi \text{ occurred in the } n^{th} \textit{ trial}\Big) \qquad 1 \le n < \infty$$

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

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

$$u_n = f_0 u_n + f_1 u_{n-1} + \ldots + f_n u_0 \quad \forall\, n \ge 1$$

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

Examples of recurrent events

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 alternative response, as, e.g., flips of the coin, with probability of success p.
We say that recurrent event ξ occurred in time n if some prescribed pattern, e.g. *ZZZ* or *ZZNZN*, occurred after n trials.

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. 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.

Repeated occurrences of recurrent events

Rem. 3:

- 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

$$\left\{f_n^{(r)}\right\} = \left\{f_n\right\}^{r\star},$$

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

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 $T_i,\ 1 \le i \le r$, 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

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

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. 4: Let $T_1, T_2, \ldots$ 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 $T_1 + T_2 + \ldots + T_r = n$.
- We match the statement recurrent event occurred in time n for the r-th time with the statement $T_1 + T_2 + \ldots + T_r = n$.

Classification of recurrent events

Def. 5: Event ξ is called recurrent 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 recurrent 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. 6: If $f = 1$, then we denote $\mu = E\,T_1 = \sum_{n=0}^{\infty} nf_n$ and interpret is as the mean renewal (return) time of ξ.

Def. 7:

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

Def. 8: 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 $P(yes) = p$. 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
- $U(x) = \sum_{n=0}^{\infty} \binom{2n}{n} \left(pqx^2\right)^n = \frac{1}{\sqrt{1-4pqx^2}}$
- $F(x) = 1 - \sqrt{1 - 4pqx^2}$
- $f_{2n-1} = 0,\ f_{2n} = \frac{2}{n}\binom{2n-2}{n-1} p^n q^n,\ n = 1, 2, \ldots$
- if $p = 1/2$ then $u_n \approx 1/\sqrt{\pi n}$
- 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:

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

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

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

Proof: Follows from Thm. 47 and Rmk 49.

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 $var\, T_1 = \sigma^2 < +\infty$. Then it holds:

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

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

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

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 $n \to \infty$, where $\mu = E\,T_1$ is the mean renewal (return) time of ξ.

Rem. 4: 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. 9: Assume independent integer valued random variables $T_1, T_2, \ldots$, where T_1 has distribution $\{b_n\}$, while $T_2, T_3, \ldots$ have distribution $\{f_n\}$. We say that recurrent event with delay ξ occurred in time n for the r-th time, if

$$T_1 + T_2 + \ldots + T_r = n \tag{3}$$

Analogously, recurrent event with delay ξ occurred in time n if there exists r such that (3) holds.

Rem. 5: In Def. 9 is random variable T_1 interpreted as time to the first arrival of ξ, while $T_2, T_3, \ldots$ 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 $T_2, T_3, \ldots$ 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

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

Rem. 6: Recall equivalence between following events

$$\Big[\,\xi \text{ occurred at time } n\,\Big] \equiv$$

$$\bigcup_{k=1}^{n-1} \Big[\,\xi \text{ occurred at time } n \text{ and the last occurrence}$$

$$\text{before that event at time } k\,\Big]$$

$$\cup \Big[\,\xi \text{ occurred at time } n \text{ for the first time}\,\Big]$$

Renewal equation

Rem. 7: 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. 10: Let $a_0, a_1, a_2, \ldots$ and $b_0, b_1, b_2, \ldots$ be two sequences of real numbers such that
$a_0 = 0, 0 \le a_n \le 1, b_n \ge 0,\ n = 0, 1, 2, \ldots,\ \sum_{i=n}^{\infty} b_n < \infty$. Put
$u_n = b_n + a_0 u_n + a_1 u_{n-1} \ldots + a_n u_0,\ n = 0, 1, 2, \ldots$, i.e.

$$\{u_n\} = \{b_n\} + \{a_n\} \star \{u_n\} \tag{4}$$

Relation (4) is in the literature usually called *renewal equation*.

