AndrásVetier

Q U E U I N G

January 25, 2000

Supported by

Pro Renovanda Cultura Hunariae Alapítvány

Klebelsberg Kunó Emlékére Szakalapitvány

2000

Contents

Introduction

The theory of stochastic processes is, in a sense, a part of, in other sense, the continuation of probability theory. In both theories models are constructed about observations of random phenomena. In a first probability course, the observed phenomena are simpler: the sample space (the set of all possible outcomes) is either a finite or countably infinite set (discrete models) or a subset of a finite dimensional Euclidian space (continuous models). In a stochastic processes course the observed models may be more complicated: the sample space may consist of sequences or functions or even more complicated mathematical objects.

For example, observing the height and weight of a person randomly chosen from a given population, the outcomes are pairs of real numbers, and, as a model, we can use a two-dimensional normal distribution. As we learnt in probability theory, we may, for example, estimate the weight from the height by the so called regression line.

However, observing the movement of a particle in a big tank filled with gas for a given time interval, the possible outcomes of this so called Brown motion are vector valued functions. In fact, each outcome of this observation is a function describing how the position of the particle changes as time passes. Putting an emphasis on the fact that the domain of the possible outcomes is a continuous time interval, such processes are called **continuous-time processes**.

On the contrary, if we observe the position of the chosen particle only at time instants $1, 2, \ldots$, then any result of our observation is an infinite sequence of points. Since any sequence is a function defined on the set of natural numbers, such processes are called **discrete-time processes**.

In this textbook, we deal with such important types of discrete-time processes as Markov chains, stationary sequences, sequences of independent observations. Such continuous-time processes will be discussed as jump processes, birth and death processes, and the famous Wiener process related to the Brown motion. The Poisson point processes will differ from all the other processes, because its realizations are not functions, but point systems.

The continuous-time analogies of Markov chains, called Markov processes, and those of stationary sequences, called stationary processes, in some sense, would be quite natural to include in the course, but technically the discrete-time processes are easier to handle. This is why we shall rather deal with the discrete-time Markov chains and stationary sequences only.

Chapter 1

Markov Chains

1. Transition Matrix

Assume that a process has a finite or countable infinite number of possible states, which will be identified with the positive integers in this textbook. The process is observed at discrete time instants, which will be denoted by the non-negative integers or, in some cases, by the positive integers. Let the state in which the process at time instant t is be denoted by ξ_t $(t = 0, 1, 2, \ldots)$. The process $\xi_0 , \xi_1 , \xi_2 , \ldots$ is called a **Markov chain** if, for any time instant t (called here the **present time instant**) and any states $i_0 , i_1 , \ldots , i_{t-1}$ (called here the **states of the process in the past** or briefly, **past states**), and any two states i , j (called here the **present** and **future states**, respectively), the conditional probability

$$\Pr \left(\xi_{t+1} = j \mid \xi_0 = i_0 , \xi_1 = i_1 , \ldots , \xi_{t-1} = i_{t-1} , \xi_t = i \right)$$

does not depend on the states $i_0 , i_1 , \ldots , i_{t-1}$ but it is equal to

$$\Pr \left(\xi_{t+1} = j \mid \xi_t = i \right) .$$

This property, called **Markov property**, can be verbalized like this: on condition that the present state of the process is given, the conditional probability of any future state does not depend on the past states of the process.

If, for any time instant t and any states i , j , the conditional probability

$$\Pr \left(\xi_{t+1} = j \mid \xi_t = i \right)$$

does not depend on t , that is, it is equal to

$$\Pr \left(\xi_1 = j \mid \xi_0 = i \right) ,$$

then the chain is called **homogeneous**. Homogeneity means that the value of the conditional probability of a state in the future does not depend on the value of the present time instant. In these lecture notes we shall develop only the theory of homogeneous Markov chains.

1. Example: SUMS OF RANDOM NUMBERS

Let the numbers $\rho_0 , \rho_1 , \rho_2 , \ldots$ be random numbers generated by independent tosses of a fair die, and let the random variables $\alpha_t , \beta_t , \gamma_t$ be defined by

a) $\alpha_t = \rho_t + \rho_{t+1}$ $(t = 0, 1, 2, \ldots)$,
b) $\beta_t = \rho_0 + \rho_1 + \rho_2 + \ldots + \rho_t$ $(t = 0, 1, 2, \ldots)$,
c) $\gamma_t = \rho_0 + 2 \cdot \rho_1 + 3 \cdot \rho_2 \ldots + (t + 1) \cdot \rho_t$ $(t = 0, 1, 2, \ldots)$.

Which of the three sequences constitutes a (homogeneous) Markov chain? Before doing or reading the calculations, try, on a heuristic basis, conjecture whether the future of these processes depends or does not depend on their past, on condition that their present is given.

Solution.

a) The process α_t $(t = 0, 1, 2, \dots)$ is not a Markov process. If the time instants 0, 1, 2 represent past, present and future time instants, respectively, then it is easy to check that the conditional probabilities

$$\Pr\left(\alpha_2 = 12 \mid \alpha_0 = 12 , \alpha_1 = 11 \right),$$

$$\Pr\left(\alpha_2 = 12 \mid \alpha_0 = 11 , \alpha_1 = 11 \right)$$

are not equal, which breaks the Markov property:

$$\Pr\left(\alpha_2 = 12 \mid \alpha_0 = 12 , \alpha_1 = 11 \right) = \frac{\Pr\left(\alpha_0 = 12 , \alpha_1 = 11 , \alpha_2 = 12 \right)}{\Pr\left(\alpha_0 = 12 , \alpha_1 = 11 \right)}$$

$$= \frac{\Pr\left(\rho_0 = 6 , \rho_1 = 6 , \rho_2 = 5 , \rho_3 = 7 \right)}{\Pr\left(\rho_0 = 6 , \rho_1 = 6 , \rho_2 = 5 \right)} = 0 ,$$

and

$$\Pr\left(\alpha_2 = 12 \mid \alpha_0 = 11 , \alpha_1 = 11 \right) = \frac{\Pr\left(\alpha_0 = 11 , \alpha_1 = 11 , \alpha_2 = 12 \right)}{\Pr\left(\alpha_0 = 11 , \alpha_1 = 11 \right)}$$

$$= \frac{\Pr\left(\rho_0 = 6 , \rho_1 = 5 , \rho_2 = 6 , \rho_3 = 6 \right) + \Pr\left(\rho_0 = 5 , \rho_1 = 6 , \rho_2 = 5 , \rho_3 = 7 \right)}{\Pr\left(\rho_0 = 6 , \rho_1 = 5 , \rho_2 = 6 \right) + \Pr\left(\rho_0 = 5 , \rho_1 = 6 , \rho_2 = 5 \right)}$$

$$= \frac{\left(\frac{1}{6}\right)^4 + 0}{\left(\frac{1}{6}\right)^3 + \left(\frac{1}{6}\right)^3} = \frac{1}{12} .$$

b) The process β_t $(t = 0, 1, 2, \dots)$ is a homogeneous Markov process. We shall first calculate the conditional probability for the Markov property when $t = 2$ is considered as the present time instant. If 2 is the present time instant, then time instants 0 and 1 represent the past, and time instant 3 represents the future time instant:

$$\Pr\left(\beta_3 = j \mid \beta_0 = i_0 , \beta_1 = i_1 , \beta_2 = i \right)$$

$$= \frac{\Pr\left(\beta_0 = i_0 , \beta_1 = i_1 , \beta_2 = i , \beta_3 = j \right)}{\Pr\left(\beta_0 = i_0 , \beta_1 = i_1 , \beta_2 = i \right)}$$

$$= \frac{\Pr\left(\rho_0 = i_0 , \rho_1 = i_1 - i_0 , \rho_2 = i - i_1 , \rho_3 = j - i \right)}{\Pr\left(\rho_0 = i_0 , \rho_1 = i_1 - i_0 , \rho_2 = i - i_1 \right)}$$

$$= \frac{\Pr\left(\rho_0 = i_0\right)\ \Pr\left(\rho_1 = i_1 - i_0\right)\ \Pr\left(\rho_2 = i - i_1\right)\ \Pr\left(\rho_3 = j - i\right)}{\Pr\left(\rho_0 = i_0\right)\ \Pr\left(\rho_1 = i_1 - i_0\right)\ \Pr\left(\rho_2 = i - i_1\right)}$$

$$= \Pr\left(\rho_3 = j - i\right).$$

One can check that, for the general present time instant t, the conditional probability is

$$\Pr\left(\beta_{t+1} = j \mid \beta_0 = i_0, \beta_1 = i_1, \dots, \beta_{t-1} = i_{t-1}, \beta_t = i\right)$$

$$= \Pr\left(\rho_{t+1} = j - i\right) = \Pr\left(\rho_0 = j - i\right).$$

Since the last formula does not contain $i_0, i_1, \dots, i_{t-1}$, the conditional probability in question does not depend on the past states $i_0, i_1, \dots, i_{t-1}$, which means that the process is a Markov process. Since the value of the conditional probability does not depend on the present time instant t either, the Markov chain is homogeneous.

c) Let us analyze the process γ_t $(t = 0, 1, 2, \dots)$. We shall show – in a similar way – that it is a Markov process, but non homogeneous.

$$\Pr\left(\gamma_3 = j \mid \gamma_0 = i_0, \gamma_1 = i_1, \gamma_2 = i\right)$$

$$= \frac{\Pr\left(\gamma_0 = i_0, \gamma_1 = i_1, \gamma_2 = i, \gamma_3 = j\right)}{\Pr\left(\gamma_0 = i_0, \gamma_1 = i_1, \gamma_2 = i\right)}$$

$$= \frac{\Pr\left(\rho_0 = i_0, 2\rho_1 = i_1 - i_0, 3\rho_2 = i - i_1, 4\rho_3 = j - i\right)}{\Pr\left(\rho_0 = i_0, 2\rho_1 = i_1 - i_0, 3\rho_2 = i - i_1\right)}$$

$$= \frac{\Pr\left(\rho_0 = i_0\right)\ \Pr\left(2\rho_1 = i_1 - i_0\right)\ \Pr\left(3\rho_2 = i - i_1\right)\ \Pr\left(4\rho_3 = j - i\right)}{\Pr\left(\rho_0 = i_0\right)\ \Pr\left(2\rho_1 = i_1 - i_0\right)\ \Pr\left(3\rho_2 = i - i_1\right)}$$

$$= \Pr\left(4\rho_3 = j - i\right).$$

One can check that, for the general present time instant t, the conditional probability is

$$\Pr\left(\gamma_{t+1} = j \mid \gamma_0 = i_0, \gamma_1 = i_1, \dots, \gamma_{t-1} = i_{t-1}, \gamma_t = i\right)$$

$$= \Pr\left((t+2)\cdot\rho_{t+1} = j - i\right) = \Pr\left((t+2)\cdot\rho_0 = j - i\right).$$

Since the last formula does not contain $i_0, i_1, \dots, i_{t-1}$, the conditional probability in question does not depend on the past states $i_0, i_1, \dots, i_{t-1}$, which means that the process is a Markov process. Since the value of the conditional probability does depend on the present time instant t, the Markov chain is non-homogeneous. □

The conditional probabilities

$$p_{ij}(n) = \Pr\left(\xi_n = j \mid \xi_0 = i \right),$$

called **n -step transition probabilities**, have the following meaning: on condition that the process is in the state i at a certain time instant, the probability that, n units of time later, it will be in state j is $p_{ij}(n)$. The matrix

$$P(n) = \left(p_{ij}(n) \right)$$

is called the **n -step transition matrix**. An n -step transition matrix is always a square matrix, its size is the number of the states, of the process, and the row-sums of the matrix are equal, obviously, to 1 .

2. Example: LENGTH OF RUN

Use a false coin to generate a random head-tail sequence. Assume that for each toss, the probability of a head or a tail is p and q , respectively. For each n let the random variable ξ_n be equal to the length of the head-run at the n -th toss, that is,

$\xi_n = 0$, if the n -th toss is tail;
$\xi_n = 1$, if the n -th toss is a head , but the $(n-1)$ -th toss is a tail;
$\xi_n = 2$, if the n -th and $(n-1)$ -th tosses are heads , but the $(n-2)$ -th toss is a tail;
and so on.

It is easy to see that the sequence $\xi_1, \xi_2, \ldots$ is a homogeneous Markov chain. Find the n -step transition probabilities and the n -step transition matrix.

Solution.

If the process is in a state i at present, then n units of time later it may be either in state $i+n$ or in any of the states $0, 1, 2, \ldots, n-1$. State $i+n$ occurs if all the n tosses yield heads, state j $(j = 0, 1, 2, \ldots, n-1)$ occurs if the last j tosses yield heads, but the one preceding these tosses is a tail. This means that

$$\begin{aligned} p_{ij}(n) &= p^n && \text{if } j = i+n, \\ p_{ij}(n) &= q \cdot p^j && \text{if } j = 0, 1, 2, \ldots, n-1, \\ p_{ij}(n) &= 0 && \text{otherwise.} \end{aligned}$$

(To draw the n -step transition matrix is now an easy exercise for the reader. Check that the row-sums are equal to 1 .) □

The 1 -step transition probabilities are denoted also by p_{ij} , and the 1 -step transition matrix is denoted also by P :

$$p_{ij} = p_{ij}(1) = \Pr\left(\xi_1 = j \mid \xi_0 = i \right) = \Pr\left(\xi_{t+1} = j \mid \xi_t = i \right),$$

$$P = \left(p_{ij} \right).$$

Using the total probability formula and induction, one can be convinced that

$$\begin{aligned} P(1) &= P^1 \\ P(2) &= P^2 \\ P(3) &= P^3 \end{aligned}$$

$$\vdots$$

that is, the n-step transition matrix is equal to the n-th power of the 1-step transition matrix:

$$P(n) = P^n .$$

4. Example: TWO-STATE MARKOV CHAINS – TRANSITION MATRICES

If a Markov chain has only two states, then its 1-step transition matrix is of the form

$$P = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix},$$

where $a = p_{12}$, $b = p_{21}$. Find an explicit formula for the n-step transition probabilities.

Solution.

Since $P(n) = P^n$, we have the recursive relation

$$P(n+1) = P(n) \cdot P$$

between the transition matrices, which holds also for $n = 0$, if $P(0)$ is , by definition, the 2×2 unit matrix . The recursive relation for the transition probabilities gives that

$$p_{11}(n+1) = p_{11}(n) \cdot p_{11} + p_{12}(n) \cdot p_{21}$$

$$= p_{11}(n) \cdot \Big(1-a\Big) + \Big(1-p_{11}(n)\Big) \cdot b ,$$

that is,

$$p_{11}(n+1) = p_{11}(n) \cdot \Big(1-a-b\Big) + b ,$$

which is a recursive formula for the transition probabilities $p_{11}(n)$ $(n = 0, 1, 2, \ldots)$.

As it is known, for a sequence $X(n)$ $(n = 0, 1, 2, \ldots)$ defined by the recursive formula

$$X(n+1) = X(n) \cdot A + B$$

there exists a closed formula:

$$X(n) = \tfrac{B}{1-A} + \Big(X(0) - \tfrac{B}{1-A}\Big) \cdot A^n \qquad \text{if } A \neq 1 \text{, and}$$

$$X(n) = X(0) + n \cdot B \qquad \text{if } A = 1 .$$

(For a proof see the Remark at the end of this section.)

Substituting $A = 1-a-b$, $B = b$, $X(0) = 1$, we get that

$$p_{11}(n) = \tfrac{b}{a+b} + \tfrac{a}{a+b} \cdot (1-a-b)^n$$

if at least one of a and b differs from 0 . Using that $p_{12}(n) = 1 - p_{11}(n)$, we get that

$$p_{12}(n) = \tfrac{a}{a+b} - \tfrac{a}{a+b} \cdot (1-a-b)^n .$$

Changing the role of the indices, and the role of a and b, the other two elements of the n-step transition matrix can be determined, too:

$$p_{21}(n) = \frac{b}{a+b} - \frac{b}{a+b} \cdot (1 - a - b)^n .$$

$$p_{22}(n) = \frac{a}{a+b} + \frac{b}{a+b} \cdot (1 - a - b)^n .$$

If $a = b = 0$, then obviously $p_{11}(n) = 1$, $p_{12}(n) = 0$, $p_{21}(n) = 0$, $p_{22}(n) = 1$, that is $P(n)$ is the unit matrix for all n.

5. Exercise.

Study the behavior of $P(n)$ as $n \to \infty$, and think of the physical meaning of the result you get. □

Remark. (Proof of a Closed Formula for the Sequence $X(n+1) = X(n) \cdot A + B$.)

It is obvious that

$$\begin{aligned}
X(1) &= X(0) \cdot A + B \\
X(2) &= X(1) \cdot A + B = \Big(X(0) \cdot A + B\Big) \cdot A + B \\
X(3) &= X(2) \cdot A + B = \Big(\Big(X(0) \cdot A + B\Big) \cdot A + B\Big) \cdot A + B \\
X(4) &= X(3) \cdot A + B = \Big(\Big(\Big(X(0) \cdot A + B\Big) \cdot A + B\Big) \cdot A + B\Big) \cdot A + B \\
&\vdots
\end{aligned}$$

Multiplying out the parentheses, we get for $X(4)$ that

$$\begin{aligned}
X(4) &= X(0) \cdot A^4 + B \cdot A^3 + B \cdot A^2 + B \cdot A + B \\
&= X(0) \cdot A^4 + B \cdot \frac{1-A^4}{1-A} \\
&= \frac{B}{1-A} + \left(X(0) - \frac{B}{1-A}\right) \cdot A^4 \qquad \text{if } A \neq 1 ,
\end{aligned}$$

and $X(4) = X(0) + 4 \cdot B$ if $A = 1$. Similarly, by an exact induction, one can derive that

$$X(n) = \frac{B}{1-A} + \left(X(0) - \frac{B}{1-A}\right) \cdot A^n \qquad \text{if } A \neq 1 ,$$

and $X(n) = X(0) + n \cdot B$ if $A = 1$. □

2. Absolute Distributions

The (unconditional) distribution of ξ_n is called the **n -th absolute distribution**. Let $d_i(n)$ denote the probability

$$d_i(n) = \Pr\Big(\xi_n = i\Big) ,$$

and let $D(n)$ denote the n-th absolute distribution

$$D(n) = \Big(d_1(n), d_2(n), \ldots\Big) = \Bigg(\Pr\Big(\xi_n = 1\Big), \Pr\Big(\xi_n = 2\Big), \ldots\Bigg).$$

The distribution $D(0)$ of ξ_0 is called the **initial distribution**:

$$D(0) = \Big(d_1(0), d_2(0), \ldots\Big) = \Bigg(\Pr\Big(\xi_0 = 1\Big), \Pr\Big(\xi_0 = 2\Big), \ldots\Bigg).$$

Using again the total probability formula, one can check that the absolute distributions and the transition matrices are related to each other according to the following matrix products:

$$D(n) = D(m) \cdot P(n-m) = D(m) \cdot P^{n-m},$$

$$D(n) = D(0) \cdot P(n) = D(0) \cdot P^{n}.$$

6. Example: TWO-STATE MARKOV CHAINS (continued) – ABSOLUTE DISTRIBUTIONS

Find an explicit formula for the absolute distributions of a two-state Markov chain in terms of the initial distribution $D(0) = \Big(d_1(0), d_2(0)\Big)$ and the transition probabilities $p_{12} = a$, $p_{21} = b$.

Solution.

As we calculated the n-step transition probabilities, if at least one of a and b differs from 0, then

$$P(n) = \begin{bmatrix} \frac{b}{a+b} + \frac{a}{a+b} \cdot (1-a-b)^n & \frac{a}{a+b} - \frac{a}{a+b} \cdot (1-a-b)^n \\ \frac{b}{a+b} - \frac{b}{a+b} \cdot (1-a-b)^n & \frac{a}{a+b} + \frac{b}{a+b} \cdot (1-a-b)^n \end{bmatrix}.$$

Thus,

$$\Big(d_1(n), d_2(n)\Big) = D(n) = D(0) \cdot P(n) =$$

$$= (d_1(0), d_2(0)) \cdot \begin{bmatrix} \frac{b}{a+b} + \frac{a}{a+b} \cdot (1-a-b)^n & \frac{a}{a+b} - \frac{a}{a+b} \cdot (1-a-b)^n \\ \frac{b}{a+b} - \frac{b}{a+b} \cdot (1-a-b)^n & \frac{a}{a+b} + \frac{b}{a+b} \cdot (1-a-b)^n \end{bmatrix},$$

yielding that

$$d_1(n) = \frac{b}{a+b} + \frac{a\,d_1 - b\,d_2}{a+b}(1-a-b)^n,$$

$$d_2(n) = \frac{a}{a+b} - \frac{a\,d_1 - b\,d_2}{a+b}(1-a-b)^n.$$

If $a = b = 0$, then $P(n)$ is the unit matrix, and thus, $D(n) = D(0)$ for all n.