Rem. 8: For generating functions of sequences introduced in Def. 10 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. 11: 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

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

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

$$u_n \approx \left. \frac{B(x)}{x^{n+1} A^{\top}(x)} \right|_{x=1},$$

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 $\lim_{n\to\infty} u_n = 0$.
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 $0 \leq j < \lambda$:

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

Markov chains

Def. 12: 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

$$P\Big(E_{j_0}, E_{j_1}, \ldots, E_{j_n}\Big) = a_{j_0} p_{j_0 j_1} \ldots p_{j_{n-1} j_n}, \qquad (5)$$

where $a_k,\ k = 1, 2, \ldots$ 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. 9: 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 top know $\boldsymbol{a} \equiv \{a_k\}$ and $\mathbf{P} \equiv \Big\{p_{jk}\Big\}$. Notice that $\sum_j p_{ij} = 1 \ \ \forall i \in \mathbb{N}$.

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. 10: 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. 11: 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 $\lim_{n\to\infty} p_{jk}^{(n)}$ independent on j, then there exists also $\lim_{n\to\infty} a_k^{(n)}$ 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

$$p_{jj}^{(n)} = f_{jj}^{(0)} p_{jj}^{(n)} + f_{jj}^{(1)} p_{jj}^{(n-1)} + \ldots + f_{jj}^{(n-1)} p_{jj} + f_{jj}^{(n)} p_{jj}^{(0)}, \qquad 1 \le n < \infty$$

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

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_i, 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_i, 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_i, 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 $\lim_{n\to\infty} p_{jj}^{(n)} = 0$. In such a case $\lim_{n\to\infty} p_{ij}^{(n)} = 0 \; \forall i$.

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

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

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

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

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

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

Further, it holds:

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

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. 12: 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. 13: If we omit in a matrix of transitional probabilities $\boldsymbol{P}$ 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 is 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 $\boldsymbol{P}$ 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. 14: 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 $\boldsymbol{P}$. A distribution $\{v_j\}$ is called stationary distribution of this chain if $\forall j$

$$v_j = \sum_i v_i p_{ij} \tag{6}$$

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

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

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

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

Stationary distribution (cont.)

Rem. 15: In an irreducible MC with finitely many states stationary distribution exits, 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 $\boldsymbol{P}$. If number of states n if 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 exists.

Ex. 10:

- 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 $(\boldsymbol{P}, \boldsymbol{a})$. 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. 16: Assume irreducible reversible MC $(\boldsymbol{P}, \boldsymbol{a})$. Then it holds

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

Thm. 26: If MC $(\boldsymbol{P}, \boldsymbol{A})$ is reversible, then corresponding vector $\boldsymbol{\pi}$ is its stationary distribution.

Rem. 17: Thm. 26 does not hold in opposite direction, i.e., existence of the stationary distribution does not imply that corresponding MC is reversible.

MORE GENERAL DEFINITION OF MARKOV CHAIN

Assume MC with states $S_1, S_2, \ldots$, initial distribution $\{a_j\}$ and matrix of transition probabilities $\mathbf{P} \equiv \left\{p_{ij}\right\}$. 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 $\forall i_0, i_1, \ldots, i_{n-1}, i, j$ and $\forall n \in \mathbb{N}_0$

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

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

Rem. 18: 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. 21: Sequence of ivrv's $X_n,\ n \in \mathbb{N}_0$, is called Markov chain, if $\forall i_0, i_1, \ldots, i_{n-1}, i, j$ and $\forall n \in \mathbb{N}_0$ it holds

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

More general definition of Markov chain (cont.)

Rem. 19: Markov chains considered up to now had the property that p_{ij} do not depend on n. Such MC are called homogeneous.
If $P\big(X_{n+1} = j \,|\, X_n = i\big)$ depends on n, we will denote it $p_{ij}(n, n+1)$ and to call them non homogeneous Markov chains.
Rem. 20: Probabilities

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

change

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

Ex. 11: Let $Y_k,\ 1 \le k < \infty$, are independent ivrv's and $X_n = \sum_{k=1}^{n} Y_k$. Then sequence $\big\{X_n\big\},\ 1 \le n < \infty$ forms Markov chain. Prove it.
Ex. 12: Let $Y_k,\ 1 \le k < \infty$, are independent ivrv's and assume a sequence of moving sums $X_n^\star = \sum_{k=1}^{r} Y_{n+k},\ r$ being fixed. Then a sequence $\big\{X_n^\star,\ 1 \le n < \infty\big\}$ generally does not form Markov chain.

Reducible Markov chains

Rem. 21: 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 irreducible MC with finitely many states. Then probability that chain will be absorbed in any closed subset of states converges to 1 for $n \to \infty$, and it does not matter in which state we started.

Rem. 22: Assume irreducible MC with finitely many states described by matrix of transitional probabilities $\boldsymbol{P}$, and construct sequentially matrices $\boldsymbol{P}^2,\ \boldsymbol{P}^3, \ldots$

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

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

Reducible Markov chains (cont.)

Rem. 23: For each reducible MC described by matrix of transitional probabilities $\boldsymbol{P} = \begin{pmatrix} \boldsymbol{S} & \boldsymbol{0} \\ \boldsymbol{R} & \boldsymbol{Q} \end{pmatrix}$ corresponding matrix $\boldsymbol{I} - \boldsymbol{Q}$ has inversion and it holds

$$\boldsymbol{I} + \boldsymbol{Q} + \boldsymbol{Q}^2 + \ldots = \sum_{k=0}^{\infty} \boldsymbol{Q}^k = \left(\boldsymbol{I} - \boldsymbol{Q}\right)^{-1}$$

Matrix $\left(\boldsymbol{I} - \boldsymbol{Q}\right)^{-1}$ is called fundamental matrix.

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:

- $x_j = P\big(S_j \to C\big)$ 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. 28: Probabilities $x_j,\ j \in T$, satisfy a set of equations

$$\xi_j - \sum_{\nu \in T} p_{j\nu}\xi_\nu = x_j^{(1)} \tag{7}$$

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

$$\eta_j = \sum_{\nu \in T} p_{j\nu}\eta_\nu \tag{8}$$

Thm. 30: Set of equations (7) has unique bounded solution $\Leftrightarrow$ set of equations (8) does not have other bounded solution than the trivial one.

Thm. 31: Probabilities y_j are equal to zero $\forall\, j \in T \Leftrightarrow$ a set of equations (8) does not have other solution than the trivial one.

Thm. 32: In a chain with finitely many states all $y_j = 0$ and x_j form unique solution of set of equations (7).

Transient states (cont.)

Thm. 33: In a chain with states $S_0, S_1, S_2, \ldots$ are all states transient $\Leftrightarrow$ set of equations

$$\eta_j = \sum_{\nu=1}^{\infty} p_{j\nu}\eta_\nu, \quad 1 \leq j < \infty, \tag{9}$$

has nontrivial bounded solution.

Ising model – Notation

- G ... graph
- V ... vertexes of given graph
- $|V| = card(V)$
- 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 $\{1, \ldots, K\}$ which can describe levels of gray or colors, $|V| = card(V)$, etc.
- $\boldsymbol{\sigma} = \left(\sigma_1, \ldots, \sigma_{|V|}\right)$ describes states of the system
- State space S is $\{-1, +1\}^{|V|}$, respectively $\{1, \ldots, K\}^{|V|}$, etc.

Ising model

Def. 22: **Ising model** is probability distribution $\pi(\beta)$ on state space $S = \{-1, +1\}^{|V|}$, where

$$\pi(\beta) = \frac{e^{-\beta H(\sigma)}}{C_\beta}, \tag{10}$$

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

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. 23: 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**:

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

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & \mathbf{1} & 0 \\ 1 & 0 & 1 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 0 & 1 \\ 3 & 3 & 3 \\ 2 & 3 & 2 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 & 1 \\ 0 & \mathbf{0} & 0 \\ 1 & 0 & 1 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 2 \\ 2 & 2 & 2 \end{pmatrix}$$

S $\quad H(S)$ $\quad S$ $\quad H(S)$

- **Ising model with outside field**:

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

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:

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

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

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

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

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

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

Problem: To reconstruct true image $\boldsymbol{\sigma}$ provided we observe $\boldsymbol{Y}$ and assume that $\boldsymbol{\sigma}$ 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 $(\boldsymbol{\sigma}, \boldsymbol{Y})$ is

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

Posterior distribution is

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

What can we do:

- To generate from a posterior distribution $(\boldsymbol{\sigma} \,|\, \boldsymbol{Y})$. 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 $\widehat{\boldsymbol{\sigma}}$ that maximizes $P(\boldsymbol{\sigma} \,|\, \boldsymbol{Y})$.

Exponential distribution – repetition

Def. 24: We say that random variable (rv) X follows exponential distribution ($X \sim Exp(\lambda)$), if corresponding density has the form

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

Thm. 34: Let $X \sim Exp(\lambda)$. Then $F(x) = 1 - e^{-\lambda x},\ E\,X = \lambda^{-1}$ and $var\,X = \lambda^{-2}$.