7. Exercise.

Study the behavior of $P(n)$ and $D(n)$ as $n \to \infty$, and think of the physical meaning of the result you get. How do you interpret the result, when $a = b = 1$? □

3. The Distribution of a Markov Chain on R^∞

This and the next section can be skipped.

As known, the distribution Q on R^∞ of a random sequence $\xi_0, \xi_1, \xi_2, \xi_3 \ \ldots$ can be characterized by a primary or initial distribution Q_0 and a sequence of families of conditional distributions:

$$Q_1^{(x_0)} \ , \ Q_2^{(x_0, x_1)} \ , \ Q_3^{(x_0, x_1, x_2)} \ , \ \ldots \ .$$

Here, Q_0 represents the distribution of ξ_0, and

$$Q_n^{(x_0\, , \, x_1\, , \ldots, \, x_{n-2}\, , \, x_{n-1})}$$

represents the distribution of ξ_n on condition that $\xi_0 = x_0$, $\xi_1 = x_1$, $\ldots$, $\xi_{n-2} = x_{n-2}$, $\xi_{n-1} = x_{n-1}$. It should be clear that the random sequence $\xi_0, \xi_1, \xi_2, \xi_3, \ldots$ is a Markov chain if and only if its distribution Q on R^∞ has the property that for all $n \geq 1$ and for all $(x_0, x_1, \ldots, x_{n-2}, x_{n-1})$ the conditional distribution $Q_n^{(x_0\, , \, x_1\, , \, \ldots\, , \, x_{n-2}\, , \, x_{n-1})}$ does not depend on the parameters $x_0, x_1, \ldots, x_{n-2}$; it may depend only on the parameter x_{n-1} :

$$Q_n^{(x_0\, , \, x_1\, , \ldots, \, x_{n-2}\, , \, x_{n-1})} \ = \ Q_n^{(x_{n-1})} \ .$$

Thus, the distribution of a Markov chain is characterized by an initial distribution Q_0 and a sequence of family of special conditional distributions:

$$Q_1^{(x)} \ , \ Q_2^{(x)} \ , \ Q_3^{(x)} \ , \ \ldots \ .$$

(Since x_{n-1} represents an arbitrary real number, the index $n-1$ in x_{n-1} may be omitted.)

Homogeneity means that $Q_n^{(x)}$ does not depend on n :

$$Q_1^{(x)} \ = \ Q_2^{(x)} \ = \ Q_3^{(x)} \ = \ \ldots \ = \ Q^{(x)} \quad .$$

Thus, the distribution of a homogeneous Markov chain is characterized by an initial distribution Q_0 and a family of conditional distributions.

If, for the Markov chain, there exists only a finite or countably infinite number of states identified with the positive integers, then the initial distribution and each conditional distribution is a discrete distribution.

4. How to Generate a Markov Chain

This section can be skipped.

A random sequence of states, that is, a physical representation of a Markov chain with a given initial distribution Q_0 and sequence of families of conditional distributions $Q_1^{(x)}$, $Q_2^{(x)}$, $Q_3^{(x)}$, $\ldots$ can be generated as follows. First, choose the initial state according to the distribution Q_0 . Then, if the generated state turns out to be x_0 , then use the distribution $Q_1^{(x_0)}$ to generate the next state. If this is x_1 , then, regardless of x_0 , use $Q_2^{(x_1)}$ to generate the next element of the random sequence. Now, regardless of x_0 and x_1 , taking into account only the last generated random number x_2 , generate x_3 according to the conditional distribution $Q_3^{(x_2)}$. And so on, each element x_n of the random sequence is generated according to a distribution which is specified by the previous element of the random sequence.

If the chain is homogeneous, then the conditional distributions $Q_n^{(x)}$ ($n = 1, 2, ...$) do not depend on n :

$$Q_1^{(x)} = Q_2^{(x)} = Q_3^{(x)} = \dots = Q^{(x)} \qquad \text{for any state } x, \text{ for all } n .$$

For a homogeneous Markov chain, the element x_n , except x_0 , is generated according to the distribution $Q^{(x_{n-1})}$.

In order to perform this method in practice, let us denote by H_0 the upper quantile function, that is, the generalized inverse of the cumulative distribution function of the distribution Q_0 . For a given n and x let us denote by $H_n(x, y)$ the value of the upper quantile function of the distribution $Q_n^{(x)}$ at y . If the random variables $\eta_0, \eta_1, \eta_2, \eta_3, \dots$ are independent and uniformly distributed between 0 and 1 , then the random variables $\xi_0, \xi_1, \xi_2, \xi_3 \dots$ defined by

$$\begin{aligned} \xi_0 &= H_0(\eta_0) \\ \xi_1 &= H_1(\xi_0, \eta_1) \\ \xi_2 &= H_2(\xi_1, \eta_2) \\ \xi_3 &= H_3(\xi_2, \eta_3) \\ &\vdots \end{aligned}$$

constitute a Markov chain for which the primary distribution is Q_0 and the sequence of families of conditional distributions is $Q_1^{(x)}$, $Q_2^{(x)}$, $Q_3^{(x)}$, $\dots$. Clearly, the chain is homogeneous if, for $n \geq 1$, $H_n(x, y)$ does not depend on n , that is

$$\begin{aligned} \xi_0 &= H_0(\eta_0) \\ \xi_1 &= H(\xi_0, \eta_1) \\ \xi_2 &= H(\xi_1, \eta_2) \\ \xi_3 &= H(\xi_2, \eta_3) \\ &\vdots \end{aligned}$$

5. Stationarity

A Markov chain is called **stationary** if all its absolute distributions are the same, that is, $D(n)$ does not depend on n . It is easy to see that a Markov chain is stationary if and only if its initial distribution $D(0)$ satisfies the matrix equation:

$$D = D \cdot P .$$

This matrix equation clearly represents a system of linear equations for the terms $d_1, d_2, \dots$ of the distribution D . Since D is a normalized distribution, this system can be extended by the equation $d_1 + d_2 + \dots = 1$. For a given transition matrix P , a distribution D satisfying $D = DP$ and $d_1 + d_2 + \dots = 1$ is called a **stationary absolute distribution** for the transition matrix P . If the initial distribution $D(0)$ is a stationary absolute distribution D for the transition matrix, then the Markov chain is stationary because $D(n) = D$ for all n .

8. Example: COLOR OF THE CONTAINER

The M-bugs, which are tiny creatures, change their color by chance, independently of each other from day to day. They are either red or yellow or blue. An M-bug which is red today, tomorrow will be red or yellow with probability $\frac{1}{2}$, $\frac{1}{2}$. A yellow one changes its color to red or remains yellow with probability $\frac{1}{4}$, $\frac{1}{4}$, and becomes blue with probability $\frac{1}{2}$. A blue M-bug, the next day, turns into yellow or remains blue with probability $\frac{1}{2}$, $\frac{1}{2}$. Last year, I collected a large amount of such bugs, and have been keeping them in a glass container, where they are so mixed that the color of the whole container, from a certain distance, seems uniform all the time. You may be surprised, but in spite of the fact that each M-bug changes its color from day to day, this year the color of the container does not change from day to day, but it is a constant color. What color is it ?

Solution.

The color of an M-bug, observed each day, constitutes a Markov chain with states 'red', 'yellow', 'blue' and transition matrix

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} .$$

The large number of bugs guarantees that the proportion of M-bugs of a color in the container on a day is approximately equal to the absolute probability of that color on that day. Since the color of the container does not change from day to day, the proportion of red, yellow and blue M-bugs must be approximately constant from day to day. This means that the absolute probabilities of the colors are constants from day to day, that is, the absolute distribution is constant from day to day. This is why we calculate now the stationary absolute distribution $D = (d_1 , d_2 , d_3)$. Here is the equation $D P = D$:

$$\begin{bmatrix} d_1 & d_2 & d_3 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} d_1 & d_2 & d_3 \end{bmatrix} ,$$

or, the equivalent system of equations:

$$\begin{aligned} \tfrac{1}{2} d_1 + \tfrac{1}{4} d_2 \qquad\qquad &= d_1 \\ \tfrac{1}{2} d_1 + \tfrac{1}{4} d_2 + \tfrac{1}{2} d_3 &= d_2 \\ \tfrac{1}{2} d_2 + \tfrac{1}{2} d_3 &= d_3 \end{aligned}$$

It is easy to see that any of these three equations can be expressed by the other two, so any of them can be omitted. The remaining two, extended with normalization equation

$$d_1 + d_2 + d_3 = 1$$

constitutes a system of three equations with three unknowns, which can be uniquely solved. The solution is:

$$d_1 = \tfrac{1}{5} , \quad d_2 = \tfrac{2}{5} , \quad d_3 = \tfrac{2}{5} .$$

This means that the color of the container is the color which arises when 20 % red, 40 % yellow and 40 % blue colors are mixed. □

9. Example: DIFFUSION MODEL OF EHRENFEST (MOLECULE IS CHOSEN)

There are N molecules in a container, which is theoretically divided into two parts, say, left and right. Each time instant one molecule, chosen at random (each has the same probability) is put into the other half of the container. Consider the process defined by $\xi_t =$ the number of molecules in the left side at time instant t ($t = 0, 1, 2, \ldots$). It is easy to see that the process is a homogeneous Markov chain. Find the transition probabilities, and show that if the initial distribution is binomial with parameters N and 0.5, then the chain is stationary.

Solution.

If the process is in state i, then there are i molecules in the left part of the container. Thus, the probability of a transition into state $i-1$ is equal to

$$p_{i,i-1} = \frac{i}{N} ,$$

and the probability of a transition into state $i+1$ is

$$p_{i,i+1} = \frac{N-i}{N} .$$

These formulas can be written also as

$$p_{j+1,j} = \frac{j+1}{N} , \quad \text{and} \quad p_{j-1,j} = \frac{N-(j-1)}{N} .$$

Any other transition is impossible:

$$p_{i,j} = 0 \quad \text{if } |j-i| \neq 1 .$$

In order to prove that the distribution

$$d_i = \binom{N}{i} \frac{1}{2^N} \quad (i = 0, 1, 2, \ldots , N)$$

is a stationary absolute distribution, we calculate the j-th element of the matrix product $D \cdot P$, which is

$$\sum_{i=1}^{N} d_i \, p_{i,j} = d_{j-1} \cdot p_{j-1,j} + d_{j+1} \cdot p_{j+1,j}$$

$$= \binom{N}{j-1} \frac{1}{2^N} \cdot \frac{N-(j-1)}{N} + \binom{N}{j+1} \frac{1}{2^N} \cdot \frac{j+1}{N}$$

$$= \left(\frac{N!}{(j-1)!\,(N-j+1)!} \cdot \frac{N-j+1}{N} + \frac{N!}{(j+1)!\,(N-j-1)!} \cdot \frac{j+1}{N} \right) \frac{1}{2^N}$$

$$= \left(\frac{(N-1)!}{(j-1)!\,(N-j)!} + \frac{(N-1)!}{j!\,(N-j-1)!} \right) \frac{1}{2^N}$$

$$= \left(\binom{N-1}{j-1} + \binom{N-1}{j} \right) \frac{1}{2^N}$$

$$= \binom{N}{j} \frac{1}{2^N} .$$

This calculation shows that D P is equal to D , that is, D is really a stationary absolute distribution. □

The result of the next example can be used for many Markov chains.

10. Example: STAY OR JUMP ONLY ONTO ADJACENT STATES

Let the states of a Markov chain be the non-negative integers or an initial interval of the non-negative integers. Assume that jumps are possible only onto adjacent states, more precisely, $p_{ij} = 0$, whenever $|i-j| > 1$, and $p_{ij} > 0$, whenever $|i-j| = 1$.

a) Prove that the stationary absolute distribution $(d_0 , d_1 , d_2 , \ldots)$ of the transition matrix satisfies the relations:

$$\frac{d_1}{d_0} = \frac{p_{01}}{p_{10}} \quad , \quad \frac{d_2}{d_1} = \frac{p_{12}}{p_{21}} \quad , \quad \frac{d_3}{d_2} = \frac{p_{23}}{p_{32}} \quad , \quad \ldots \; .$$

b) Show that, if the sum

$$1 + \frac{p_{01}}{p_{10}} + \frac{p_{01}}{p_{10}}\frac{p_{12}}{p_{21}} + \frac{p_{01}}{p_{10}}\frac{p_{12}}{p_{21}}\frac{p_{23}}{p_{32}} + \ldots$$

is finite, then

$$d_1 = \frac{p_{01}}{p_{10}} \cdot d_0 \quad , \quad d_2 = \frac{p_{01}}{p_{10}}\frac{p_{12}}{p_{21}} \cdot d_0 \quad , \quad d_3 = \frac{p_{01}}{p_{10}}\frac{p_{12}}{p_{21}}\frac{p_{23}}{p_{32}} \cdot d_0 \quad , \quad \ldots \; ,$$

where

$$d_0 = \left(1 + \frac{p_{01}}{p_{10}} + \frac{p_{01}}{p_{10}}\frac{p_{12}}{p_{21}} + \frac{p_{01}}{p_{10}}\frac{p_{12}}{p_{21}}\frac{p_{23}}{p_{32}} + \ldots \right)^{-1} .$$

Solution.

a) Using the fact that $p_{ij} = 0$, whenever $|i-j| > 1$, the equation $D = D \cdot P$ means the following system of equations:

$$\begin{aligned}
d_0 &= d_0\, p_{00} + d_1\, p_{10} \\
d_1 &= d_0\, p_{01} + d_1\, p_{11} + d_2\, p_{21} \\
d_2 &= d_1\, p_{12} + d_2\, p_{22} + d_3\, p_{32} \\
d_3 &= d_2\, p_{23} + d_3\, p_{33} + d_4\, p_{43}
\end{aligned}$$

and so on.

The first of these equations yields that

$$d_0\,(1 - p_{00}) = d_1\,p_{10}\,, \qquad \text{that is,} \qquad d_0\,p_{01} = d_1\,p_{10}\,.$$

Substituting the last equation into the second of the above equations yields that

$$d_1 = d_1\,p_{10} + d_1\,p_{11} + d_2\,p_{21}\,,$$

from where

$$d_1\,(1 - p_{10} - p_{11}) = d_2\,p_{21}\,, \qquad \text{that is,} \qquad d_1\,p_{12} = d_2\,p_{21}\,.$$

Substituting the last equation into the third of the above equations yields that

$$d_2 = d_2\,p_{21} + d_2\,p_{22} + d_3\,p_{32}\,,$$

from where

$$d_2\,(1 - p_{21} - p_{22}) = d_3\,p_{32}\,, \qquad \text{that is,} \qquad d_2\,p_{23} = d_3\,p_{32}\,.$$

And so on, we get the sequence of the formulas

$$d_0\,p_{01} = d_1\,p_{10}\,, \qquad d_1\,p_{12} = d_2\,p_{21}\,, \qquad d_2\,p_{23} = d_3\,p_{32}\,, \qquad \ldots\;.$$

This sequence of formulas is equivalent to the sequence of formulas we wanted to prove.

b) Expressing d_1 , then d_2 , then d_3 , and so on, in terms of d_0 , we get that

$$d_1 = \frac{p_{01}}{p_{10}} \cdot d_0 \quad , \quad d_2 = \frac{p_{01}}{p_{10}}\,\frac{p_{12}}{p_{21}} \cdot d_0 \quad , \quad d_3 = \frac{p_{01}}{p_{10}}\,\frac{p_{12}}{p_{21}}\,\frac{p_{23}}{p_{32}} \cdot d_0 \quad , \quad \ldots\;,$$

Since the distribution $(d_0\,, d_1\,, d_2\,, d_3\,, \ldots\,)$ should be normalized, the statement follows from the equation

$$d_0 \cdot \left(1 + \frac{p_{01}}{p_{10}} + \frac{p_{01}}{p_{10}}\,\frac{p_{12}}{p_{21}} + \frac{p_{01}}{p_{10}}\,\frac{p_{12}}{p_{21}}\,\frac{p_{23}}{p_{32}} + \ldots \right) = 1\,.$$

11. Exercise

Pretend that the stationary absolute distribution for the Diffusion Model of Ehrenfest were not known. Using the method explained in the previous example, determine the stationary absolute distribution. ☐

The next three problems are slight modifications of each other.

12. Example: RANDOM WALK WITH REFLECTIVE BARRIERS

A random walk on the states $0\,,1\,,2\,,3$ with reflective barriers means that from the endpoints 0 and 3 a jump is possible only into the states 1 and 2 , respectively, and for all other states both a left or right jump has a probability p and q , respectively $(\,p + q = 1\,)$. Find the transition matrix and the stationary absolute distribution for the symmetric walk, which means $p = \frac{1}{2}$.

Solution.

The transition matrix is obviously

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Thus, the system of equations

$$d_0\, p_{01} = d_1\, p_{10}\ ,\quad d_1\, p_{12} = d_2\, p_{21}\ ,\quad d_2\, p_{23} = d_3\, p_{32}$$

for the stationary absolute distribution is

$$d_0 = \tfrac{1}{2}\, d_1\ ,\quad \tfrac{1}{2}\, d_1 = \tfrac{1}{2}\, d_2\ ,\quad \tfrac{1}{2}\, d_2 = d_3\ .$$

Since the stationary absolute distribution should be normalized, the solution is

$$d_0 = \tfrac{1}{6}\ ,\quad d_1 = \tfrac{2}{6}\ ,\quad d_2 = \tfrac{2}{6}\ ,\quad d_3 = \tfrac{1}{6}\ .$$

13. Exercise.

Find the stationary absolute distribution for the non-symmetric case, too. □

14. Example: RANDOM WALK WITH ABSORBING BARRIERS

Modify the random walk in the previous problem the following way: from the endpoints 0 and 3 a transition is possible only into themselves, and from all other states both a left or right jump has a probability p and q , respectively $(\, p + q = 1\,)$. Find the transition matrix and the stationary absolute distributions.

Solution.

The transition matrix is obviously

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ p & 0 & q & 0 \\ 0 & p & 0 & q \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Thus, the system of equations

$$d_0\, p_{01} = d_1\, p_{10}\ ,\quad d_1\, p_{12} = d_2\, p_{21}\ ,\quad d_2\, p_{23} = d_3\, p_{32}\ .$$

for the stationary absolute distribution is

$$0 = p\, d_1\ ,\quad q\, d_1 = p\, d_2\ ,\quad q\, d_2 = 0\ .$$

This yields that $d_1 = d_2 = 0$ and d_0 and d_3 are non-negative numbers whose sum is equal to 1 . Realize that the stationary absolute distribution is not unique for this process. □

15. Example: RANDOM WALK WITH SEMI-REFLECTIVE BARRIERS

Mix the random walks of the previous two problems by assuming that for the endpoints both a jump into the adjacent state and staying at the same endpoint has a probability 0.5 , and for all other states both a left and a right jump has a probability 0.5 . Find the transition matrix and the stationary absolute distribution.

Solution.

The transition matrix is obviously

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

Thus, the system of equations

$$d_0\, p_{01} = d_1\, p_{10}\,, \quad d_1\, p_{12} = d_2\, p_{21}\,, \quad d_2\, p_{23} = d_3\, p_{32}$$

for the stationary absolute distribution is

$$\frac{1}{2}\, d_0 = \frac{1}{2}\, d_1\,, \quad \frac{1}{2}\, d_1 = \frac{1}{2}\, d_2\,, \quad \frac{1}{2}\, d_2 = \frac{1}{2}\, d_3\,,$$

which yields that the stationary absolute distribution is the uniform distribution:

$$d_0 = d_1 = d_2 = d_3 = \frac{1}{4}\,.$$

16. Exercise.

Find the stationary absolute distribution for the non-symmetric case, too. □

6. Recurrence

Let $f_{ij}(n)$ denote the conditional probability that starting from state i at time instant 0 , the process avoids state j before time instant n , and at time instant n it is in state j . It is clear that the sum

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}(n)$$

is the probability that starting from state i , the process will ever get to state j in the future.

As a special case, f_{ii} means the probability that starting from state i , the process will ever return to state i . If $f_{ii} = 1$, then the state is called **recurrent**. If $f_{ii} < 1$, then the state i is called **non-recurrent**.

For a recurrent state i , $\left(f_{ii}(1) , f_{ii}(2) , \dots \right)$ is clearly the distribution of the number of steps needed to return into state i , and the finite or infinite number

$$m_i = \sum_{n=1}^{\infty} n \cdot f_{ii}(n)$$

clearly means the **expected value of steps** (or **length of time) needed to return into state i in the future**. For a non-recurrent state, let $m_i = \infty$. The state i is called **positive** if $m_i < \infty$. If $m_i = \infty$, then the state i is called a **null state**. The names 'positive' and 'null' will become justified by the statement of the theorem bellow which implies that, under some conditions,

$$\lim_{n \to \infty} d_i(n) = d_i(\infty) > 0 \quad \text{if } m_i < \infty ,$$

and

$$\lim_{n \to \infty} d_i(n) = d_i(\infty) = 0 \quad \text{if } m_i = \infty .$$

17. Example: LUNCH MONEY

John keeps some dollars in his schoolbag so that he could pay for his lunch at school every day. The lunch costs one dollar each day. When John's money has gone, then in the evening he asks his father to give him money. In order to generate the amount, his father tosses a special die until the first occurrence of the face with 'MONEY' on it, and then gives his son as many dollars as many tosses were needed. The probability of the occurrence of the face with 'MONEY' is p at each toss. Observing the amount of money John has in his schoolbag in the morning of each school-day, we clearly get a homogeneous, recurrent Markov chain with states 1, 2, 3, $\dots$ and transition matrix

$$\begin{bmatrix} p & pq & pq^2 & \dots \\ 1 & 0 & 0 & \\ 0 & 1 & 0 & \\ \vdots & & \ddots & \end{bmatrix} ,$$

where $q = 1 - p$. Clearly, John has to ask for money in the evening of a day if that day he went to school with 1 dollar only.

We shall determine the distribution and the expected value of the number of days between the days when John has to ask his father for money, more precisely, we shall determine the distribution $\left(f_{11}(1) , f_{11}(2) , f_{11}(3) , \dots \right)$ and the expected value m_1 of the number of steps needed to return into state 1 .

Solution.

A return to state 1 needs n steps if and only if the father tosses 'MONEY' the first time at the n -th toss. Thus, $f_{11}(n) = pq^{n-1}$ $(n = 1, 2, 3, \dots)$, that is, $\left(f_{11}(1) , f_{11}(2) , f_{11}(3) , \dots \right)$ is the geometrical distribution with parameter p . As we know well its expected value is $m_1 = \frac{1}{p}$.

18. Exercise.

Determine the distribution and the expected value of the number of days between the mornings when John goes to school with 2 dollars in his bag, more precisely, determine the distribution $\left(f_{22}(1), f_{22}(2), f_{22}(3), \dots \right)$ and the expected value m_2 of the number of steps needed to return into state 2 . Prove that $f_{22}(1) = 0$, and

$$f_{22}(n) = (n-1)\,\frac{1}{2^n} \quad \text{if } p = \frac{1}{2} \ ,$$

and

$$f_{22}(n) = \sum_{k=1}^{n-1} p^{n-k} q^k = p\,q\,\frac{p^{n-1} - q^{n-1}}{p-q} \quad \text{if } p \neq \frac{1}{2} \ ,$$

$(n = 2, 3, \dots)$, and then find a closed formula for m_2 . □

19. Exercise.

Find $f_{ij}(n)$ and m_i for a two-state chain. □

The number of time instants in the future when the process is in state j on condition that it starts from state i is the sum of an infinite number of indicator random variables: the n-th indicator random variable is equal to 1 if the process is in state j at time instant n , and 0 otherwise. Since the expected value of such a sum is clearly equal to the sum of the parameters of the indicator distributions, we get that the sum

$$g_{ij} = \sum_{n=1}^{\infty} p_{ij}(n)$$

can be interpreted as the **expected value of the number of time instants in the future when the process is in state j on condition that it starts from state i** . Specifically, g_{jj} is the **expected value of the number of time instants in the future when the process returns into state j on condition that it starts from state j** .

It is clear that starting from state i , the probability of stepping onto state j exactly k times in the future is $1 - f_{ij}$ for $k = 0$, and $f_{ij} \cdot (f_{jj})^{k-1} \cdot (1 - f_{jj})$ for $k = 1, 2, \dots$. The expected value of this distribution is

$$g_{ij} = \frac{f_{ij}}{1-f_{jj}} \ .$$

Specifically, when $i = j$, we get that

$$g_{jj} = \frac{f_{jj}}{1-f_{jj}} \ .$$

This relation shows that if $f_{ii} < 1$, then $g_{ii} < \infty$. Equivalently: if $g_{ii} = \infty$, then $f_{ii} = 1$, that is, state i is recurrent.

20. Exercise.

Show that $g_{ij} = f_{ij} \cdot (1 + g_{jj})$. □

21. Example: RANDOM WALK ON THE LINE

Consider the random walk on the integer numbers of the real line, when the probability of a left or right step is p or q , respectively. Consider a state i .

a) Find the n-step transition probability $p_{ii}(n)$.

b) Use the Stirling formula to approximate $p_{ii}(n)$.

c) Using the approximation, prove that the series $\sum_{n=1}^{\infty} p_{ii}(n)$ is divergent if $p = q = 0.5$, and convergent otherwise.

d) Draw the conclusion about the problem of recurrence or non-recurrence.

Solution.

a) $p_{ii}(n)$ is the probability that starting from state i, the process, n units of time later, is back in i. This is impossible if n is odd. If n is even, and $n = 2N$, then a return in $2N$ steps means that among the $2N$ steps there are N left and N right steps. The probability of this event is

$$p_{ii}(2N) = \binom{2N}{N} p^N q^N .$$

b) The Stirling formula states that

$$N! \sim \left(\frac{N}{e}\right)^N \sqrt{2\pi N} .$$

Using this approximation, we get that

$$\binom{2N}{N} p^N q^N = \frac{(2N)!}{(N!)^2} p^N q^N$$

$$\sim \frac{\left(\frac{2N}{e}\right)^{2N} \sqrt{2\pi\, 2N}}{\left(\left(\frac{N}{e}\right)^N \sqrt{2\pi N}\right)^2} p^N q^N = \frac{1}{\sqrt{\pi N}} \left(4pq\right)^N .$$

that is,

$$p_{ii}(2N) \sim \frac{1}{\sqrt{\pi N}} \left(4pq\right)^N .$$

c) Using this approximation, we get that

$$\sum_{n=1}^{\infty} p_{ii}(n)$$

is convergent if and only if

$$\sum_{N=1}^{\infty} \frac{1}{\sqrt{\pi N}} \left(4pq\right)^N$$

is convergent.

If $p = q = \frac{1}{2}$, then this series reduces to

$$\sum_{N=1}^{\infty} \frac{1}{\sqrt{\pi N}} \,,$$

which is clearly divergent.

If $p \neq \frac{1}{2}$, then $4pq < 1$, and thus

$$\sum_{N=1}^{\infty} \frac{1}{\sqrt{\pi N}} \left(4pq \right)^{N}$$

is obviously convergent.

d) The above investigation shows that the symmetric random walk on the real line is recurrent, and any non-symmetric one is non-recurrent. □

22. Example: SYMMETRIC RANDOM WALK WITH INDEPENDENT COORDINATES ON THE PLANE

Consider the symmetrical random walk on the points with integer coordinates on the plane, when the two coordinates of the displacements of the 'moving bug' are independent of each other: both coordinates change by $+1$ or -1 with probability $\frac{1}{2}$. Prove that the process is recurrent.

Solution.

For a return into a point on the plane, both coordinates have to return. Since the coordinates are independent, and the probability of a return in $n = 2N$ steps is

$$\binom{2N}{N} \left(\frac{1}{2} \right)^{2N}$$

for the one-dimensional symmetrical random walk, the probability of a return in $n = 2N$ steps is

$$\left(\binom{2N}{N} \left(\frac{1}{2} \right)^{2N} \right)^{2}$$

for the two-dimensional symmetrical random walk. In order to show that the process is recurrent, the sum

$$\sum_{N=1}^{\infty} \left(\binom{2N}{N} \left(\frac{1}{2} \right)^{2N} \right)^{2}.$$

should be examined. Since this sum is equiconvergent to the sum

$$\sum_{N=1}^{\infty} \left(\frac{1}{\sqrt{\pi N}} \right)^{2} = \frac{1}{\pi} \sum_{N=1}^{\infty} \frac{1}{N} \,,$$

which is divergent, we draw the conclusion that the process is recurrent. □

23. Example: SYMMETRIC RANDOM WALK WITH INDEPENDENT COORDINATES IN THE SPACE

Consider the symmetrical random walk on the points with integer coordinates in the space, when the three coordinates of the displacements of the 'moving bug' are independent of each other: each coordinate changes by $+1$ or -1 with probability $\frac{1}{2}$. Prove that the process is non-recurrent.

Solution.

Knowing the previous problem, it is clear that the series

$$\sum_{N=1}^{\infty}\left(\binom{2N}{N}\left(\frac{1}{2}\right)^{2N}\right)^3 .$$

should be examined. Since this sum is equiconvergent to the sum

$$\sum_{N=1}^{\infty}\left(\frac{1}{\sqrt{\pi N}}\right)^3 = \frac{1}{\pi^{\frac{3}{2}}}\sum_{N=1}^{\infty}\frac{1}{N^{\frac{3}{2}}} ,$$

which is convergent, we draw the conclusion that the process is non-recurrent. □

7. Irreducibility, Aperiodicity

We say that state j is **available** from state i if the probability of getting ever to j from i is positive: $f_{ij} > 0$. This is obviously equivalent to the condition that there exists such an n that $p_{ij}(n) > 0$. A Markov chain is called **irreducible** if all states are available from any state.

The **period** δ_i of a state i is the greatest common divisor (abbreviated by g.c.d.) of those step-numbers n for which the return into the state i with n steps has a positive probability:

$$\delta_i = \text{g. c. d.}\left\{ n : p_{ii}(n) > 0 \right\} .$$

It can be shown that the period of a state i is δ_i if and only if, for any sufficiently large step-number n which is a multiple of δ_i , the return into the state i has a positive probability, that is, there exists such a critical step-number n_0 , which may depend on i , that $p_{ii}(n) > 0$ whenever $n = k \cdot \delta_i$ and $n \geq n_0$.

A state i is called **aperiodic**, if its period is equal to 1 , that is, there exists a critical step-number n_0 , which may depend on i , that $p_{ii}(n) > 0$ whenever $n \geq n_0$.

It can be shown that the properties introduced now either hold simultaneously for all or for none of the states of an irreducible Markov chain. We refer to this fact by saying that these properties are **class properties**. (The notion of a **class** is described is the problem entitled Classes of States.)

Theorem. (The properties listed below are Class Properties)

If an irreducible Markov chain has a recurrent state, then all the states are recurrent. (So the whole chain can be called **recurrent**.)

If an irreducible Markov chain has a positive state, then all the states are positive. (So the whole chain can be called **positive**.)

The period of any two states of an irreducible Markov chain is the same. (So the whole chain can be assigned a certain **period**.)

If an irreducible Markov chain has an aperiodic state, then all the states are aperiodic. (So the whole chain can be called **aperiodic**.)

24. Example: RANDOM WALKS WITH DIFFERENT BARRIERS

Examine the random walk on the states 0 , 1 , 2 , 3 with a) reflective, b) absorbing, c) semi-reflective barriers. Decide which of these Markov chains are irreducible. Find the period of the irreducible ones.

Solution.

The reader should be convinced that the random walk with reflective barriers is irreducible, and its period is 2 . The random walk with absorbing barriers is not irreducible, because from states 0 and 3 none of the other states are available. The random walk with semi-reflective barriers is also irreducible, and its period is 1 . □

Notice that all the above introduced notions with the exception of the absolute distributions and stationarity are related only to the transition matrix of the chain, and are not related to the initial distribution of the chain. Thus, it is meaningful to call a **transition matrix recurrent or non-recurrent, positive or null, irreducible, aperiodic**.

8. Stability

A transition matrix P is called **stable** if its n -th power, that is, the n -step transition matrix $P(n)$, when $n \to \infty$, approaches such a limit matrix that each row of the limit matrix consists of the same, normalized distribution consisting of positive terms:

$$\lim_{n \to \infty} p_{ij}(n) = p_j(\infty) > 0 \qquad (i, j = 1, 2, \ldots),$$

$$\sum_{j=1}^{\infty} p_j(\infty) = 1 .$$

A transition matrix P is called **transient** if its n -th power, that is, the n -step transition matrix $P(n)$, when $n \to \infty$, approaches the 0 matrix:

$$\lim_{n \to \infty} p_{ij}(n) = p_j(\infty) = 0 \qquad (i, j = 1, 2, \ldots).$$

A Markov chain is called **stable** if the n -th absolute distribution $D(n)$, when $n \to \infty$, approaches a normalized distribution consisting of positive terms:

$$\lim_{n \to \infty} d_j(n) = d_j(\infty) > 0 \qquad (j = 1, 2, \ldots),$$

$$\sum_{j=1}^{\infty} d_j(\infty) = 1 .$$

A Markov chain is called **transient** if the n -th absolute distribution $D(n)$, when $n \to \infty$, approaches the 0 vector:

$$\lim_{n \to \infty} d_j(n) = d_j(\infty) = 0 \qquad (j = 1, 2, \ldots).$$

The following theorem describes the asymptotic behavior of irreducible, aperiodic Markov chains.

Theorem. (Limit Theorem for Markov Chains)

1. (For Finite Markov Chains)

For an irreducible, aperiodic Markov chain with a finite number of states the following statements hold:

i) The chain is a positive chain, that is, $m_j < \infty$ for all j.

ii) The transition matrix as well as the chain itself are stable, and the limit of the transition and absolute probabilities is equal to the reciprocal of the average return times:

$$\lim_{n \to \infty} p_{ij}(n) = p_j(\infty) = \lim_{n \to \infty} d_j(n) = d_j(\infty) = \frac{1}{m_j} \qquad (i, j = 1, 2, \ldots).$$

iii) There is a unique normalized distribution D solving the equation $D = D \cdot P$, namely,

$$D = \left(\frac{1}{m_1}, \frac{1}{m_2}, \ldots \right)$$

is the only stationary absolute distribution for the transition matrix.

2. (For Infinite Markov Chains)

An irreducible, aperiodic Markov chain with an infinite number of states is

either

i) a positive chain, that is, $m_j < \infty$ for all j, and

ii) the transition matrix as well as the chain itself are stable, and the limit of the transition and absolute probabilities is equal to the reciprocal of the average return times:

$$\lim_{n \to \infty} p_{ij}(n) = p_j(\infty) = \lim_{n \to \infty} d_j(n) = d_j(\infty) = \frac{1}{m_j} \qquad (i, j = 1, 2, \ldots), \text{ and}$$

iii) there is a unique normalized distribution D solving the equation $D = D \cdot P$, namely,

$$D = \left(\frac{1}{m_1}, \frac{1}{m_2}, \ldots \right)$$

is the only stationary absolute distribution for the transition matrix,

or

i) a null chain, that is, $m_j = \infty$ for all j, and

ii) the transition matrix as well as the chain itself are transient:

$$\lim_{n \to \infty} p_{ij}(n) = p_j(\infty) = \lim_{n \to \infty} d_j(n) = d_j(\infty) = 0 \qquad (i, j = 1, 2, \ldots), \text{ and}$$

iii) there is no normalized solution D to the equation $D = D \cdot P$, and thus, no stationary absolute distribution exists for the transition matrix.

Proof.

We shall prove only the statement for finite Markov chains. (See the last section of this chapter and the problem entitled 'Average Return Time = Reciprocal of Limit Probability'.) □

One can check the statement of the theorem for the very special Markov chains consisting of two states only.

25. Example: TWO-STATE MARKOV CHAINS (continued) – ASYMPTOTIC BEHAVIOR

Solution.

We calculated that, if at least one of a and b differs from 0 , then

$$p_{11}(n) = \frac{b}{a+b} + \frac{a}{a+b} \cdot (1-a-b)^n ,$$

$$p_{12}(n) = \frac{a}{a+b} - \frac{a}{a+b} \cdot (1-a-b)^n ,$$

$$p_{21}(n) = \frac{b}{a+b} - \frac{b}{a+b} \cdot (1-a-b)^n ,$$

$$p_{22}(n) = \frac{a}{a+b} + \frac{b}{a+b} \cdot (1-a-b)^n .$$

If, moreover, at least one of a and b differs not only from 0 , but from 1 as well, then $|1-a-b| < 1$, which guarantees that $(1-a-b)^n \to 0$, when $n \to \infty$, that is,

$$\lim_{n \to \infty} P(n) = \begin{bmatrix} \frac{b}{a+b} & \frac{a}{a+b} \\ \frac{b}{a+b} & \frac{a}{a+b} \end{bmatrix} .$$

Such a chain is stable.

If $a = b = 0$, then obviously $p_{11}(n) = 1$, $p_{12}(n) = 0$, $p_{21}(n) = 0$, $p_{22}(n) = 1$, that is $P(n)$ is the unit matrix for all n . Such a chain is not irreducible.

If $a = b = 1$, then the chain is not aperiodic, because its period is obviously equal to 2 . □

26. Example: RANDOM WALK ON THE POSITIVE HALF OF THE REAL LINE

The states of this random walk are the non-negative integers. From any state i , except 0 , a left or a right jump has a probability $p_{i,i-1} = p$ and $p_{i,i+1} = q$, respectively ($p, q > 0$, $p + q = 1$) . Staying in 0 has a probability $p_{0,0} = p$, and the probability of a right jump from 0 is $p_{0,1} = q$. This Markov chain is clearly irreducible and aperiodic. Thus, according to the above theorem, it is either stable, or instable. Decide how stability or instability depends on the parameter p .

Solution.

Let's try to decide whether, for a given value of p , stationary absolute distribution exists or does not exist. Since a jump is possible only onto adjacent states (see problem entitled 'Stay or Jump Only onto Adjacent States'), the relations

$$\frac{d_1}{d_0} = \frac{p_{0,1}}{p_{1,0}} , \quad \frac{d_2}{d_1} = \frac{p_{1,2}}{p_{2,1}} , \quad \frac{d_3}{d_2} = \frac{p_{2,3}}{p_{3,2}} , \quad \ldots$$

should hold for the stationary absolute distribution $(d_0, d_1, d_2, \ldots)$. Since the right side of each of these relations is equal to $\frac{q}{p}$, the sequence $d_0, d_1, d_2, \ldots$ constitutes a geometrical sequence:

$$d_n = d_0 \left(\frac{q}{p} \right)^n \quad (n = 0, 1, 2 \ldots) .$$

Thus, the distribution $(d_0 , d_1 , d_2 , \ldots)$ is normalizable if and only if $\frac{q}{p} < 1$, that is, if and only if $p > \frac{1}{2}$. This means that the chain has a stationary absolute distribution and it is stable if and only if $p > \frac{1}{2}$.

Remark.

Use your common sense to realize that this result is quite natural. □

27. Example: DIFFUSION MODEL OF BERNOULLI (continued)

Recall the model. Assume now that initially all the N red molecules are in the left side. After a long time you look at the system, and count the number of red molecules in the left side. What can you say about the distribution of this random variable?

Solution.

The chain is irreducible and aperiodic. (Check why.) The limit $\lim_{n \to \infty} d_j(n) = d_j(\infty)$ $(j = 0 , 1 , 2 , \ldots)$, of the absolute distribution is the stationary absolute distribution, which, as we calculated, is the hypergeometrical distribution with parameters $2N, N, N$. Thus, the random variable in question is distributed approximately according to this distribution, that is, observing the system at a large time instant

$$\Pr(\text{there are } j \text{ red molecules in the left side}) = \frac{\binom{N}{j}\binom{N}{N-j}}{\binom{2N}{N}} \qquad (j = 0 , 1 , 2 , \ldots , N) .$$

□

28. Exercise: DIFFUSION MODEL OF EHRENFEST (continued)

Recall the model. Assume now that initially

a) all the molecules are in the left side;

b) the number of balls at the left side is distributed according to an initial distribution $D(0) = \Big(d_1(0) , d_2(0) , \ldots \Big)$.

Realize that the chain is irreducible but not aperiodic, because the period is 2. After a long time you look at the system, and count the number of molecules in the left side. What can you say about the distribution of this random variable? □

9. Reducible, Periodic Markov Chains

In order to describe non-irreducible or periodic Markov chains, the phase space should be partitioned into classes and sub-classes.

By definition, two different states i and j belong to the same **class**, if the two states are **communicating** with each other, which means that both f_{ij} and f_{ji} are positive. It can be shown that the periods of states belonging to the same class are equal, so a class can be assigned a period. If the period of a class C is δ, then the class can be cyclically ordered into δ sub-classes $C_1 , \ldots , C_\delta$ so that, for any two states i and j belonging to the class, $p_{ij} > 0$ holds only if the sub-class of j follows the sub-class of i in the cyclic order on the sub-classes $C_1 , \ldots , C_\delta$.

By means of these notions, the above theorems can be generalized. We omit this generalization. (For more detail see problems entitled 'Classes of States' and 'Recurrence and Positivity as Class Properties'.)