Thm. 35: No memory property. Let $X \sim Exp(\lambda)$. 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.

$$P(X > x + y \,|\, X > x) = P(X > y), \qquad \forall\, x, y > 0 \tag{11}$$

Rem. 26: Exponential distribution is the only continuous distribution for which (11) holds. Among discrete distributions the only one with the same property is geometric distribution.

Intensity function

Def. 25: 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. 27: 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:

$$P\big(x < X \leq x + \Delta \,|\, X > x\big) = \frac{P(x < X \leq x + \Delta)}{P(X > x)} = \frac{F(x + \Delta) - F(x)}{1 - F(x)}$$
$$= \frac{F(x + \Delta) - F(x)}{1 - F(x)} \frac{\Delta}{\Delta} \overset{\Delta \to 0}{\approx} \Delta \frac{f(x)}{1 - F(x)}$$

Rem. 28: Let rv $X \sim Exp(\lambda)$. 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.

- $S_n \to S_{n+1} \ldots$ **birth**
- $S_n \to S_{n-1} \ldots$ **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

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

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 $p_n = \lim_{t\to\infty} P_n(t)$.

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

$$
\begin{aligned}
P_0'(t) &= -\lambda_0 P_0(t) + \mu_1 P_1(t) && (12)\\
P_n'(t) &= -\big(\lambda_n + \mu_n\big)P_n(t) + \mu_{n+1}P_{n+1}(t) + \lambda_{n-1}P_{n-1}(t), \;\; n \geq 1
\end{aligned}
$$

- 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

$$
\begin{aligned}
0 &= -\lambda_0 p_0 + \mu_1 p_1 && (13)\\
0 &= -\big(\lambda_n + \mu_n\big)p_n + \mu_{n+1}p_{n+1} + \lambda_{n-1}p_{n-1} \;\; n \geq 1
\end{aligned}
$$

which we receive if we set in (12) $P_n'(t) = 0, \;\; n \geq 0$, and replace $P_n(\infty)$ by p_n.

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. 29: 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. 36: System of equations (12) for considered process has solution

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

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

$$A(t) = \frac{\mu\left(e^{(\lambda-\mu)t} - 1\right)}{\lambda e^{(\lambda-\mu)t} - \mu} \qquad \textit{and} \qquad B(t) = \frac{\lambda\left(e^{(\lambda-\mu)t} - 1\right)}{\lambda e^{(\lambda-\mu)t} - \mu}$$

Linear process of growth and death (cont.)

Model

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

$\boldsymbol{Q}$ =

	S_0	S_1	S_2	S_3	$\dots$
S_0	$\cdot$	λ_0	$\cdot$	$\cdot$	$\cdot$
S_1	μ_1	$\cdot$	λ_1	$\cdot$	$\cdot$
S_2	$\cdot$	μ_2	$\cdot$	λ_2	$\cdot$
S_3	$\cdot$	$\cdot$	μ_3	$\cdot$	λ_3

$\boldsymbol{Q}^\star$ =

	S_0	S_1	S_2	S_3	S_4	$\dots$	$\sum$
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.)

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

	S_0	S_1	S_2	S_3	S_4	
S_0	$1-\lambda_0$	μ_1	0	0	0	$\dots$
S_1	λ_0	$1-(\lambda_1+\mu_1)$	μ_2	0	0	$\dots$
S_2	0	λ_1	$1-(\lambda_2+\mu_2)$	μ_3	0	$\dots$
S_3	0	0	λ_2	$1-(\lambda_3+\mu_3)$	μ_3	$\dots$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
$\sum$	1	1	1	1	1	$\dots$

$\underbrace{\hphantom{xxxxxxxx}}_{\boldsymbol{I}+\boldsymbol{Q}^{\star\prime}}$

System of Kolmogorov's differential equations has the form

$$\boldsymbol{P}'(t) = \big(\boldsymbol{I} + \boldsymbol{Q}^{\star\prime}\big)\boldsymbol{P}(t) \qquad \& \qquad \boldsymbol{0} = \big(\boldsymbol{I} + \boldsymbol{Q}^{\star\prime}\big)\boldsymbol{p}$$

$$\boldsymbol{P}'(t) = \big(P_0'(t), P_1'(t), \dots\big)', \ \boldsymbol{P}(t) = \big(P_0(t), P_1(t), \dots\big)', \ \boldsymbol{p} = \big(p_0, p_1, \dots\big)'$$