10. Foster's Criterion for Stability

The following theorem offers a sufficient condition for the stability of a type of Markov chains which often occurs in applications.

If a Markov chain is at a state i , and then jumps from i onto j , then the difference $j - i$ is called the **move** from i . Clearly, the mean value of the move, called **average move**, from i is

$$\sum_j (j - i) \cdot p_{ij} \text{ , or equivalently, } \left(\sum_j j \cdot p_{ij} \right) - i .$$

Let the states of the homogeneous Markov chain $\xi_0 , \xi_1 , \xi_2 , \ldots$ be the nonnegative integers. Suppose that the chain is irreducible and aperiodic. Assume that the average move from any state i is finite. Assume also that there exist such a critical state I and such a negative number $-\delta$ that the average move from any state larger than I is less than $-\delta$:

$$\sum_j (j - i) \cdot p_{ij} \le -\delta \qquad (i > I) .$$

We shall prove that under these conditions the chain is stable.

Remark.

The above assumptions can be verbalized by saying that "the average move from any state is finite, and the average move from sufficiently large states is strictly negative". This interpretation may make the statement quite plausible.

Proof.

If the chain were not stable but transient, then the limit of $P(\xi_n \le J)$, when $n \to \infty$, would be 0 for all J . We shall show that the limit of $P(\xi_n \le I)$, when $n \to \infty$, is positive. Let us suppose that, on the contrary, $\lim_{n \to \infty} P(\xi_n \le I) = 0$. We show that this assumption leads to a contradiction.

Let us denote the maximum of $\sum_j j \cdot (p_{ij} - i)$ $(i = 0, 1, \ldots I)$ by C . We shall analyze the sequence of the expected values of the absolute distributions

$$m_n = \sum_i i \cdot d_i(n) \qquad (n = 1, 2, \ldots)$$

The following argument should be clear:

$$m_{n+1} = \sum_j j \cdot d_j(n+1) = \sum_j j \cdot \left(\sum_i d_i(n) \cdot p_{ij} \right) =$$

$$= \sum_i d_i(n) \cdot \left(\sum_j j \cdot p_{ij} \right)$$

Taking the difference, we get that

$$m_{n+1} - m_n = \sum_i d_i(n) \cdot \left(\sum_j (j - i) \cdot p_{ij} \right) =$$

$$= \sum_{i \le I} d_i(n) \cdot \left(\sum_j (j - i) \cdot p_{ij} \right) + \sum_{i > I} d_i(n) \cdot \left(\sum_j (j - i) \cdot p_{ij} \right)$$

For $i \le I$, $\sum_j (j - i) \cdot p_{ij} \le C$, and for $i > I$, $\sum_j (j - i) \cdot p_{ij} \le -\delta$. Thus

$$m_{n+1} - m_n \le \left(\sum_{i \le I} d_i(n) \right) \cdot C + \left(\sum_{i > I} d_i(n) \right) \cdot \left(-\delta \right).$$

Since we assumed that $\lim_{n \to \infty} P(\xi_n \le I) = 0$, we get

$$\left(\sum_{i \le I} d_i(n) \right) \cdot C \le \frac{\delta}{4}, \text{ and } \left(\sum_{i > I} d_i(n) \right) \cdot \left(-\delta \right) \le -\frac{3}{4}\delta.$$

These yield that for sufficiently large n

$$m_{n+1} - m_n \le -\frac{\delta}{2}.$$

This implies that $m_n \to -\infty$ contradicting to the fact $\xi_n \ge 0$. □

11. Stability of the Length of the Queue

Let the random variables ξ_0 and $\theta_1, \theta_2, \ldots$ and $\eta_1, \eta_2, \ldots$ be independent, nonnegative integer valued random variables, and let each of the sequences $\theta_1, \theta_2, \ldots$ and $\eta_1, \eta_2, \ldots$ consist of identically distributed random variables. Define the process $\xi_0, \xi_1, \xi_2, \ldots$ by recursion:

$$\xi_n = (\xi_{n-1} - \theta_n)^+ + \eta_n \qquad (n = 1, 2, \ldots),$$

where $(x)^+$ means the positive part of x, that is, x if x is positive, and 0 if x is zero or negative. Clearly, the process $\xi_0, \xi_1, \xi_2, \ldots$ is a homogeneous Markov chain with states $0, 1, 2, \ldots$. The random variable ξ_n can be interpreted as the **length of a queue** at time instant n, if $\theta_1, \theta_2, \ldots$ mean the capacity of the shop to serve customers during each time interval, and $\eta_1, \eta_2, \ldots$ mean how many customers arrive to the shop during each time interval. The formula for ξ_n expresses the assumption that the η_n customers arriving between the time instants $n-1$ and n are not served before the time instant n, even if the capacity θ_n of the shop exceeds the length ξ_{n-1} of the queue.

Theorem. (Sufficient Condition for the Stability of the Length of the Queue)

Suppose that the homogeneous Markov chain $\xi_0, \xi_1, \xi_2, \ldots$ defined above is irreducible and aperiodic. If, moreover, $M(\eta_1) < M(\theta_1) < \infty$, then the conditions of the Foster criterion hold, and thus, the process $\xi_0, \xi_1, \xi_2, \ldots$ is stable.

Remark.

Heuristically the statement about the stability of the length of the queue is clear, since the assumption $M(\eta_1) < M(\theta_1)$ means that the capacity of the shop to serve customers is greater than the average number of the arriving customers.

Proof.

It is clear that that the move from state i can be written like this:

$$(i - \theta_1)^+ + \eta_1 - i = \eta_1 - \min(i, \theta_1)$$

Since the expected value of $\min(i , \theta_1)$ approaches the expected value of θ_1 , when $i \to \infty$, we get that the expected value of the move from state i approaches $M (\eta_1) - M (\theta_1)$. This guarantees that, on one hand, the expected value of the move from state i is finite for all i , and on the other hand, there exists a negative number $(- \delta)$ such that the expected value of the move from state i , for sufficiently large i , is less than $(- \delta)$. □

29. Example: AVERAGE WAITING TIME FOR A PERSON IN QUEUE

At the end of this example it will turn out that the average of the amount of time that a person who gets into the queue has to spend in the queue is approximately equal to the quotient of the expected value of the stationary queue length and the expected value of the number of people arriving to the shop during a time interval.

For any time instant t , let us introduce the following random variables:

ρ_t = the total amount of time that people who get into the queue before the time instant t altogether spend in the queue;

α_t = the total amount of time that people who get into the queue before the time instant t altogether have spent in the queue before the time instant t ;

β_t = the total amount of time that people who get into the queue before the time instant t altogether will spend in the queue after the time instant t ;

ν_t = the number of people getting into the queue before the time instant t ;

τ_t = the average of the amount of time that people who get into the queue before the time instant t have to spend in the queue. Be convinced that

$$\rho_t = \alpha_t + \beta_t ;$$

$$\alpha_t = \xi_0 + \xi_1 + \ldots + \xi_t ;$$

$$\nu_t = \xi_0 + \eta_1 + \ldots + \eta_t ;$$

a) Check that the following equalities are correct:

$$\tau_t = \frac{\rho_t}{\nu_t} = \frac{\xi_0 + \xi_1 + \ldots + \xi_t + \beta_t}{\xi_0 + \eta_1 + \ldots + \eta_t} = \frac{\frac{\xi_0}{t} + \frac{\xi_1 + \ldots + \xi_t}{t} + \frac{\beta_t}{t}}{\frac{\xi_0}{t} + \frac{\eta_1 + \ldots + \eta_t}{t}} .$$

b) Try to find a heuristic explanation why $\frac{\xi_0}{t}$ and , under some circumstances, $\frac{\beta_t}{t}$ may be negligible for large values of t , and thus, by the law of large numbers the average of the amount of time that a person who gets into the queue has to spend in the queue is

$$\tau_t = \frac{\frac{\xi_0}{t} + \frac{\xi_1 + \ldots + \xi_t}{t} + \frac{\beta_t}{t}}{\frac{\xi_0}{t} + \frac{\eta_1 + \ldots + \eta_t}{t}} \approx \frac{0 + m_\xi + 0}{0 + m_\eta} = \frac{m_\xi}{m_\eta} ,$$

where m_ξ means the expected value of the stationary queue length, and m_η means the expected value of the number of people arriving to the shop during a time interval.

Solution.

The steps to figure out are clearly exposed. The details are left for the reader. □

30. Example: STATIONARY QUEUE LENGTH (continued)

Assume now that the common distribution of θ_1 , θ_2 , ... is the indicator distribution with parameter q , and the common distribution of η_1 , η_2 , ... is the indicator distribution with parameter p . Find the transition matrix and the stationary absolute distribution for the process ξ_0 , ξ_1 , ξ_2 , ...

Solution.

The transition matrix looks like as this:

$$\begin{bmatrix} 1-p & p & 0 & 0 & \\ (1-p)q & (1-p)(1-q)+pq & p(1-q) & 0 & \\ 0 & (1-p)q & (1-p)(1-q)+pq & p(1-q) & \\ & & \ddots & \ddots & \ddots \end{bmatrix}$$

Thus, the relations

$$\frac{d_1}{d_0} = \frac{p_{01}}{p_{10}} \quad , \quad \frac{d_2}{d_1} = \frac{p_{12}}{p_{21}} \quad , \quad \frac{d_3}{d_2} = \frac{p_{23}}{p_{32}} \quad , \quad \ldots$$

for the stationary absolute distribution are now

$$\frac{d_1}{d_0} = \frac{p}{(1-p)\,q} \quad , \quad \frac{d_2}{d_1} = \frac{p\,(1-q)}{(1-p)\,q} \quad , \quad \frac{d_3}{d_2} = \frac{p\,(1-q)}{(1-p)\,q} \quad , \quad \ldots ,$$

which means that, neglecting the term d_0 from the sequence d_0 , d_1 , d_2 , ... , the sequence d_1 , d_2 , d_3 , ... constitutes a geometrical sequence with quotient

$$\frac{p\,(1-q)}{(1-p)\,q} \; .$$

Thus, the distribution (d_0 , d_1 , d_2 , ...) is normalizable if and only if this quotient is less than 1 , that is, if and only if $p < q$.

31. Exercise.

Use your common sense to realize that this result is quite natural. Give a heuristic explanation for the fact that the stationary absolute distribution does not exist if $p \geq q$. □

12. Proof of Stability for Finite Markov Chains

In this section we give the proof of a part of the Limit Theorem given in Section 8.

Theorem. (Stability of Finite Markov Chains)

An irreducible, aperiodic Markov chain with a finite number of states is stable, that is,

$$\lim_{n \to \infty} p_{ij}(\,n\,) \;=\; p_j(\infty) \;>\; 0 \qquad (\,\forall\;\, i\,,j\,)\,,$$

Proof.

Define $m_j(n)$ and $M_j(n)$ as the minimum and the maximum of the elements in the j-th column of the n-step transition matrix:

$$m_j(n) = \min_i p_{ij}(n) \qquad \text{and} \qquad M_j(n) = \max_i p_{ij}(n).$$

First we prove that $m_j(1), m_j(2), \ldots$ is an increasing, $M_j(1), M_j(2), \ldots$ is a decreasing sequence. In order to prove this, by the total probability formula we write

$$p_{ij}(n+1) = \sum_k p_{ik} \cdot p_{kj}(n).$$

Since $p_{kj}(n) \leq M_j(n)$, we can overestimate the second factor of each term by $M_j(n)$, and we get that

$$p_{ij}(n+1) = \sum_k p_{ik} \cdot p_{kj}(n) \leq \sum_k p_{ik} \cdot M_j(n) = \left(\sum_k p_{ik}\right) \cdot M_j(n) = M_j(n).$$

Since $p_{ij}(n+1) \leq M_j(n)$ holds for all i, we get that $M_j(n+1) \leq M_j(n)$. In a similar fashion it can be shown that $m_j(n+1) \geq m_j(n)$.

Now we know that

$$\lim_{n \to \infty} m_j(n) \qquad \text{and} \qquad \lim_{n \to \infty} M_j(n)$$

exist $(j = 1, 2, \ldots)$. In order to prove the stability of the transition matrix of the Markov chain, it is enough to show that

$$\lim_{n \to \infty} \left(M_j(n) - m_j(n) \right) = 0.$$

Since $\lim_{t \to \infty} \left(M_j(n) - m_j(n) \right)$ exists, it is enough to show that

$$\lim_{n \to \infty} \left(M_j(n \cdot s) - m_j(n \cdot s) \right) = 0$$

for some s.

Since the chain is assumed to be irreducible and aperiodic, there exists such an s_0 that for all $s \geq s_0$ all elements of the s-step transition matrix are positive. (See problem entitled "All Elements of $P(n)$ Are Positive".) Let us choose and fix such an s. By the total probability formula we write for any x and y:

$$p_{xj}(s+n) = \sum_k p_{xk}(s) \cdot p_{kj}(n),$$

$$p_{yj}(s+n) = \sum_k p_{yk}(s) \cdot p_{kj}(n).$$

Taking the difference we get that

$$p_{xj}(s+n) - p_{yj}(s+n) = \sum_k \Big(p_{xk}(s) - p_{yk}(s) \Big) \cdot p_{kj}(n).$$

Let us divide the terms in this sum into two groups according to the sign of the factor $p_{xk}(s) - p_{yk}(s)$. The first sum contains those terms for which $p_{xk}(s) - p_{yk}(s) > 0$, the second contains those for which $p_{xk}(s) - p_{yk}(s) < 0$:

$$p_{xj}(s+n) - p_{yj}(s+n) =$$

$$= \sum_k^{(+)} \Big(p_{xk}(s) - p_{yk}(s) \Big) \cdot p_{kj}(n) + \sum_k^{(-)} \Big(p_{xk}(s) - p_{yk}(s) \Big) \cdot p_{kj}(n).$$

Now, using that $m_j(n) \le p_{kj}(n) \le M_j(n)$ and taking into account the signs of the factors $p_{xk}(s) - p_{yk}(s)$, we get that

$$\sum_k^{(+)} \Big(p_{xk}(s) - p_{yk}(s) \Big) \cdot p_{kj}(n) + \sum_k^{(-)} \Big(p_{xk}(s) - p_{yk}(s) \Big) \cdot p_{kj}(n) \le$$

$$\le \sum_k^{(+)} \Big(p_{xk}(s) - p_{yk}(s) \Big) \cdot M_j(n) + \sum_k^{(-)} \Big(p_{xk}(s) - p_{yk}(s) \Big) \cdot m_j(n) =$$

$$= \left(\sum_k^{(+)} \Big(p_{xk}(s) - p_{yk}(s) \Big) \right) \cdot M_j(n) + \left(\sum_k^{(-)} \Big(p_{xk}(s) - p_{yk}(s) \Big) \right) \cdot m_j(n).$$

Since

$$\left(\sum_k^{(+)} \Big(p_{xk}(s) - p_{yk}(s) \Big) \right) + \left(\sum_k^{(-)} \Big(p_{xk}(s) - p_{yk}(s) \Big) \right) =$$

$$= \sum_k \Big(p_{xk}(s) - p_{yk}(s) \Big) = \sum_k p_{xk}(s) - \sum_k p_{yk}(s) = 1 - 1 = 0,$$

we get that

$$\sum_k^{(-)} \Big(p_{xk}(s) - p_{yk}(s) \Big) = -\sum_k^{(+)} \Big(p_{xk}(s) - p_{yk}(s) \Big).$$

Thus,

$$\left(\sum_k^{(+)} \left(p_{xk}(s) - p_{yk}(s) \right) \right) \cdot M_j(n) + \left(\sum_k^{(-)} \left(p_{xk}(s) - p_{yk}(s) \right) \right) \cdot m_j(n) =$$

$$= \left(\sum_k^{(+)} \left(p_{xk}(s) - p_{yk}(s) \right) \right) \cdot \left(M_j(n) - m_j(n) \right).$$

Let us define h by

$$h = \max_{x,y} \left(\sum_k^{(+)} \left(p_{xk}(s) - p_{yk}(s) \right) \right).$$

Since $p_{xk}(s)$ and $p_{yk}(s)$ are positive, we have for all x and y that

$$\sum_k^{(+)} \left(p_{xk}(s) - p_{yk}(s) \right) = \sum_k^{(+)} p_{xk}(s) - \sum_k^{(+)} p_{yk}(s) < \sum_k p_{xk}(s) = 1.$$

Since there is a finite number of x, y pairs, this guarantees that

$$h < 1.$$

Thus, we get that

$$p_{xj}(s+n) - p_{yj}(s+n) \le$$

$$\le \left(\sum_k^{(+)} \left(p_{xk}(s) - p_{yk}(s) \right) \right) \cdot \left(M_j(n) - m_j(n) \right) \le$$

$$\le h \cdot \left(M_j(n) - m_j(n) \right).$$

Since the inequality

$$p_{xj}(s+n) - p_{yj}(s+n) \le h \cdot \left(M_j(n) - m_j(n) \right)$$

holds for all x and y, and the right side of the inequality does not depend on x and y, we get that

$$M_j(s+n) - m_j(s+n) \le h \cdot \left(M_j(n) - m_j(n) \right)$$

holds for all n. Taking $n = s, n = 2s, n = 3s, \ldots$, we get that

$$\left(M_j(2s) - m_j(2s) \right) \le h \cdot \left(M_j(s) - m_j(s) \right),$$

$$\Big(M_j (3s) - m_j (3s) \Big) \le h \cdot \Big(M_j (2s) - m_j (2s) \Big) ,$$

$$\Big(M_j (4s) - m_j (4s) \Big) \le h \cdot \Big(M_j (3s) - m_j (3s) \Big) ,$$

$$\vdots$$

Since $h < 1$, this implies that

$$\lim_{n \to \infty} \Big(M_j (n \cdot s) - m_j (n \cdot s) \Big) = 0 . \qquad \square$$

13. PROBLEMS

1. PHYSICAL INTERPRETATIONS

a) Using a die, give a physical interpretation of the Markov chain which has three states, its initial distribution is $(0 , \frac{1}{2} , \frac{1}{2})$, and its transition matrix is

$$P = \begin{bmatrix} \frac{1}{6} & \frac{1}{6} & \frac{4}{6} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} .$$

b) Give a simple interpretation and a good name for a Markov chain with transition matrix

$$\begin{bmatrix} 0 & q & 0 & 0 & p \\ p & 0 & q & 0 & 0 \\ 0 & p & 0 & q & 0 \\ 0 & 0 & p & 0 & q \\ q & 0 & 0 & p & 0 \end{bmatrix}$$

2. CHAPMAN-KOLMOGOROV EQUATION

The equation $P(m+n) = P(m) \cdot P(n)$ is obvious for a homogeneous Markov Chain. Try to generalize it for a non-homogeneous Markov chain. The generalized formula is called **Chapman-Kolmogorov equation**.

3. POLYA'S URN MODEL/1

4. GENERAL URN MODEL/2

There is a red and a white ball in a box. Pick one ball from the box, observe its color, put it back with one more balls of the same color. Repeat the action again and again. Let the random variables ξ_n and η_n be defined by

ξ_n = the number of red balls in the box after the n-th action,
$\eta_n = 1$ if the ball at the n-th drawing is red, 0 otherwise.

a) Show that the sequence $\xi_1, \xi_2, \ldots$ is a non-homogeneous Markov chain, and find the transition probabilities and the absolute distributions.

b) Show that the sequence $\eta_1, \eta_2, \ldots$ is not a Markov chain.

5. POLYA'S URN MODEL/2

a) There are A red and B white balls in a box. Pick one ball from the box, observe its color, put it back with C more balls of the same color. (C is an arbitrary positive integer.) Repeat the action again and again. Let the states of the process be the pairs of numbers consisting of the number of red and white balls, respectively. Show that this is a Markov chain, and find the one-step transition probabilities.
%%%

6. POLYA'S URN MODEL/2 – COMPUTER SIMULATION

There are A red and B white balls in a box. Pick one ball from the box, observe its color, put it back with C more balls of the same color. (C is an arbitrary positive integer.) Suggested values for A, B and C are:

$$A = 1, B = 1, C = 1;$$
$$A = 2, B = 1, C = 1;$$
$$A = 3, B = 2, C = 1;$$
$$A = 1, B = 1, C = 3;$$
$$A = 1, B = 2, C = 3.$$

For a given A, B, C triple, draw again and again, and observe the relative frequency of the times you pick a red ball. Performing a large number of experiments you can feel (what can be proven!) that the random sequence of the relative frequencies converges to a limit. However, as you can realize from your experiments, even for a given A, B, C triple, this limit value is not a constant number, but it is also a random number. Setting up a histogram for the distribution of the limit value check that

a) its distribution is the uniform distribution, when $A = B = C$;

b) its distribution has a shape of a symmetrical hill, when $A = B > C$;

c) its distribution has a shape of an asymmetrical hill, when $A \neq B > C$;

d) its distribution has a shape of a symmetrical valley, when $A = B < C$;

e) its distribution has a shape of an asymmetrical valley, when $A \neq B < C$.

f) What kind of shape du you expect when C is between A and B ?

f) The arising limit distributions constitute the family of beta-distributions. The beta-distribution with parameters $a > 0$ and $b > 0$ is defined by its density function of the form

$$\text{constant} \cdot x^{a-1} (1-x)^{b-1} ,$$

where the constant can be expressed in terms of a and b . Make drawings of these density densities for different parameter values. The parameters a and b should be expressed in terms of A , B and C by $a = \frac{A}{C}$, $b = \frac{B}{C}$. Compare the histogram you have got by the simulation to the graph of the theoretical density function.