Little's formula

- L ... mean number of clients in system
- L_Q ... mean number of clients waiting in queue
- W ... mean time client spends in system
- W_Q ... mean time client spends in queue
- λ_a ... average arrival rate of entering customers
- ARE ... average rate at which system earns
- AAP ... average amount an entering customer pays
- $N(t)$ ... number of clients that entered to the system until time t

Little's formula

- $\lambda_a = \lim_{t\to\infty} \frac{N(t)}{t}$ & $ARE = \lambda_a \cdot AAP$
- $L = \lambda_a W$ & $L_Q = \lambda_a W_Q$

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 $k = 1, \ldots, M$ 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 $Bi\big(N, \mu/(\mu+\lambda)\big)$, i.e.

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

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 $B \to A$ arise with probability zero.

Tasks:

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

Queuing system

Def. 26: 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. 30: 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. 27: 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. 37: System **M/M/x** can be described by the general process of birth and death.

Rem. 31: For model:

- M/M/1 : $\lambda_j = \lambda,\ 0 \le j < \infty$ and $\mu_j = \mu,\ 1 \le j < \infty$
- M/M/c : $\lambda_j = \lambda,\ 0 \le j < \infty,\ \mu_j = j\mu,\ 0 \le j \le c$ and
  $\mu_j = c\mu,\ c \le j < \infty$
- M/M/∞ : $\lambda_j = \lambda,\ 0 \le j < \infty,\ \mu_0 = 0$ and $\mu_j = j\mu,\ 0 \le j < \infty$

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

M/M/1, M/M/c AND M/M/∞

Thm. 38: For systems **M/M/x** it holds:

- M/M/1 : limit probabilities p_n follow geometric distribution with parameter $1 - \lambda/\mu$.
- M/M/c : limit probabilitiesp_n follow truncated Poisson distribution with parameters $\big(c + 1,\ \lambda/\mu\big)$.
- M/M/∞ : limit probabilities p_n follow Poisson distribution with parameter λ/μ.

Thm. 39: 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 λ!

Rem. 32: 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. 33: Assume model **M/M/1**. We know from Theorem 38 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 $\frac{\lambda}{\mu}\Big/\left(1 - \frac{\lambda}{\mu}\right)$
- Variance of rv describing number of clients in system is $\frac{\lambda}{\mu}\Big/\left(1 - \frac{\lambda}{\mu}\right)^2$
- Mean length of the queue is $\sum_j jp_{j+1} = \left(\frac{\lambda}{\mu}\right)^2\Big/\left(1 - \frac{\lambda}{\mu}\right)$

Rem. 34: 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. 35: 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:

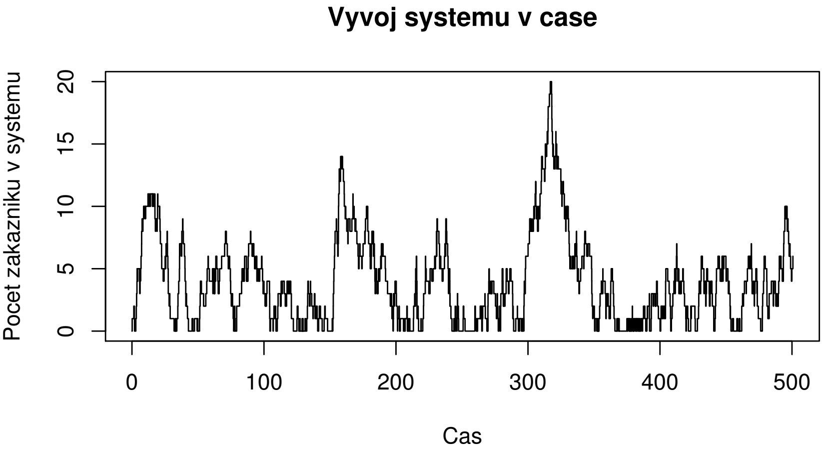
$$P(T \leq t) = 1 - e^{-(\mu-\lambda)t}, \quad t \geq 0$$