7. POLYA'S URN MODEL/3

There are A red and B white balls in a box. Pick one ball from the box, observe its color, put it back with C more balls of the same color, and D balls of the other color. (C and D are arbitrary integers. If they are negative, then the number of balls may decrease to 0 , too.) Repeat the action again and again. Let the states of the process be the pairs of numbers consisting of the number of red and white balls, respectively. Show that this is a Markov chain, and find the one-step transition probabilities.

8. PASSIVE DETECTOR

The following experiment is a rough model how a detector may work.

Use a false coin to generate a random infinite H-T sequence. This sequence represents the sequence of signals to be detected. Our instrument is supposed to generate a 1 for each H , and a 0 for each T . However, the instrument is a little bit 'passive' which means that when a 1 is generated for an H , then the next 3 numbers generated by the instrument are all 0 -s, whatever the results of the next 3 tosses are.

a) Decide whether the generated 0-1 sequence is a Markov chain or not?

b) If it is, then find the one-step transition matrix.

9. SUPER-ACTIVE DETECTOR

Now assume that, on the contrary, our instrument is 'super-active', which means that when a 1 is generated for an H , then the next 3 numbers generated by the instrument are all 1 -s, whatever the results of the next 3 tosses are. Answer the same questions as before.

10. LARGEST NUMBER TOSSED

Toss a die, and for each n let the random variable ξ_n be equal to the largest number thrown until the n-th toss.

a) Prove that the sequence $\xi_1, \xi_2, \ldots$ is a homogeneous Markov chain.

b) Find an explicit formula for the n-step transition matrix. (Hint: Start with $n = 1, 2, 3$ and then take an arbitrary n.)

c) Find an explicit formula for the n-th absolute distribution.

d) Find the limit of the n-step transition matrix and the n-th absolute distribution.

e) Find out a heuristic explanation for the result obtained.

11. POLYA'S URN MODEL – n-STEP TRANSITION PROBABILITIES

There are A red and B white balls in a box. Pick one ball from the box, observe its color, put it back with C more balls of the same color ($C > 0$). Repeat the action again and again. Let the state of the process be the pair of numbers consisting of the number of red and white balls, respectively. Find an explicit formula for the n-step transition probabilities.

12. BANKRUPTCY (or: RANDOM WALK WITH ABSORBING ENDPOINTS) / 1

Assume that you have Z forints, and your opponent has $T - Z$ forints (thus, the total you two have is T forints.) Toss a coin. Assume that at each toss the probability of heads is p, the probability of tails is q. For each head you get 1 forint from your opponent, for each tail you pay 1 forint to your opponent until you or your opponent becomes bankrupt. Analyze the event "you become bankrupt".

a) Denote the probability of this event by $B_T(Z)$. Show that $B_T(Z)$ satisfies the equation

$$B_T(Z) = p \cdot B_T(Z+1) + q \cdot B_T(Z-1).$$

b) Show that the general solution to this equation for $p = 0.5$ is

$$B_T(Z) = A + B \cdot Z.$$

c) Show that the general solution to this equation for $p \neq 0.5$ is

$$B_T(Z) = A + B \cdot \left(\frac{q}{p}\right)^Z.$$

d) Be convinced that we are interested in the particular solution for which

$$B_T(0) = 1, \quad B_T(T) = 0.$$

e) Prove that for $p = 0.5$ the particular solution is

$$B_T(Z) = 1 - \frac{Z}{T}.$$

f) Prove that for $p \neq 0.5$ the particular solution is

$$B_T(Z) = \frac{\left(\frac{q}{p}\right)^T - \left(\frac{q}{p}\right)^Z}{\left(\frac{q}{p}\right)^T - 1}.$$

g) Find the probability that your opponent becomes bankrupt.

h) Find the probability that none of you becomes bankrupt.

i) Taking the limit when $T \to \infty$, analyze the problem. What is the real life meaning of the limit value $B_\infty(Z)$?

j) Realize that in the above calculations you did not use the finite dimensional distributions to construct a measure P on the infinite dimensional space. What made it possible to determine the measures of some non-cylindrical subsets?

k) To develop your real life experiences – find the numerical value of the above probabilities for some values of T, Z and p.

13. LENGTH OF THE GAME (continued)

Denote the expected value of the length of the game in the previous problem by $L_T(Z)$.

a) Realize that the length of the game is a random variable defined on the set of all 0-1-sequences.

b) Prove that $L_T(Z)$ satisfies the equation

$$L_T(Z) = 1 + p \cdot L_T(Z+1) + q \cdot L_T(Z-1).$$

c) Show that the general solution to this equation for $p = 0.5$ is

$$L_T(Z) = A + B \cdot Z - Z^2.$$

d) Show that the general solution to this equation for $p \neq 0.5$ is

$$L_T(Z) = \frac{Z}{q-p} \cdot A + B \cdot \left(\frac{q}{p}\right)^Z.$$

e) Be convinced that we are interested in the particular solution for which

$$L_T(0) = 0, \quad L_T(T) = 0.$$

f) Prove that for $p = 0.5$ the particular solution is

$$L_T(Z) = Z \cdot (T - Z).$$

g) Prove that for $p \neq 0.5$ the particular solution is

$$L_T(Z) = \frac{Z}{q-p} - \frac{T}{q-p} \cdot \frac{1-\left(\frac{q}{p}\right)^Z}{1-\left(\frac{q}{p}\right)^T}.$$

h) Taking the limit when $T \to \infty$, analyze the problem. What is the real life meaning of the limit value $L_\infty(Z)$?

i) Realize that in the above calculations the expected value of the length of the game was calculated without using its distribution.

14. BANKRUPTCY (SOLUTION BY RECURSIVE FORMULAS)

Prove that

$$B_T(Z) = p \cdot B_{T-Z}(1) \cdot B_T(Z) +$$

$$+ q \cdot \left(B_Z(Z-1) + \left(1 - B_Z(Z-1)\right) \cdot B_T(Z) \right),$$

and realize that from this equation $B_T(Z)$ can be expressed by other $B_T(Z)$-s with smaller T or Z values, which offers a recursive formula for $B_T(Z)$.

15. BANKRUPTCY FOR $T = 4$ AND $Z = 2$ (TABLE TENNIS GAME FROM 20:20)

Show that $B_4(2)$ can be expressed as the sum of a geometrical series:

$$B_4(2) = p^2 + 2pq + (2pq)^2 + \ldots = \frac{p^2}{1-2pq}.$$

16. BANKRUPTCY (SOLUTION BY A SYSTEM OF LINEAR EQUATIONS)

Let's take $T = 6$, and let $B_6(i)$ be denoted by x_i ($i = 0, \ldots, 6$). Clearly $x_0 = 1$, $x_6 = 0$. Set up a system of linear equations for x_i ($i = 1, \ldots, 5$).

17. RANDOM WALK WITH AN ABSORBING ENDPOINT

Consider the random walk on the set of all non-negative integers, when 0 is an absorbing state. From any state different from 0, the probability of a right jump is p, and the probability of a left jump is q $(p+q=1)$. Let $B_\infty(z)$ denote the probability that the state 0 will be reached ever in the future on condition that we start from state z. Show that

a) $B_\infty(2) = \left(B_\infty(1)\right)^z$ for $z = 1, 2, \ldots$;

b) $B_\infty(1) = p + q \cdot B_\infty(2)$.

c) Prove that if $q < p$, then $B_\infty(1) = \frac{q}{p}$, and $B_\infty(1) = 1$ otherwise.

d) For $z = 1, 2, \ldots$ conclude that if $q < p$, then $B_\infty(z) = \left(\frac{q}{p}\right)^z$, and $B_\infty(z) = 1$ otherwise.

18. RANDOM WALK, RECURRENT OT NOT?

a) Consider the random walk on the set of integer numbers, when the probability of a left or right step is p or q, respectively. Find f_{ij} , and decide whether the process is recurrent or not.

b) Consider the random walk on the set of non-negative integer numbers. Find f_{ij} , and decide whether the process is recurrent or not.

19. BUG ON THE REAL LINE

A bug is moving on the real line: it makes jumps one unit to the right or to the left with probability p or q, respectively, independently of the previous jumps $(p+q=1)$. Assume that at time instant 0 the bug was at the origin.

a) What is the probability that the bug reaches the point x before it reaches the point y ? (x and y are integers; $x < 0 < y$.)

b) What is the probability that the bug will ever reach the point x ?

(Hint: discover the connection of this problem to the bankruptcy problem above.)

20. ABSORPTION PROBABILITIES

Consider a Markov-chain in which absorbing states exist. Order the states so that first you list the absorbing ones, then the others.

a) Check that that the one- step transition matrix has the following form:

$$P = \begin{bmatrix} E & 0 \\ R & Q \end{bmatrix},$$

where E is a unit matrix, and 0 is a zero matrix.

b) Check that that the n - step transition matrix has the following form:

$$P(n) = \begin{bmatrix} E & 0 \\ R(n) & Q(n) \end{bmatrix} .$$

c) Prove the relation $Q(n+1) = Q^n$.

d) Prove the relations $R(n+1) = R(n) + Q(n) \cdot R$ and $R(n+1) = R + Q \cdot R(n)$.

e) Prove the relation $R(n) = R + Q \cdot R + Q^2 \cdot R + \ldots + Q^n \cdot R$.

f) Show that the sequence of matrices $R(n)$ $(n = 1, 2, \ldots)$ constitutes an increasing sequence, meaning that each entry of the matrix $R(n+1)$ is greater than or equal to the corresponding entry of $R(n)$.

g) Setting up an equation for $R(\infty)$ show that

$$R(\infty) = (E - Q)^{-1} \cdot R .$$

f) Figure out that the element in the i -th row and j -th column of the matrix $R(\infty)$ is the probability that if the process starts at the non-absorbing state i , then sooner or later it will get stuck in the absorbing state j .

21. INDEPENDENT TRANSITIONS FOR THE COORDINATES

Consider two Markov chains. Imagine that the first coordinate of a new process changes according to the first transition rule, while the second coordinate changes according to the second transition rule. Show that the resulting process is also a Markov chain, and find its transition matrix.

22. WHICH RUN WINS?

Generate a random 0-1-sequence by a false coin so that the probability of 1 and 0 at each toss are p and q , respectively. Analyze the event "a run of eight successive 1-s occurs before a run of five successive 0-s" .

a) Let u denote the probability of this event on condition that the first element of the sequence is 1 , and let v denote the probability of this event on condition that the first element of the sequence is 0 . Prove that the following system of equations holds for u and v :

$$u = p^7 + (1 - p^7) \cdot v$$

$$v = (1 - q^4) \cdot u ,$$

Hint: use the total probability formula, and the summation formula for a finite geometrical series.

b) Solve the system to calculate the conditional probabilities.

c) Combining the conditional probabilities find the (unconditional) probability of the above event.

d) Find now the probability that "five 0 -s occur successively before eight 1 -s occur successively".

e) How much is the probability that "five 0 -s or eight 1 -s never occur".

f) Realize that in the above calculations you did not use the finite dimensional distributions to construct a measure P on the infinite dimensional space.. What made it possible to determine the measures of some non-cylindrical subsets?

g) Replace the numbers "eight" and "five" in the definition of the above event by arbitrary positive integers M and N , and – to develop your real life experiences – find the numerical value of the corresponding probabilities for some values of p , M and N .

23. LENGTH OF EACH RUN

a) Consider a box which contains A red and B white balls. Pick balls with replacement, and each time observe its color and the length of the run of that color. If, for example, you pick

R, R, W, W, W, W, W, R , W, R ,

then the result of your observation will be:

R1, R2, W1, W2, W3, W4, W5, R1 , W1, R1 .

Identify the states, find the one-step transition probabilities and determine the stationary absolute distribution.

b) Generalize the problem when there are balls with more colors.

24. LIMIT PROBABILITIES FOR MARTINGALES

The states of a homogeneous Markov chain $\xi_0 , \xi_1 , \xi_2 , \ldots$ are the integers between a and b , a and b included. Assume that the transition matrix satisfies the so called "martingale property":

$$\sum_{j=a}^{b} j \cdot p_{ij} = i \quad \text{for all } i .$$

a) Prove that $p_{aa} = p_{bb} = 1$.

b) Assume that, for all i between a and b , the probability of getting from i into a and the probability of getting from i into b are positive. Prove that with the exception of a finite number of steps $\xi_n = a$ or $\xi_n = b$ with probability 1 .

c) Prove that for all n

$$\sum_{j=a}^{b} j \cdot p_{ij}(n) = i \quad \text{for all } i .$$

d) Prove that

$$\lim_{n \to \infty} p_{ia}(n) = \tfrac{b-i}{b-a} ,$$

$$\lim_{n \to \infty} p_{ij}(n) = 0 \quad (a < j < b) ,$$

$$\lim_{n \to \infty} p_{ib}(n) = \tfrac{i-a}{b-a} .$$

25. GENETICS MODEL: 'DRAW WITH REPLACEMENT'

The contents of a sequence of boxes are defined at random as follows. The 0 -th box contains 3 red and 3 white balls. Assume that the contents of the n -th box has already been defined. Then make 6 drawings with replacement from the n -th box, and observe the color of each ball, and fill the $(n+1)$ -th box according to the colors you have got. Let ξ_{n+1} mean the number of red balls in the $(n+1)$ -th box.

a) Show that the sequence $\xi_0 , \xi_1 , \xi_2 , \ldots$ is a Markov chain.

b) Find the transition matrix.

c) Prove that the martingale property (see problem entitled 'Limit Probabilities for Martingales') holds, and analyze the real life meaning of the limit relations which follow from the martingale property.

26. GENETICS MODEL: 'DOUBLE THE CONTENTS AND DRAW WITHOUT REPLACEMENT'

The contents of a sequence of boxes are defined at random as follows. The 0 -th box contains 2 red and 3 white balls. Assume that the contents of the n -th box has already been defined. In order to define the contents of the $(n+1)$ -th box, double the contents of the n -th box, and choose 5 out of the 10 , without replacement. Let ξ_{n+1} mean the number of red balls in the $(n+1)$ -th box.

a) Show that the sequence $\xi_0 , \xi_1 , \xi_2 , \ldots$ is a Markov chain.

b) Find the transition matrix.

c) Prove that the martingale property (see problem entitled 'Limit Probabilities for Martingales') holds, and analyze the real life meaning of the limit relations which follow from the martingale property.

27. GENETICS MODEL: 'DRAW WITHOUT REPLACEMENT AND DOUBLE THE CONTENTS'

The contents of a sequence of boxes are defined at random as follows. The 0 -th box contains 4 red and 6 white balls. Assume that the contents of the n th box has already been defined. In order to define the contents of the $(n+1)$ -th box, choose 5 out of the 10 from the n -th box , without replacement, and then double what you got, and put them into the $(n+1)$ -th box. Let ξ_{n+1} mean the number of red balls in the $(n+1)$ -th box.

a) Show that the sequence $\xi_0 , \xi_1 , \xi_2 , \ldots$ is a Markov chain.

b) Find the transition matrix.

c) Prove that the martingale property (see problem entitled 'Limit Probabilities for Martingales') holds, and analyze the real life meaning of the limit relations which follow from the martingale property.

28. GETTING STUCK

The states of a homogeneous Markov chain are $1 , 2 , \ldots , 6$. The transition matrix is

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 & 0 \\ \frac{1}{16} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{16} & \frac{1}{8} & 0 \\ 0 & 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

a) Show that the process, with probability 1 , gets stuck either in state 1 or state 5 .

b) Find the probability that, starting from state 3 , the process gets stuck in state 1 .

c) Find the probability that, starting from state i , the process gets stuck in state 1 .

29. STATIONARY ABSOLUTE DISTRIBUTION

Find the stationary absolute distribution for the transition matrix:

$$P = \begin{bmatrix} .75 & .25 & 0 \\ .25 & 0 & .75 \\ .25 & .25 & .50 \end{bmatrix}.$$

30. STATIONARY?

For a homogeneous Markov chain $\xi_0 , \xi_1 , \xi_2 , \ldots$ the distribution of ξ_1 is the same as the distribution of ξ_2 .

a) Is the chain necessarily stationary?

b) What about the chain $\xi_1 , \xi_2 , \xi_3 , \ldots$?

31. STATIONARY BUT NON-MARKOV

Consider the stationary random walk $\xi_0 , \xi_1 , \xi_2 , \ldots$ on the set $1 , 2 , 3$ with reflecting barriers at 1 and 3, for which the initial distribution and the one-step transition matrix are

$$\begin{bmatrix} 0.25 & 0.50 & 0.25 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 1 & 0 \\ 0.5 & 0 & 0.5 \\ 0 & 1 & 0 \end{bmatrix},$$

respectively. Define the process $\eta_0 , \eta_1 , \eta_2 , \ldots$ by $\eta_n = 1$ if $\xi_n = 1$, and $\eta_n = 0$ otherwise.

a) Show that the process $\eta_0 , \eta_1 , \eta_2 , \ldots$ is not a Markov chain, but it is a stationary process.

b) Will the process $\eta_0 , \eta_1 , \eta_2 , \ldots$ remain stationary if the process $\xi_0 , \xi_1 , \xi_2 , \ldots$ is a homogeneous Markov chain with the given transition matrix, but with another initial distribution?

32. LENGTH OF RUN (continued, see Section 1 , Example entitled 'Length of Run')

a) In order to determine the stationary absolute distribution, find the limit of the n-step transition matrix.

b) Find out a heuristic explanation for the result obtained.

33. DIFFUSION MODEL OF EHRENFEST (MOLECULE IS CHOSEN) (continued, see Section 5 , Example entitled 'Diffusion Model of Ehrenfest')

Pretend that the stationary absolute distribution is not known for this model, and use the ideas of the Example entitled 'Stay or Jump Only onto Adjacent States' in Section 5. to find it.

34. DIFFUSION BASED ON COIN TOSSING (HALF OF THE CONTAINER IS CHOSEN)

There are N molecules in a container, which is theoretically divided into two parts, say, left and right. Each time instant we toss a coin, and according to the result, one molecule is put from the left into the right or from the right into the left side. If the part from which the molecule should be taken is empty, then nothing happens. Consider the process defined by ξ_t $=$ the number of molecules in the left side at time instant t $(t = 0, 1, 2, \dots)$.

a) Show that the process is a homogeneous Markov chain.

b) Find the transition probabilities.

c) Assume that initially the molecules are distributed somehow in the container. After a long time you look at the system, and count the number of molecules in the left side. What can you say about the distribution of this random variable?

35. DIFFUSION MODEL OF BERNOULLI AND LAPLACE (MOLECULES COLORED, CHOSEN IN BOTH PARTS)

There are $2N$ molecules in a container, which is theoretically divided into two parts, say, left and right. Half of the molecules is red, the others are white. Half of the molecules is in the left side, the others are in the right side. Each time instant one molecule is chosen at random from both sides (all have the same probability, and the two choices are independent), and then they are exchanged. Consider the process defined by ξ_t $=$ the number of red molecules in the left side at time instant t $(t = 0, 1, 2, \dots)$.

a) Show that the process is a homogeneous Markov chain.

b) Find the transition probabilities.

c) Show that if initially the molecules are placed into the container at random so that the distribution of the number of red molecules in the left side is hypergeometrical with parameter $(2N, N, N)$, then the chain is stationary.

d) Pretend that the stationary absolute distribution were not given in part c) above, and use the ideas of the Example entitled 'Stay or Jump Only onto Adjacent States' in Section 5. to find it.

e) Assume now that initially all the red molecules are in the left side. After a long time you look at the system, and count the number of red molecules in the left side. What can you say about the distribution of this random variable?

36. DOUBLE STOCHASTIC MATRICES

Assume that a Markov chain has only a finite number of states $1, 2, \dots, N$ and its transition matrix is double stochastic, which means that not only the rows, but the columns as well are normalized, that is, for every state j

$$\sum_{i=0}^{N} p_{ij} = 1 .$$

Prove that the uniform distribution is a stationary absolute distribution.

37. IMMEDIATE RETURN IS PROHIBITED

Consider the points

$$\begin{matrix} (3,1) & (3,2) & (3,3) \\ (2,1) & (2,2) & (2,3) \\ (1,1) & (1,2) & (1,3) \end{matrix}$$

on the plane. Put a bug on $(1,2)$ or $(2,2)$ with probability p or q, respectively $(p+q=1)$. At the first step the bug steps onto one of the adjacent points, and then it walks on the given nine points so that 'Immediate Return Is Prohibited', which means that it always steps onto one of the adjacent points with equal probabilities except the one from which it came.

a) Considering the sequence of the positions of the bug do we get a Markov chain?

b) There are 24 possible moves for the bug. Show that considering the sequence of the moves of the bug we do get a Markov chain.

c) Prove that the transition matrix of this Markov chain is a double stochastic matrix. (For this notion, see problem entitled 'Double Stochastic Matrices'.)

d) Prove that the transition matrix is irreducible with period 2, thus, considering every second move, we get an aperiodic chain, for which the uniform distribution on the 24 possible moves is the limit distribution.

e) Using this result try to describe how the position of the bug will be distributed after a long period of time.

38. LUNCH MONEY (continued, see Section 6 for Example entitled 'Lunch Money')

What is the distribution of the amount of money John has in his schoolbag at the end of the school year.

39. INFINITELY MANY RETURNS

Prove that the number of returns to a recurrent state is ∞ with probability 1.

40. IRREDUCIBLE AND APERIODIC

Prove that if the elements in and above and below the main diagonal of the transition matrix are positive, then the chain is irreducible and aperiodic.

41. $p_{ii}(n) > 0$ IF n IS A LARGE MULTIPLE OF THE PERIOD

Let d_i the period of a state i. Prove that there exists such an n_0 that $p_{ii}(n) > 0$ whenever $n \geq n_0$ and n is a multiple of d_i. Use the following fact from number theory:

If $D \subseteq \{1, 2, 3, \dots\}$ and d is the greatest common divisor of D, then there exists a finite number of elements $d(1), \dots, d(r)$ in D that, whenever $n \geq n_0$ and n is a multiple of d, then n is expressible as linear combination of $d(1), \dots, d(r)$ with non-negative coefficients:

$$n = \sum_{k=0}^{r} c(k)\, d(k) \quad \Big(c(k) \geq 0 \,, \ (k = 1, \dots, r) \Big) .$$

42. ALL ELEMENTS OF $P(n)$ ARE POSITIVE

Prove that for an irreducible, aperiodic, finite Markov chain there exists such an n_0 that, for $n \geq n_0$, all elements of the n-step transition matrix are positive: $p_{ij}(n) > 0$.

43. AVERAGE RETURN TIME = RECIPROCAL OF LIMIT PROBABILITY

Consider a recurrent state i of a homogeneous Markov chain. Recall that on condition that at time instant 0 the process is in this state, $p_{ii}(n)$ denotes the probability that n units of time later the process is again in this state. We assume that the limit $p_{ii}(\infty) = \lim_{n \to \infty} p_{ii}(n)$ exists and it is positive. Recall that $f_{ii}(k)$ denotes the probability that k is the first time instant when the process returns to this state. Let $q_{ii}(n)$ denote the probability that the process does not return to i during the time interval $[1, n]$. Recall that the expected value of the steps (or length of time) needed to return into state i is denoted by m_i:

$$m_i = \sum_{k=1}^{\infty} k \cdot f_{ii}(k) .$$

At the end of the following problems it will turn out that $m_i = \frac{1}{p_{ii}(\infty)}$.

a) Show that

$$q_{ii}(n) = \sum_{k=n+1}^{\infty} f_{ii}(k) .$$

b) Show that

$$m_i = \sum_{n=0}^{\infty} q_{ii}(n) .$$

c) Show that

$$\sum_{k=0}^{n} p_{ii}(n-k) \cdot q_{ii}(k) = 1 \ .$$

d) Lebesgue's Dominated Convergence Theorem states that, if $\lim_{n \to \infty} c(n,k)$ exists for all k, and $|c(n,k)| \leq C(k)$ for all k and n, where $\sum_k C(k) < \infty$, then

$$\lim_{n \to \infty} \left(\sum_k c(n,k) \right) = \sum_k \left(\lim_{n \to \infty} c(n,k) \right) .$$

Using Lebesgue's Dominated Convergence Theorem and the formula shown in part c) , prove that

$$\lim_{n \to \infty} \sum_{k=0}^{n} p_{ii}(n-k) \cdot q_{ii}(k) = \left(\lim_{n \to \infty} p_{ii}(n) \right) \cdot \left(\sum_{k=0}^{\infty} q_{ii}(k) \right) ,$$

from which it follows that

$$p_{ii}(\infty) \cdot \left(\sum_{k=0}^{\infty} q_{ii}(k) \right) = 1 , \quad \text{that is ,} \quad m_i = \frac{1}{p_{ii}(\infty)} \ .$$

44. AVERAGE RETURN TIMES

The transition matrix of a homogeneous Markov chain is

$$\begin{bmatrix} \frac{1}{6} & \frac{1}{6} & \frac{4}{6} \\ \frac{1}{6} & 0 & \frac{5}{6} \\ \frac{1}{6} & \frac{2}{6} & \frac{3}{6} \end{bmatrix}$$

What is the average length of time of a return for each state?

45. PERIOD OF A CHAIN

The period of an irreducible Markov chain $\xi_1 , \xi_2 , \xi_3 , \xi_4 \ldots$ is 6 .

a) What is the period of the chain $\xi_1 , \xi_3 , \xi_5 , \ldots$?

b) What is the period of the chain $\xi_2 , \xi_4 , \xi_6 , \ldots$?

c) What is the period of the chain $\xi_3 , \xi_6 , \xi_9 , \ldots$?

d) What is the period of the chain $\xi_6 , \xi_{12} , \xi_{18} , \ldots$?

46. CLASSES OF STATES

Two different states i and j are said to **communicate** with each other if both f_{ij} and f_{ji} are positive. Any state is said to **communicate** with itself.

a) Show that communication is an equivalence relation, partitioning the phase space into equivalence classes.

b) Define Markov chains which consists of one class only.

c) Define Markov chains which consists of more than one class.

47. SUB-CLASSES OF A CLASS

The **period** δ_i of a state i is the greatest common divisor of those step-numbers n for which the return into the state i with n steps has a positive probability:

$$\delta_i = \text{g. c. d.} \{ n : p_{ii}(n) > 0 \} .$$

a) Prove that states communicating with each other have the same period. Thus, the period is a characteristic of a class of states.

b) Show that if the period of a class C is δ, then the class can be partitioned into δ cyclically ordered sub-classes $C_1 , \ldots , C_\delta$ so that, for any two states i and j belonging to the class, $p_{ij} > 0$ holds only if the sub-class of j follows the sub-class of i in the cyclic order.

48. ESSENTIAL AND UNESSENTIAL STATES

A state i is called **essential** if, for any other state j, $f_{ij} > 0$ implies that $f_{ji} > 0$. A state i is called **unessential** if it is not essential, which means that, for some state j, $f_{ij} > 0$, but $f_{ji} = 0$.

a) Show that if a state of a class is essential, then all states of that class are essential, thus essentiality is a characteristic of a class of states.

b) Show that if i is essential and j is unessential, then $p_{ij}(n) = 0$ for all n, that is $f_{ij} = 0$.

c) Show that if a state j is unessential, then $\lim_{n \to \infty} p_{ij}(n) = 0$ for any state i.

49. LIMIT PROBABILITIES OF CLASSES

a) Show that if a Markov chain consists of more than one class, and a class C is unessential, then $\lim_{n \to \infty} P(\xi_n \in C) = 0$.

b) Show that if a class C is essential, then the sequence $P(\xi_n \in C)$ $(n = 0, 1, 2, \dots)$ of probabilities is increasingly converging to a limit.

c) Show that if C_1 is a sub-class of an essential class C with period δ, then the sequence $P(\xi_{k\delta+r} \in C)$ $(k = 0, 1, 2, \dots)$ of probabilities increasingly converges to a limit for any $r = 0, 1, \dots, \delta - 1$.

50. RECURRENCE AND POSITIVITY AS CLASS PROPERTIES

a) Show that if a state of a class is recurrent, then all states of that class are recurrent, thus recurrence is a characteristic of a class of states. (Hint: use the criterion for recurrence stated in one of the above problems.)

b) Show that if a state of a class is positive, then all states of that class are positive, thus positivity is a characteristic of a class of states.

51. LENGTH OF THE QUEUE MAXIMIZED, FEW CUSTOMERS

Let the random variables ξ_0 and $\theta_1, \theta_2, \dots$ and $\eta_1, \eta_2, \dots$ independent. Assume that the common distribution of $\theta_1, \theta_2, \dots$ is the indicator distribution with parameter p, and the common distribution of $\eta_1, \eta_2, \dots$ is the indicator distribution with parameter q. ξ_0 is an arbitrary integer valued random variable. Let L be a positive number. Define the process $\xi_0, \xi_1, \xi_2, \dots$ by recursion:

$$\xi_n = \min\Big((\xi_{n-1} - \theta_n)^+ + \eta_n , L \Big) \qquad (n = 1, 2, \dots).$$

a) Prove that the process $\xi_0, \xi_1, \xi_2, \dots$ is a homogeneous Markov chain.

b) Interpret ξ_n as the length of a queue at time instant n, if $\theta_1, \theta_2, \dots$ mean the capacity of the shop to serve customers during each time interval, and $\eta_1, \eta_2, \dots$ mean how many customers arrive to a shop during each time interval, and the length of the queue cannot exceed the maximal value L : customers, arriving when the length of the queue is L, are driven away. Since the number of the arriving customers in each time interval is equal to 0 or 1, "Few Customers" is added to the title of the problem.

c) Find the transition matrix for the process $\xi_0, \xi_1, \xi_2, \dots$.

d) Find the stationary absolute distribution.

e) How much is the probability that at a large time instant the queue is full, that is, $\xi_n = L$?

52. LENGTH OF THE QUEUE MAXIMIZED, MANY CUSTOMERS

We modify the previous problem. We assume again that the common distribution of $\theta_1 , \theta_2 , \ldots$ is the indicator distribution with parameter p , but the common distribution of the random variables $\eta_1 , \eta_2 , \ldots$ is assumed to be an arbitrary normalized distribution $Q = (q_0 , q_1 , q_2 , \ldots)$ on $\{ 0 , 1 , 2 ,\ldots \}$. Since the number of the arriving customers in each time interval may be equal to $0 , 1 , 2 , \ldots$, "Many Customers" is added to the title of the problem.

a) Answer the questions in the previous problem when $L = 3$, and the distribution Q is the 0.5 -parametrical geometrical distribution on $\{ 0 , 1 , 2 , \ldots \}$;

b) Answer the questions in the previous problem when $L = 3$, and the distribution Q is the q -parametrical geometrical distribution on $\{ 0 , 1 , 2 , \ldots \}$.

c) Find an algorithm to answer the questions in the previous problem for any finite value L and distribution Q . Hint: introduce the notation

$$p_{i,\geq j} = \sum_{k:\, k \geq j} p_{i,k} \, .$$

53. LEFT JUMP ONLY ONTO THE NEIGHBOR

Let the states of an irreducible Markov chain be the non-negative integers (infinite case) or the integers $0 , 1 , 2 , \ldots , L$ (finite case). Assume that left jumps are possible only onto adjacent states, that is, $p_{i,j} = 0$, whenever $j < i - 1$.

a) Realize that the previous problem belongs to the finite case of this problem.

b) Set up a system of equations in terms of the transition probabilities p_{ij} for the stationary absolute distribution and solve it when L equals 2 or 3 . Hint: introduce the notation

$$p_{i,\text{ not } i} = \sum_{j:\, j \neq i} p_{i,j} \, .$$

c) Find an algorithm to determine the stationary distribution for arbitrary L .

d) Modify the previous problem by omitting the upper bound L for the queue length so that the arising problem will belong to the infinite case of this problems.

e) Find an algorithm to determine the stationary distribution for the infinite case, too. Try to conjecture a criterion for the stability of the chain?

f) What is the advantage of the assumption that the common distribution of $\theta_1 , \theta_2 , \ldots$ is an indicator distribution? Why would it be difficult to find the stationary absolute distribution otherwise?

54. DOES THE FAIRY WAKE UP?

While there are Dream Viruses in the blood of the Good Fairy, she is sleeping and sleeping. Each of this kind of virus, independently from the others, dies with probability p, or splits with probability q and becomes two new viruses exactly one minute after it came into being. Assume that one virus gets into the blood of the Fairy.

a) Let $a(n)$ denote the probability that n minutes after the infection the Fairy is already awake. Set up a recursive formula for $a(n)$.

b) Using the recursive formula find the limit of $a(n)$ when $n \to \infty$. What is the meaning of the limit value $s(\infty)$?

55. STRONGER FORMS OF THE MARKOV PROPERTY

Verbalize and try to prove the following stronger forms of the Markov property:

a) If $s < t$, then

$$\Pr\left(\xi_{t+1} = j \mid \xi_s = i_s , \xi_{s+1} = i_{s+1} , \dots , \xi_{t-1} = i_{t-1} , \xi_t = i\right)$$

$$= \Pr\left(\xi_{t+1} = j \mid \xi_t = i\right);$$

b) If $t_1 < t_2 < \dots < t_{n-1} < t_n < t_{n+1}$, then

$$\Pr\left(\xi_{t_{n+1}} = i_{n+1} \mid \xi_{t_1} = i_1 , \xi_{t_2} = i_2 , \dots , \xi_{t_{n-1}} = i_{n-1} , \xi_{t_n} = i_n\right)$$

$$= \Pr\left(\xi_{t_{n+1}} = i_{n+1} \mid \xi_{t_n} = i_n\right) .$$

56. CONDITIONAL INDEPENDENCE BETWEEN THE PAST AND THE FUTURE

a) Prove that the past (represented by the time instant $t - n$) and the future (represented by the time instant $t + n$) are independent under the condition that the present (represented by the time instant t) is given, that is,

$$\Pr\left(\xi_{t-n} = i_{t-n} , \xi_{t+n} = i_{t+n} \mid \xi_t = i_t\right)$$

$$= \Pr\left(\xi_{t-n} = i_{t-n} \mid \xi_t = i_t\right) \cdot \Pr\left(\xi_{t+n} = i_{t+n} \mid \xi_t = i_t\right)$$

b) Try to generalize this conditional independence for the case when the past, the present and the future are represented by more than one time instant.

Chapter 2

Stationary Processes

1. Stationarity

A process η_t ($t = 0, 1, 2, \ldots$) is called **stationary in the strong sense** (or **strictly stationary**) if, the distribution of η_t does not depend on t, and moreover, for all n and $t_1, t_2, \ldots t_n$, the distribution of ($\eta_{t_1}, \eta_{t_2}, \ldots, \eta_{t_n}$) depends on $t_1, t_2, \ldots t_n$ only through the differences between $t_1, t_2, \ldots t_n$, that is, the distribution of ($\eta_{t_1+\theta}, \eta_{t_2+\theta}, \ldots, \eta_{t_n+\theta}$) does not depend on θ.

Notice that not the value of the vector ($\eta_{t_1+\theta}, \eta_{t_2+\theta}, \ldots, \eta_{t_n+\theta}$) is assumed not to depend on θ, only its distribution should not depend on θ.

Stationarity implies that the distribution of η_t may not depend on t, and the distribution of (η_s, η_t) may depend on s and t only through $t - s$. It is clear that if the expected value and the covariance functions exist for a strongly stationary process, then the expected value function is constant and the covariance function depends on s and t only through the time difference $t - s$. This is why we introduce the notion of weak stationarity.

A process η_t ($t = 0, 1, 2, \ldots$) is called **stationary in the weak sense** (or **weakly stationary**) if the expected value $m(t)$ of η_t and the covariance $b(s, t)$ between η_s and η_t exist for all s and t, and $m(t)$ does not depend on t, and $b(s, t)$ depends on s and t only through $t - s$:

$$m(t) = m \qquad \text{and} \qquad b(s,t) = B(t-s),$$

where m is a constant, and $B(u)$ is a function of one variable only.

Because of the symmetry of the covariance, the function B is an even function: $B(-u) = B(u)$.

If the expected value or the covariance of a strongly stationary process does not exist, then the process is not stationary in the weak sense. However, if the expected value and the covariance of a strongly stationary process exist, then the process is stationary in the weak sense, too. It is easy to construct processes which are stationary in the weak sense but not in the strong sense.

A homogeneous Markov chain is obviously stationary in the strong sense as soon as the distribution of η_t does not depend on t. This is why a homogeneous Markov chain was called stationary in the other chapter if the distribution of η_t does not depend on t.

1. Example: WEAKLY STATIONARY BUT NOT STRONGLY

Let the random variable ψ take the values $0, 1, 2, 3$ with equal probabilities. Define a process by $\eta_t = \cos\left(\frac{\pi}{4} t + \frac{\pi}{2} \psi \right)$ ($t = 0, 1, 2, \ldots$).

a) What are the possible realizations of this process? Draw a figure to visualize them.

b) Show that the process is not stationary in the strong sense.

c) Show that the process is weakly stationary.

Solution.

a) The expected value function is

$$m(t) = M\left(\cos\left(\frac{\pi}{4}t + \frac{\pi}{2}\psi\right)\right)$$

$$= \frac{1}{4}\left(\cos\left(\frac{\pi}{4}t + \frac{\pi}{2}0\right) + \cos\left(\frac{\pi}{4}t + \frac{\pi}{2}1\right)\right.$$

$$\left. + \cos\left(\frac{\pi}{4}t + \frac{\pi}{2}2\right) + \cos\left(\frac{\pi}{4}t + \frac{\pi}{2}3\right)\right) = 0,$$

because the first and the third, and the second and fourth terms cancel each other.

The covariance function is

$$b(s,t) = M\left(\cos\left(\frac{\pi}{4}s + \frac{\pi}{2}\psi\right) \cdot \cos\left(\frac{\pi}{4}t + \frac{\pi}{2}\psi\right)\right).$$

Using the identity $\cos\alpha \cdot \cos\beta = \frac{1}{2}\left(\cos(\alpha+\beta) + \cos(\alpha-\beta)\right)$, the covariance function can be written as

$$b(s,t) = \frac{1}{2} M\left(\cos\left(\frac{\pi}{4}(t+s) + \pi\psi\right) + \cos\left(\frac{\pi}{4}(t-s)\right)\right)$$

$$= \frac{1}{2} M\left(\cos\left(\frac{\pi}{4}(t+s) + \pi\psi\right)\right) + \frac{1}{2} M\left(\cos\left(\frac{\pi}{4}(t-s)\right)\right).$$

Now the first expected value is obviously 0 (do a similar calculation as we did for the expected value function). In the second, the expected value of a constant stands, which is equal to that constant. This is how we get that

$$b(s,t) = \frac{1}{2}\cos\left(\frac{\pi}{4}(t-s)\right).$$

Since the expected value function is a constant, and the covariance function depends only on $t-s$, the process turns out to be weakly stationary.

b) The following table shows how the values of ξ_0 and ξ_1 depend on the values of ψ:

ψ	0	1	2	3
ξ_0	1	0	-1	0
ξ_1	$\frac{\sqrt{2}}{2}$	$\frac{-\sqrt{2}}{2}$	$\frac{-\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$

The table clearly shows that the distributions of ξ_0 and ξ_1 are clearly different, which breaks strong stationarity.

2. Exercise.

Sketch the possible realizations of the process. □

3. Example: BOTH WEAKLY AND STRONGLY STATIONARY

Let the random variable ψ be uniformly distributed on the interval $[\,0\,,\,2\pi\,]$. Define a process by $\eta_t = \cos\Big(2\pi\,t + \psi\,)\Big)$ $\quad(\,t = 0\,,1\,,2\,,\dots\,)$.

a) Show that the process is weakly stationary.

b) Show that the process is strongly stationary, too.