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

TANDEM = M/M/1 + M/M/1

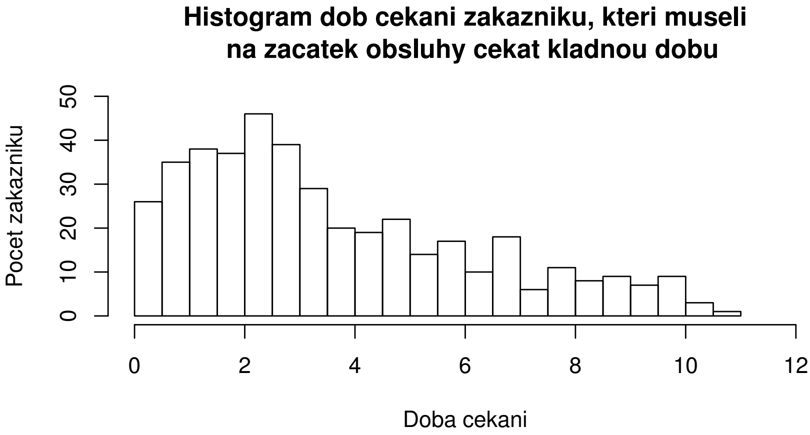
Serial system composed of two independent M/M/1 systems is called **tandem**. Details will be added later.

M/M/1



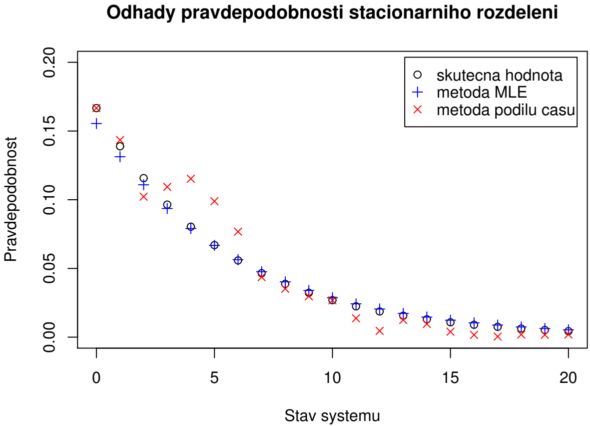
Evolution of number of clients in system M/M/1 with parameters $\lambda = 1$ a $\mu = 1.2$.

M/M/1



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



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

Counting process

Def. 28: 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. 36: 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. 29: 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. 30: 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.

Poisson process

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

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

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

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

Poisson process – properties

Thm. 40: Definitions 31 and 32 of Poisson process are equivalent.

Rem. 37: Conditions (ii) – (iv) can be replaced by the following set of equivalent conditions:

- **iii)** Process has independent increments.
- **iv)** $P\Big(N(t+h) - N(t) = 1\Big) = \lambda h + o(h)$
- **v)** $P\Big(N(t+h) - N(t) \geq 2\Big) = o(h)$

Rem. 38: Poisson process has stationary increments and $EN(t) = \lambda t$.

Rem. 39: The fact that $N(t)$ follow Poisson distribution is a consequence of the approximation of binomial distribution by the Poisson one.

Poisson process – times between events

Def. 33: Assume Poisson process with intensity λ. Denote by $T_n,\ n = 1, 2, \ldots$ 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. 41: Sequence of times between the events follow exponential distribution with intensity λ.

Thm. 42: 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 (10), from where

$$N(t) \geq n \quad \Longleftrightarrow \quad S_n \leq t$$

and differentiate corresponding distribution function.

Poisson process 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. 40: 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.

Branching process alias How viruses can propagate

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

$$P(U = j) = p_j, \quad j = 0, 1, 2, \dots \tag{14}$$

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 descendents behave independently each from other.

Problems.

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

Branching process (cont.)

Model. Let $X_0 = 1$, $X_1 (\equiv U)$ follow (14) 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 (14). Let these rv's are mutually independent and independent on X_1, i.e.

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

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. descendents of X_2 rv's with distribution as X_1, i.e. (14) etc.

Thm. 43: Probability generation function $P_n(x)$ of X_n fulfil recurrent relation

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

and, moreover, it holds

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

Markov processes with discrete states and continuous time

Def. 34: 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. 41: 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

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

Basic characterizations of random variables

Rem. 42: 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. 35: Assume probability space $\left(\Omega, \mathcal{A}, \mathcal{P}\right)$. 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. 36: Let X and Y are two discrete rv's. Then:

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

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

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

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

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

$$E\left[\, X|Y = y \,\right] = \sum_{x} x \cdot \mathcal{P}\big(X = x|Y = y\big) = \sum_{x} x \cdot p_{X|Y}(x|y)$$

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

$$var\left[\, X|Y = y \,\right] = E\Big[\Big(X - E\left[\, X|Y = y \,\right]\Big)^2 |Y = y\Big]$$

Conditional characteristics for continuous rv's

Def. 37: 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

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

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

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

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

$$var\big[\,X|Y = y\,\big] = E\Big[\big(X - E\big[\,X|Y = y\,\big]\big)^2 |Y = y\Big]$$

Generating functions

Def. 38: 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. 43: If $\{a_j\}$ is bounded, then $\mathcal{A}(x)$ evidently converge at least in the interval $(-1, 1)$.