Solution.

a) The expected value function is

$$m\,(\,t\,) = M\left(\cos\,(\,2\pi\,t + \psi\,)\right) = \frac{1}{2\pi}\int\limits_0^{2\pi}\cos\,(\,2\pi\,t + x\,)\,dx = 0\,.$$

The covariance function is

$$b\,(\,s\,,t\,) = M\left(\cos\,(\,2\pi\,s + \psi\,)\cdot\cos\,(\,2\pi\,t + \psi\,)\right).$$

Using th identity $\cos\alpha\cdot\cos\beta = \frac{1}{2}\Big(\cos\,(\alpha+\beta) + \cos\,(\alpha-\beta)\Big)$, the covariance function can be written as

$$b\,(\,s\,,t\,) = \frac{1}{2}\,M\left(\cos\Big(2\pi\,(s+t) + 2\psi\Big) + \cos\Big(2\pi\,(t-s)\Big)\right)$$

$$= \frac{1}{2}\,M\left(\cos\Big(2\pi\,(s+t) + 2\psi\Big)\right) + \frac{1}{2}\,M\left(\cos\Big(2\pi\,(t-s)\Big)\right)$$

Now the first expected value is obviously 0 (do a similar calculation as we did for the expected value function). In the second, the expected value of a constant stands, which is equal to that constant. This is how we get that

$$b\,(\,s\,,t\,) = \frac{1}{2}\cos\Big(2\pi\,(t-s)\Big)\,.$$

Since the expected value function is a constant, and the covariance function depends only on $t-s$, the process turns out to be weakly stationary.

b) For any fixed t, consider the random point P on the circumference of the unit circle whose polar coordinates are 1 and $\phi = 2\pi t + \phi$. It is clear that ξ_t is the projection of this point onto the horizontal axis. Since the distribution of P is the same uniform distribution on the circumference for any t, the distribution of the projection of P, that is, the distribution of ξ_t does depend on t. Similarly, it can be shown that not only the distribution of η_t does not depend on t, but, for all n and $t_1\,,t_2\,,\dots\,t_n$, the distribution of $(\,\eta_{t_1+\theta}\,,\eta_{t_2+\theta}\,,\,\dots\,,\eta_{t_n+\theta}\,)$ does not depend on θ, either. We omit the details of this calculation.

4. Exercise.

Sketch some possible realizations of the process. □

2. Ergodicity of the Expected Value

The **ergodicity of the expected value** of a weakly stationary process $\eta_t \; (t = 0, 1, 2, \ldots)$, by definition, means that the average value of $\eta_1, \eta_2, \ldots, \eta_T$ approaches the expected value m in some sense when $T \to \infty$:

$$\lim_{T \to \infty} \left(\frac{1}{T} \cdot \sum_{t=0}^{T-1} \eta_t \right) = m .$$

An often used sense for limits in the theory of stochastic processes is the so called **mean square sense**, which means that the expected value of the square of the difference between

$$\frac{1}{T} \cdot \sum_{t=0}^{T-1} \eta_t$$

and m goes to 0, as $T \to \infty$:

$$M \left(\left(\left(\frac{1}{T} \cdot \sum_{t=0}^{T-1} \eta_t \right) - m \right)^2 \right) \to 0 .$$

The expression on the left side of this relation is obviously the variance σ_T^2 of

$$\frac{1}{T} \cdot \sum_{t=0}^{T-1} \eta_t .$$

This is why we calculate this variance. Using the fact that the variance of a sum can be expressed in terms of the covariance function, we get that

$$\sigma_T^2 = \frac{1}{T^2} \sum_{s=0}^{T-1} \sum_{t=0}^{T-1} b(s, t) .$$

Using stationarity, we get

$$\sigma_T^2 = \frac{1}{T^2} \sum_{s=0}^{T-1} \sum_{t=0}^{T-1} B(t - s) .$$

Using that $B(u)$ is an even function, simple algebraic calculation yields that

$$\sigma_T^2 = \frac{1}{T^2} \left(T \cdot B(0) + 2 \cdot \sum_{u=1}^{T-1} \left(T - u \right) \cdot B(u) \right)$$

$$= \frac{1}{T} B(0) + \frac{2}{T} \cdot \sum_{u=1}^{T-1} \frac{T-u}{T} \cdot B(u) .$$

This implies that

$$\sigma_T^2 \leq \frac{1}{T} \left| B(0) \right| + \frac{2}{T} \cdot \sum_{u=1}^{T-1} \left| B(u) \right| .$$

Now, we can conclude that the condition

$$\lim_{T \to \infty} \frac{1}{T} \cdot \sum_{u=1}^{T-1} | B(u) | = 0$$

is sufficient for the ergodicity of the expected value.

Let us notice that the relation $B(u) \to 0$ when $u \to \infty$ is enough to guarantee this condition. Thus, the condition

$$B(u) \to 0 \quad \text{when} \quad u \to \infty$$

guarantees the ergodicity of the expected value for the weakly stationary process η_t $(t = 0, 1, 2, \ldots)$.

3. Ergodicity of a State in a Stationary Markov Chain

Let ξ_t $(t = 0, 1, 2, \ldots)$ be a stationary Markov chain, and let us choose and fix a state of it. Let us denote by p the probability that the Markov chain is in the chosen state at a given time instant.

Let η_t be defined by the formula:

$$\eta_t = 1 \text{ if } \xi_t \text{ equals the chosen state, and } 0 \text{ otherwise.}$$

Obviously, the process η_t $(t = 0, 1, 2, \ldots)$ is stationary both in the weak and the strong sense (Explain why?). The expected value of η_t is equal to p. In order to show that

$$\lim_{T \to \infty} \left(\frac{1}{T} \cdot \sum_{t=0}^{T-1} \eta_t \right) = p$$

in the mean square sense, let us calculate the covariance $B(u)$ between η_0 and η_u:

$$B(u) = M(\eta_0 \cdot \eta_u) - M(\eta_0) \cdot M(\eta_u) =$$

$$= P(\eta_0 = 1, \eta_u = 1) - P(\eta_0 = 1) \cdot P(\eta_u = 1) =$$

$$= P(\eta_0 = 1) \cdot P(\eta_u = 1 \mid \eta_0 = 1) - P(\eta_0 = 1) \cdot P(\eta_u = 1).$$

$$= P(\eta_0 = 1) \cdot \Big(P(\eta_u = 1 \mid \eta_0 = 1) - P(\eta_u = 1) \Big).$$

If the Markov chain is stable, then $P(\eta_u = 1 \mid \eta_0 = 1) \to P(\eta_u = 1)$, and so $B(u) \to 0$, if $u \to \infty$, which yields the following theorem.

Theorem. (Ergodicity of a State in a Stationary Markov Chain)

If ξ_t $(t = 0, 1, 2, \ldots)$ is a stationary, stable Markov chain, and a state is chosen, then the event that the process is at the chosen state is ergodic, that is, if η_t and p are defined as above, then

$$\lim_{T \to \infty} \left(\frac{1}{T} \cdot \sum_{t=0}^{T-1} \eta_t \right) = p$$

in the mean square sense.

6. PROBLEMS

1. TEXT IN A GIVEN LANGUAGE

Write a 0 for each consonant an a 1 for each vowel in a sufficiently (say, infinitely) long everyday text written in a given language. Would you consider the arising process strongly stationary and ergodic?

2. TEXT IN A RANDOMLY CHOSEN LANGUAGE (continued)

Choose now the language at random: Hungarian or English with equal probabilities. Be convinced that, though the process is strongly stationary, it is not ergodic.

3. WAVY WATER SURFACE

Consider a sequence of points (denoted by $0, 1, 2, \ldots$) on a line in the middle of a large lake or the ocean. Assume that the distance between any two adjacent points is the same. At any time instant, the waves make the water level at each point a random variable, so we have a process $\xi_0, \xi_1, \xi_2, \ldots$ describing the water level at the given points. Can you give an explanation why this process can be considered strongly stationary? Do not be confused: the physical meaning of the parameter t in ξ_t $(t = 0, 1, 2, \ldots)$ is now place, not time as usual!

4. STATIONARY PROCESS WITH RANDOM PERIOD

Choose a number k, called period, according to a given distribution. Then, if k is given, choose a $0-1$ **-**sequence of length k, called the seed of the sequence, according to some distribution given for each k. Finally, choose a number i according to uniform distribution on the set $\{1, 2, \ldots, k\}$. Consider now the infinite random sequence with period k and seed $x_i, x_{i+1}, \ldots, x_{k-1}, x_k, x_1, \ldots, x_{i-1}$:

$$x_i, x_{i+1}, \ldots, x_{k-1}, x_k, x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_{k-1}, x_k, x_1, \ldots, x_{i-1}, \ldots$$

Prove that the random infinite process is strictly stationary. Is it ergodic?

5. RANDOMLY SHIFTED DETERMINISTIC PERIODICAL SEQUENCE

Take a periodical infinite sequence with period d. Then apply a random shift to the sequence so that its amount is uniformly distributed on $\{0, 1, \ldots, d-1\}$. Show that the generated process is stationary.

6. MIXTURE OF STATIONARY PROCESSES

Show that the mixture of stationary processes is stationary, too. Be convinced that a process defined as in the problem entitled "Stationary Process with Random Period" is a mixture of "Randomly Shifted Deterministic Periodical Sequences"

7. BOTH WEAKLY AND STRONGLY STATIONARY

Let the random variable ψ take the values $0, 1, 2, 3$ with equal probabilities. Define a process by $\xi_t = \cos\left(\frac{\pi}{2}(t+\psi)\right) \quad (t = 0, 1, 2, \ldots)$.

a) Sketch the possible realizations of the process.

b) Show that the process is weakly stationary.

c) Show that the process is stationary in the strong sense, too.

8. MOVING AVERAGE

Let the random variables η_n $(n = 0, \pm 1, \pm 2, \dots)$ be uncorrelated with expected value 0 and variance σ^2 . In each case below, show that the process ξ_t $(t = 0, \pm 1, \pm 2, \dots)$ is weakly stationary and the ergodicity of the expected value holds. Find also the one-variable covariance function of the process:

a) $\xi_t = \eta_t + \eta_{t-1}$;

b) $\xi_t = 5\,\eta_t + 4\,\eta_{t-1} - 3\,\eta_{t-2}$;

c) $\xi_t = a_0\,\eta_t + a_1\,\eta_{t-1} + \dots + a_n\,\eta_{t-n}$;

Processes defined like this are called **moving averages**.

9. MOVING AVERAGE OF A STATIONARY PROCESS

Let the process η_n $(n = 0, \pm 1, \pm 2, \dots)$ be weakly or strongly stationary. In each case below, show that the process ξ_t $(t = 0, \pm 1, \pm 2, \dots)$ is stationary in the same sense:

a) $\xi_t = \eta_t + \eta_{t-1}$;

b) $\xi_t = 5\,\eta_t + 4\,\eta_{t-1} - 3\,\eta_{t-2}$;

c) $\xi_t = a_0\,\eta_t + a_1\,\eta_{t-1} + \dots + a_n\,\eta_{t-n}$;

Processes defined like this might be called **moving averages** average of the original process.

10. Let the random variables η_n $(n = 0, \pm 1, \pm 2, \dots)$ be independent, identically distributed random variables. In each case below, show that the moving average process ξ_t $(t = 0, \pm 1, \pm 2, \dots)$ is strongly stationary.

a) $\xi_t = \eta_t + \eta_{t-1}$;

b) $\xi_t = 5\,\eta_t + 4\,\eta_{t-1} - 3\,\eta_{t-2}$;

c) $\xi_t = a_0\,\eta_t + a_1\,\eta_{t-1} + \dots + a_n\,\eta_{t-n}$;

11. AUTOREGRESSIVE PROCESS

Let the random variables η_n $(n = 0, \pm 1, \pm 2, \dots)$ be uncorrelated with expected value 0 and variance σ^2 , and let λ be a real number. Let us define the process ξ_t $(t = 0, \pm 1, \pm 2, \dots)$ by the infinite sum

$$\xi_t = \eta_t + \lambda\,\eta_{t-1} + \lambda^2\,\eta_{t-2} + \lambda^3\,\eta_{t-3} + \dots .$$

Assume that the series is convergent.

a) Exchange the summation and the expected value operation to show that the process is weakly stationary, and find the one-variable covariance function of the process.

b) Prove that the process satisfies the equation $\xi_t = \lambda \xi_{t-1} + \eta_t$.

Processes defined like this are called **first order autoregressive processes**.

12. AUTOREGRESSIVE PROCESS OF HIGHER ORDER

Let the random variables η_n ($n = 0, \pm 1, \pm 2, \ldots$) be uncorrelated with expected value 0 and variance σ^2 , and let a_n ($n = 0, 1, 2, \ldots$) be real numbers. Let us define the process ξ_t ($t = 0, 1, 2, \ldots$) by the infinite sum

$$\xi_t = a_0 \eta_t + a_1 \eta_{t-1} + a_2 \eta_{t-2} + a_3 \eta_{t-3} + \ldots .$$

(Assume that this sum is convergent, and summation and the expected value operation are interchangeable.)

a) Show that the process is weakly stationary, and find the one-variable covariance function of the process.

Find relations between the numbers a_n so that the process ξ_t ($t = 0, 1, 2, \ldots$) satisfies the equation

b) $\xi_t = 0.4 \xi_{t-1} + 0.7 \xi_{t-2} + \eta_t$;

c) $\xi_t = \lambda_1 \xi_{t-1} + \lambda_2 \xi_{t-2} + \eta_t$;

d) $\xi_t = \lambda_1 \xi_{t-1} + \lambda_2 \xi_{t-2} + \lambda_3 \xi_{t-3} + \eta_t$;

e) $\xi_t = \lambda_1 \xi_{t-1} + \lambda_2 \xi_{t-2} + \ldots + \lambda_p \xi_{t-p} + \eta_t$.

A process satisfying the equation given in e) is called **autoregressive processes of order p** .

13. INFINITE SUM/1

Let the random variables η_n ($n = 0, \pm 1, \pm 2, \ldots$) be independent and identically distributed, and let λ be a real number. Let us define the process ξ_t ($t = 0, \pm 1, \pm 2, \ldots$) by the infinite sum

$$\xi_t = \eta_t + \lambda \eta_{t-1} + \lambda^2 \eta_{t-2} + \lambda^3 \eta_{t-3} + \ldots .$$

Assume that the series is convergent. Show that the process is strongly stationary.

14. INFINITE SUM/2

Let the random variables η_n ($n = 0, \pm 1, \pm 2, \ldots$) be independent and identically distributed, and let a_n ($n = 0, 1, 2, \ldots$) be real numbers. Let us define the process ξ_t ($t = 0, 1, 2, \ldots$) by the infinite sum

$$\xi_t = a_0 \eta_t + a_1 \eta_{t-1} + a_2 \eta_{t-2} + a_3 \eta_{t-3} + \ldots .$$

(Assume that the sum is convergent). Show that the process is strongly stationary.

15. SIN-COSINE PROCESS

Assume that A and B are uncorrelated random variables with 0 expected value and common variance σ^2 . Consider a frequency ω from $[\,0 , \pi\,]$. Show that the process

$$\xi_t = A \cos (t \omega) + B \sin (t \omega) \qquad (t = \pm 0 , \pm 1 , \pm 2 , \ldots)$$

is weakly stationary, and its one-variable cavernous function is

$$B(u) = \sigma^2 \cos\left(u\,\omega\right) .$$

16. SUM OF SINE-COSINE PROCESSES

Assume that $A_1, A_2, \ldots, A_n$ and $B_1, B_2, \ldots, B_n$ are uncorrelated random variables with 0 expected value so that the variance of A_k and B_k is the same σ_k . Let $\omega_1, \omega_2, \ldots, \omega_n$ be frequencies from the interval $[0, \pi]$. Show that the process

$$\xi_t = \sum_{k=1}^{n} \left(A_k \cos(t\,\omega_k) + B_k \sin(t\,\omega_k) \right) \qquad (t = \pm 0, \pm 1, \pm 2, \ldots)$$

is weakly stationary, and its one-variable covariance function is

$$B(u) = \sum_{k=1}^{n} \sigma_k^2 \cos\left(u\,\omega_k\right) .$$

Remark.

It can be shown that the one-variable covariance function of any weakly stationary process can be written as

$$B(u) = \int_0^{\pi} \cos\left(u\,\omega\right) F(d\omega) ,$$

where F is the cumulative distribution function of a finite distribution on $[0, \pi]$, called spectral distribution of the frequencies.

Chapter 3

Poisson Processes

1. Point Processes

Let S be an arbitrary set, which later will be mainly a subset of the real line or some higher dimensional Euclidian space.

A finite or countably infinite subset of S will be called a **point system** in S. In many real life problems, point systems in some set S occur at random. Such random point systems are called **point processes**.

Let A be a subset of S, and let us consider a point process. At each outcome of the process a random number of points fall into A. This random variable will be denoted by ν_A, and it will be called the **occurrence-number indicator** of A. The possible values of ν_A are the nonnegative integers or the symbol "infinity". The expected value of ν_A will be denoted by $m(A)$. The set function m, being nonnegative and σ **-**additive, is a measure, not necessarily normalized on S. It is called the **expected value measure** of the process. If $A_1, A_2, \ldots, A_n$ are subsets of S, then $(\nu_{A_1}, \nu_{A_2}, \ldots, \nu_{A_n})$ represents an n **-**dimensional random variable.

It is a basic fact about point processes that, in order to describe the distribution of the point process on the set of all point systems in S, it is enough to know the distribution of all vectors of the form $(\nu_{A_1}, \nu_{A_2}, \ldots, \nu_{A_n})$, where $A_1, A_2, \ldots, A_n$ are subsets of S.

It easy to be convinced that in the previous statement the subsets $A_1, A_2, \ldots, A_n$ may be assumed to be disjoint, or more specifically, may constitute a partition of S.

If S is an Euclidian space, and A is a subset and b is a vector in S, then the set A can be shifted by the vector b. Let $A+b$ denote the subset generated by the shift. A point process is called **homogeneous** if the distribution of ν_{A+b} is the same as the distribution of ν_A for all A and b.

A point process is called a process with **independent increments** if for disjoint subsets $A_1, A_2, \ldots, A_n$, the random variables $\nu_{A_1}, \nu_{A_2}, \ldots, \nu_{A_n}$ are independent.

2. Poisson Processes

A point process is called a **Poisson process** if it is of independent increments, and for every subset A, the random variable ν_A has a Poisson distribution. A Possion process is, clearly, characterized by its expected value measure m. If m is a measure on S, then the m **-**parametric Poisson process is the point process with independent increments, for which the number of points falling in a subset A follows Poisson distribution with parameter $m(A)$.

If S is the real line, and λ is a positive number, then the **homogeneous** Poisson process with parameter λ is the Poisson process for which the expected value measure is λ times the length measure. This means that for every subset A, the expected value of the number of points falling in A is λ times the length of A.

If S is the plane, and λ is a positive number, then the **homogeneous** Poisson process with parameter λ is the Poisson process for which the expected value measure is λ times the area measure. This means that for every subset A, the expected value of the number of points falling in A is λ times the area of A.

More generally , if S is a subset of the Euclidian space, and λ is a positive number, then the **homogeneous** Poisson process with parameter λ is the Poisson process for which the expected value measure is λ times the Lebesgue measure.

3. Differential Equations

Let us consider a homogeneous point process with independent increments on the positive half of the real line. Let $d_k(t)$ denote the probability that exactly k points fall in the interval $[0, t]$. If the process is a homogeneous Poisson process, then – one can check easily – the following facts are true:

$$d_0(t) = P\Big(\text{no points in } [0,t]\Big) = 1 - \lambda t + o(t),$$

$$d_1(t) = P\Big(1 \text{ point in } [0,t]\Big) = \lambda t + o(t),$$

$$d_k(t) = P\Big(k \text{ points in } [0,t]\Big) = o(t) \qquad (k \geq 2),$$

that is,

$$d_0(0) = 1,$$

$$d_k(0) = 0 \qquad (k \geq 1),$$

and

$$d_0'(0) = -\lambda,$$

$$d_1'(0) = \lambda,$$

$$d_k'(0) = 0 \qquad (k \geq 2).$$

These facts can be verbalized by saying that:

for short time intervals, the probability of the occurrence of 0 points is less than 1 by a quantity which is approximately proportional to the length of the time interval;

for short time intervals, the probability of the occurrence of one point is approximately proportional to the length of the time interval;

if $k \geq 2$, then, for short time intervals, the probability of the occurrence of k points is negligible;

the constant of proportionality in these statements is everywhere the same λ .

We shall show that these properties are not only necessary but also sufficient for a homogeneous process with independent increments to be a homogeneous Poisson process with parameter λ . To prove this, notice that the following relations are obviously true:

$$d_0(t+h) = d_0(t) \cdot d_0(h) ,$$

$$d_k(t+h) = d_k(t) \cdot d_0(h) + d_{k-1}(t) \cdot d_1(h) + \sum_{i=2}^{k} d_{k-i}(t) \cdot d_i(h) .$$

Differentiating with respect to h , then substituting h by 0 , and using the assumptions, we get that

$$d_0'(t) = d_0(t) \cdot (-\lambda) ,$$

$$d_k'(t) = d_k(t) \cdot (-\lambda) + d_{k-1}(t) \cdot \lambda \qquad (k \geq 1) .$$

This is a system of differential equations, for which the conditions

$$d_0(0) = 1$$

$$d_k(0) = 0 \qquad (k \geq 1)$$

can be used as initial conditions. The first equation is easily solvable, and then recursively the others are solvable, too. The solution turns out to be:

$$d_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t} \qquad (k \geq 0) ,$$

which proves that the points in intervals of the type $[0, t]$ are distributed according to Poisson distribution with parameter λt . Since any interval can be generated by shifting an interval of this type, homogeneity guarantees that the number of points in any interval is distributed according to a Poisson distribution with parameter equal to λ times the length of the interval. Thus, the process is a homogeneous Poisson process with parameter λ .