Def. 39: 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. 44:

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

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 ... $\mathcal{P}(x) = \big(p/(1-qx)\big)^r$ resp. $= \big(px/(1-qx)\big)^r$
- Discrete uniform ... $\mathcal{P}(x) = (1-x^{n+1})/\big((n+1)(1-x)\big)$ resp. $= \big(x(1-x^n)\big)/\big(n(1-x)\big)$

Using these generating functions calculate corresponding expectations and variances.

Rem. 45: Recall that geometric and negative binomial distributions are the simplest models describing waiting times.

PROPERTIES OF GENERATING FUNCTIONS

Thm. 44: Denote $q_k = P(X > k) = \sum_{j>k} p_j,\ k = 0, 1, 2, \ldots$ and corresponding generating function $\mathcal{Q}(x) = \sum_{j=0}^{\infty} q_j x^j$. Then for $-1 < x < 1$ it holds that $\mathcal{Q}(x) = \big(1 - \mathcal{P}(x)\big)/(1 - x)$.

Thm. 45: For integer valued random variable X it holds that

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

Thm. 46: Let generating function $\mathcal{P}(x)$ of integer valued rv X has radius of convergence larger than one. Then it holds

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

Rem. 46: Thm. 46 holds also in the case when radius of convergence is equal to 1, provided $\lim_{x\to 1_-} \mathcal{Q}''(x) < \infty$ and derivatives in point $x = 1$ are replaced by their limits for $x \to 1_-$.

PARTIAL FRACTION DECOMPOSITION

Rem. 47: 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. 47: 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

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

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

Rem. 48: For calculation of $\rho_k,\ 1 \le k \le m$, can be used decomposition into partial fractions, embedded in programs *Maple* or *Mathematica*, e.g.

Partial fraction decomposition

Rem. 49: Assume that x_1 is that root of $V(x)$ for which $|x_1| < x_k,\ 2 \le k \le m$. Then

$$p_n = \frac{\rho_1}{x_1^{n+1}}\Big(1 + \frac{\rho_2}{\rho_1}\Big(\frac{x_1}{x_2}\Big)^{n+1} + \ldots + \frac{\rho_m}{\rho_1}\Big(\frac{x_1}{x_m}\Big)^{n+1}\Big), \qquad (15)$$

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

Rem. 50: 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 (15) 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: $Q(x) = \big(8 + 4x + 2x^2\big)/\big(8 - 4x - 2x^2 - x^3\big)$.

Convolution

Def. 40: Let $a_0, a_1, \ldots$ and $b_0, b_1, \ldots$ are two sequences of real numbers. Then a sequence $c_0, c_1, \ldots$ defined by relation

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

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

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

Thm. 48: 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

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

Rem. 51: Convolution of a sequence $\{a_j\}$ with itself is called convolution power and is denoted $\{a_j\}^{2\star}$. Analogously, n-th convolution power $\{a_j\} \star \ldots \star \{a_j\}$will be denoted $\{a_j\}^{n\star}$.

Convolution

Thm. 49: Let $X_1, X_2, \ldots, X_n$ 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 $X_1 + X_2 + \ldots + X_n$ is described by the n-th convolution power $\{p_j\}^{n\star}$, and corresponding generating function has the form

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

COMPOUND DISTRIBUTIONS

Thm. 50: Let $X_1, X_2, \ldots$ and N are independent integer valued rv's, X_i's have the same distribution $\{f_j\}$ and N has distribution $\{g_j\}$. Then $S_N = X_1 + \ldots + X_N$ is also integer valued random variable with the distribution $\{h_j\}$, where

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

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 $\mathcal{C}(x) = \mathcal{B}\big(\mathcal{A}(x)\big)$ and $E\, S_N = E\, X_1 \cdot E\, N$. Corresponding variance o can be calculated using Thm. 46.

Rem. 52: Notice that random variable $S_N = X_1 + \ldots + X_N$ 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 $Po(\lambda \cdot p)$.