4. Exponential Waiting Times

Another characterization of homogeneous Poisson processes is that the distances between adjacent points in the process are independent random variables, which are exponentially distributed with the same parameter. If the points mean time instances, then these distances are the **waiting times** between these time instants.

In order to show that the waiting times α_1 , α_2 , ... are independent and exponentially distributed with the parameter λ of the process, we determine the conditional distribution function of α_n under the condition that $\alpha_1 = x_1$, ... , $\alpha_{n-1} = x_{n-1}$.

$$P\left(\alpha_n < x \mid \alpha_1 = x_1 , \dots , \alpha_{n-1} = x_{n-1} \right) = P\left(\text{at least one point occurs in } (a , b) \right) ,$$

where $a = x_1 + \dots + x_{n-1}$, $b = x_1 + \dots + x_{n-1} + x$. Since the length of the interval (a , b) is equal to x , this probability is equal to

$$P\left(\alpha_n < x \mid \alpha_1 = x_1 , \dots , \alpha_{n-1} = x_{n-1} \right) = 1 - e^{-\lambda x} .$$

Since this conditional distribution function does not depend on $x_1 , \dots , x_{n-1}$, the waiting times are really independent. On the other hand the form of the distribution function shows that α_n is exponentially distributed with parameter λ .

In order to show that the independence and the exponentiality with identical parameters of the waiting times guarantees that the point process is a homogeneous Poisson process, let us consider the interval $[0 , t]$. Let us introduce the notation

$$\gamma_k = \text{the sum of the first } k \text{ waiting times} .$$

The event $\nu_{[0,t]} = k$ obviously means the same as the event

$$(\gamma_k < t) - (\gamma_{k+1} < t) .$$

Using the fact that $\gamma_{k+1} < t$ implies $\gamma_k < t$, this yields that

$$\Pr\left(\nu_{[0,t]} = k \right) = \Pr\left(\gamma_k < t \right) - \Pr\left(\gamma_{k+1} < t \right) .$$

Since the sum of k exponentially distributed random variables with parameter λ has gamma-distribution of order k ,

$$\Pr\left(\gamma_k < t \right) = \int_0^t \frac{\lambda^k x^{k-1}}{(k-1)!} e^{-\lambda x} \, dx .$$

Using partial integration, we get that

$$\int_0^t \frac{\lambda^k x^{k-1}}{(k-1)!} e^{-\lambda x} \, dx = \left(\frac{\lambda^k x^k}{k!} e^{-\lambda x} \right)_{x=0}^{x=t} + \int_0^t \frac{\lambda^{k+1} x^k}{k!} e^{-\lambda x} \, dx .$$

that is,

$$\int_0^t \frac{\lambda^k x^{k-1}}{(k-1)!} e^{-\lambda x} \, dx = \frac{(\lambda t)^k}{k!} e^{-\lambda t} + \int_0^t \frac{\lambda^{k+1} x^k}{k!} e^{-\lambda x} \, dx .$$

Bringing the integral on the right side to the left side, we get that

$$\int_0^t \frac{\lambda^k x^{k-1}}{(k-1)!} e^{-\lambda x} \, dx - \int_0^t \frac{\lambda^{k+1} x^k}{k!} e^{-\lambda x} \, dx = \frac{(\lambda t)^k}{k!} e^{-\lambda t} ,$$

which means that

$$\Pr\left(\gamma_k < t \right) - \Pr\left(\gamma_{k+1} < t \right) = \frac{(\lambda t)^k}{k!} e^{-\lambda t} ,$$

that is,

$$\Pr\left(\nu_{[0,t]} = k \right) = \frac{(\lambda t)^k}{k!} e^{-\lambda t} .$$

This means that the random variable $\nu_{[0,t]}$ is distributed according to Poisson distribution with parameter λ .

The case of an arbitrary interval $[A, B]$ can be reduced to the interval $[0, B-A]$ by neglecting the interval $[0, A]$: at this truncation, the waiting time between A and the first point after A remains exponentially distributed because of the memoryless property of the exponential distribution.

The memoryless property also guarantees that the process is of independent increments.

5. PROBLEMS

1. CLOSEST TO THE ORIGIN ON THE LINE

Consider the homogeneous Poisson process with parameter λ on the real line. Find the distribution of the point closest to a given point.

2. CLOSEST TO THE ORIGIN ON THE PLANE

Consider the homogeneous Poisson process with parameter λ on the plane. Find the distribution of

a) the distance between the origin and the point closest to the origin;

b) the point closest to the origin.

Generalize the problem for higher dimensions and for the non-homogeneous case, too.

3. DROPPING A RANDOM NUMBER OF POINTS

Let Q be a normalized distribution on a set S . Generate a random number of points according to λ - parametric Poisson distribution, and drop each of them independently of each other onto the set according to the distribution Q . Prove that the the point system on S is Poisson process. What is the expected value measure of the process?

4. ONE POINT IN A SET

Assume that observing a Poisson point process with expected value measure m , you find that only one point falls into a set A , for which $m(A) < \infty$. How is the position of the point distributed in A ?

5. TWO POINTS IN A SET

Assume that observing a Poisson point process with expected value measure m , you find that two points fall into a set A , for which $m(A) < \infty$. Show that they are distributed the same way as if two points were independently dropped onto A according to the normalized restriction of m onto A .

6. n POINTS IN A SET

Assume that observing a Poisson point process with expected value measure m , you find that n points fall into a set A , for which $m(A) < \infty$. Show that they are distributed the same way as if n points were independently dropped onto A according to the normalized restriction of m onto A .

7. ELECTRIC BULBS

Assume that electric bulbs you use have independent life times distributed according to exponential distribution with a common parameter λ . You start using one at time instant 0 , and whenever a bulb goes wrong, you replace it with another. Fixing a time instant t , consider the bulb being in use at time instant t .

a) The amount of time for which this bulb has been used before the time instant t , defines a random variable. Determine its distribution, and calculate its expected value.

b) The amount of time for which this bulb will be used after the time instant t , defines a random variable. Determine its distribution, and calculate its expected value.

c) The total life time of the bulb being in use at time t is the sum of the two random variables examined in questions a) and b) . Use this fact to show that the expected value of the total life time of the bulb being in use at time t is $\frac{1}{\lambda} + \frac{1}{\lambda}(1 - e^{-\lambda t})$.

d) This expected value is larger than the expected value $\frac{1}{\lambda}$ of the life time of the bulbs, which may seem a contradiction. Try to give a heuristic explanation why it is natural that the expected value of the total life time of the bulb being in use at time t is larger than the common expected value of the life time of the bulbs.

8. TRANSFORMATION OF A POINT PROCESS

Let t be an increasing transformation from R^1 into R^1 . Transform each point of a point process by the transformation t . Show that a Poisson process, under such a transformation, yields another Poisson process. What is the expected value measure of the new process, if the original process is a homogeneous Poisson process?

9. PAINTING THE POINTS RED OR GREEN

Each point of a Possion process, independently of the others, is painted red with probability p , and green with probability q .

a) Show that both the red and the green points constitute a Poisson process. Find their expected value measures.

b) Show that the process defined by the red points is independent of the process defined by the green points.

10. POINTS JUMPING OUT OF THE LINE INTO THE PLANE

Assume that a normalized distribution, which will serve as a conditional distribution, is assigned to each vertical line in the plane. Each point of a Poisson process on the horizontal axis, independently of the others, is moved vertically into a new point on the plane generated at random according to the distribution assigned to the vertical line which passes the point. Show that the new points constitute a Poisson process, too. Assuming that the expected value measure and the conditional measures have density functions, show that the new process also has one. Find their relationship.

Chapter 4

Birth and Death Processes

1. Jump Processes – Markov Processes

Consider a finite or countably infinite number of states denoted by $0, 1, 2, \ldots$. Assign each state i a non-negative parameter λ_i and a normalized distribution $(\pi_{ij} : j \neq i)$ on the set of all other states. Consider also an initial distribution $D = (d_0, d_1, d_2, \ldots)$ on the set of all states. Imagine now the following process:

First, a state i is chosen at random according to the initial distribution, and a particle is placed there. The particle spends a random amount of time, called **resting time**, at this state so that the distribution of its resting time is exponential with parameter λ_i, and then it jumps into one of the other states. The state where it goes is chosen at random according to the distribution $(\pi_{ij} : j \neq i)$. Then the particle spends again a random amount of time at the new state so that the distribution of its resting time is again exponential with the parameter assigned to this new state, and then it performs again a jump into one of the other states chosen at random according to the distribution assigned to the state from where the particle is leaving. And so on, this goes on forever. Such processes are called (continuous-time and discrete-state) **jump processes**.

Let us denote the state of the particle at time instant t by ξ_t . The memoryless property of the exponential distribution guarantees that if the present state of the particle is known, then the past has no influence on the distribution of the future of the process, that is, for any time instants $s_1 < \ldots < s_r < s$ and for any time $t > 0$ (here $s_1, \ldots, s_r$ represent time instants in the past, s represents the present, $s + t$ represents a future time instant) and any states $i_1, i_2, \ldots, i_r$, i and j, the conditional probability

$$\Pr\left(\xi_{s+t} = j \mid \xi_{s_1} = i_1, \ldots, \xi_{s_r} = i_r, \xi_s = i \right)$$

does not depend on $s_1, \ldots, s_r$ and $i_1, \ldots, i_r$, but it is equal to

$$\Pr\left(\xi_{s+t} = j \mid \xi_s = i \right).$$

The definition guarantees the homogeneity of the process, that is, the conditional probability $\Pr\left(\xi_{s+t} = j \mid \xi_s = i \right)$ does not depend on s :

$$\Pr\left(\xi_{s+t} = j \mid \xi_s = i \right) = \Pr\left(\xi_t = j \mid \xi_0 = i \right).$$

Thus, the jump process $(\xi_t : t \geq 0)$ is a **homogeneous Markov process**.

It can be shown that every homogeneous Markov process can be represented as a jump process.

2. Transition Matrix – Infinitesimal Generator

Let us agree that, in this chapter, the symbol $\approx$ means that the expressions on its two sides differ only at an amount of $o(t)$ when $t \to 0$. $o(t)$ means a quantity which divided by t goes to 0:

$$\lim_{t \to 0} \frac{o(t)}{t} = 0 .$$

Remember that two functions differ only at a quantity of $o(t)$, when $t \to 0$, if and only if their substitution values and their derivatives at 0 are equal.

The following facts are clear:

$$\Pr(\text{the particle leaves state } i \text{ before } t \mid \text{it is in state } i \text{ now}) = 1 - e^{-\lambda_i t} \approx \lambda_i t ,$$

$$\Pr(\text{the particle stays at state } i \text{ until } t \mid \text{it is in state } i \text{ now}) = e^{-\lambda_i t} \approx 1 - \lambda_i t .$$

Since the first conditional probability is, for small values of t, approximately proportional to t, and the constant of proportionality is λ_i, the coefficient λ_i may be called as the **proportionality coefficient in the probability of leaving state i**. If λ_i is close to 0, then the particle is likely not to leave state i too early, if it is a large positive constant, then the particle will leave state i with more intensity.

The second conditional probability shows that, for small values of t, the probability of remaining at state λ_i differs from 1 approximately by a quantity $(-\lambda_i) \cdot t$, thus, $(-\lambda_i)$ may be called as the **proportionality coefficient in the change of the probability of remaining at state i**. The fact that $(-\lambda_i)$ is a negative number shows that the chances of remaining at state i decrease as the time passes. If $(-\lambda_i)$ is close to 0, then the probability of remaining at state i remains, for small values of t, close to 1, if it is a large negative constant, then the probability of remaining at state i will decrease with more intensity.

Since

$$\Pr(\text{the particle leaves state } i \text{ before } t \mid \text{it is in state } i \text{ now}) \approx \lambda_i t ,$$

it is intuitively clear that the probability that more than one jumps occur in a small time interval of length t is of order t^2, thus

$$\Pr(\text{more than one jumps occur before } t) = o(t) .$$

Let us introduce the notation $p_{ij}(t)$ for the **transition probabilities**:

$$p_{ij}(t) = \Pr\Big(\xi_t = j \mid \xi_0 = i\Big) = \Pr\Big(\text{the particle is in } j \text{ at } t \mid \text{it is in } i \text{ now}\Big) .$$

The fact that the probability of more than one jump is $o(t)$ clearly implies that

$$p_{ii}(t) = \Pr(\text{the particle is in } i \text{ at } t \mid \text{it is in } i \text{ now}) \approx$$

***

$$\approx \Pr(\text{the particle stays at } i \text{ until } t \mid \text{it is in } i \text{ now}),$$

thus,

$$p_{ii}(t) \approx 1 - \lambda_i t,$$

which means that the derivative of $p_{ii}(t)$ at 0 is equal to $(-\lambda_i)$:

$$p'_{ii}(0) = (-\lambda_i).$$

The fact that the probability of more than one jump is $o(t)$ implies the first step in the following argument. The other steps should be clear. For $i \neq j$, we have

$$p_{ij}(t) = \Pr(\text{the particle is in } j \text{ at } t \mid \text{it is in } i \text{ now}) \approx$$

***

$$\approx \Pr(\text{the particle leaves } i \text{ before } t\text{, and when it leaves } i\text{, it jumps to } j \mid \text{it is in } i \text{ now}) =$$

$$= \Pr(\text{the particle leaves } i \text{ before } t \mid \text{it is in } i \text{ now}) \cdot$$

$$\cdot \Pr(\text{when it leaves } i\text{, it jumps into } j \mid \text{it is in } i \text{ now}) =$$

$$= (1 - e^{-\lambda_i t}) \cdot \pi_{ij} \approx \lambda_i t \cdot \pi_{ij} = \lambda_i \pi_{ij} t,$$

which means that, for $j \neq i$, the derivative of $p_{ij}(t)$ at 0 is equal to $\lambda_i \pi_{ij}$:

$$p'_{ij}(0) = \lambda_i \pi_{ij} \qquad (j \neq i).$$

We shall refer to the two approximate equalities marked above by three asterisks (***), as the **first order approximations of the transition probabilities**.

The matrix composed of the transition probabilities $p_{ij}(t)$ is called the **transition matrix**:

$$P(t) = \left(p_{ij}(t) \right).$$

The total probability formula combined with the Markov property yields that

$$p_{ij}(t+h) = \sum_k p_{ik}(t) \cdot p_{kj}(h)$$

for any t, h and i, j, that is, the transition matrix satisfies the so called **Chapman – Kolmogorov equation**:

$$P(t+h) = P(t) \cdot P(h).$$

Taking formally the derivative with respect to h, we get that

$$P'(t+h) = P(t) \cdot P'(h).$$

Substituting h by 0, we get that

$$P'(t) = P(t) \cdot P'(0).$$

The matrix $P'(0)$ is called the **infinitesimal generator** of the process, and is denoted by A:

$$A = P'(0).$$

The differential equation

$$P'(t) = P(t) \cdot A$$

is called the **forward Kolmogorov – Feller equation for the transition matrix**. This is a system of an infinite number of differential equations for the transition probabilities, because denoting the elements of A by a_{ij}, we get :

$$p'_{ij}(t) = \sum_k p_{ik}(t) \cdot a_{kj} \qquad (i, j = 0, 1, 2, \ldots).$$

The differential equation can be extended by the initial condition:

$$P(0) = E,$$

where E is the infinite times infinite unit matrix, that is,

$$p_{ij}(0) = \delta_{ij} \qquad (i, j = 0, 1, 2, \ldots),$$

where δ_{ij} is the Kronecker symbol: $\delta_{ij} = 1$ if $i = j$, and $\delta_{ij} = 0$ if $i \neq j$. Changing t and h in the Chapman – Kolmogorov equation, then differentiating formally with respect to h, and then replacing h by 0, we get the **backward Kolmogorov – Feller equation for the transition matrix**:

$$P'(t) = A \cdot P(t),$$

which means another system of an infinite number of differential equations for the transition probabilities:

$$p'_{ij}(t) = \sum_k a_{ik} \cdot p_{kj}(t) \qquad (i, j = 0, 1, 2, \ldots).$$

This system can be extended by the same initial conditions as above.

If the elements of A are denoted by a_{ij}, then, according to the above calculations,

$$a_{ii} = (-\lambda_i) \qquad \text{and} \qquad a_{ij} = \lambda_i \pi_{ij} \qquad (j \neq i),$$

that is, the infinitesimal generator matrix has the following form:

$$A = \begin{bmatrix} (-\lambda_0) & \lambda_0\pi_{01} & \lambda_0\pi_{02} & \ldots \\ \lambda_1\pi_{10} & (-\lambda_1) & \lambda_1\pi_{12} & \\ \lambda_2\pi_{20} & \lambda_2\pi_{21} & (-\lambda_2) & \\ \vdots & & & \end{bmatrix}.$$

We can also express λ_i and π_{ij} in terms of the elements of the infinitesimal generator:

$$\lambda_i = (-a_{ii}) \qquad \text{and} \qquad \pi_{ij} = \frac{a_{ij}}{-a_{ii}} \qquad (j \neq i).$$

Since the row-sums of the infinitesimal generator are obviously equal to 0, the negative of the diagonal element of each row is equal to the sum of all the other elements in the row:

$$-a_{ii} = \sum_{j:j\neq i} a_{ij}.$$

Thus, the distribution $(\pi_{ij} : j \neq i)$ characterizing the jumps from i can be deduced from the i-th row of the infinitesimal generator by normalizing the non-diagonal elements of the row:

$$\pi_{ij} = \frac{a_{ij}}{\sum_{j:j\neq i} a_{ij}} \qquad (j \neq i).$$

It is advantageous to abbreviate the notation of the infinitesimal generator in the examples later by writing an asterisk (*) instead of the diagonal element like this:

$$A = \begin{bmatrix} * & a_{01} & a_{02} & \cdots \\ a_{10} & * & a_{12} & \\ a_{20} & a_{21} & * & \\ \vdots & & & \end{bmatrix} .$$

This does not mean a loss of information, because

$$a_{ii} = - \left(\sum_{j\,:\,j \neq i} a_{ij} \right) .$$

In most examples, replacing the diagonal elements by $*$, the matrix becomes esthetically nicer. It is also useful to show the states in front of and above the matrix. This is why we shall give the infinitesimal generators in an arrangement like this:

$$\begin{array}{c c ccccc c} & & 0 & 1 & 2 & 3 & 4 & \\ & & & & & & & \\ 0 & & * & a_{01} & a_{02} & a_{03} & a_{04} & \\ 1 & & a_{10} & * & a_{12} & a_{13} & a_{14} & \\ 2 & & a_{20} & a_{21} & * & a_{23} & a_{24} & \cdots \\ 3 & & a_{30} & a_{31} & a_{32} & * & a_{34} & \\ \cdots & & & & & & & \end{array}$$

A jump process is called a **birth and death process** if only adjacent states are available from each state, that is, $\pi_{ij} = 0$ unless $j = i - 1$ or $j = i + 1$, or equivalently, $a_{ij} = 0$ unless $j = i - 1$ or $j = i$ or $j = i + 1$, that is, the infinitesimal generator has the following form:

	0	1	2	3	...
0	*	μ_0			
1	ν_1	*	μ_1		
2		ν_2	*	μ_2	
3			ν_3	*	
...					

It is clear that a birth and death process is the continuous-time generalization of the random walk on the real line, for which jumps occur at random time instants, and left and right jumps have a probability which may depend on the state from where the particle is leaving.

A jump process is called a **pure birth process** if only right states are available from each state, that is, $\pi_{ij} = 0$ unless $j = i + 1$, or equivalently, $a_{ij} = 0$ unless $j = i$ or $j = i + 1$, that is, the infinitesimal generator has the following form:

	0	1	2	3	...
0	*	μ_0			
1		*	μ_1		
2			*	μ_2	
3				*	
...					

It is clear that a pure birth process is a continuous-time random walk on the real line, for which jumps occur at random time instants, and only right jumps are possible from each state.

In applications, most of the jump processes are birth and death processes, so will be our examples.

1. Example: POISSON PROCESS

Assume that cars pass a certain point on a highway at random time instants according to μ -parametric homogeneous Poisson process. Let ξ_t be the number of cars passing before t . Show that the infinitesimal generator of the process is

	0	1	2	3	4	
0	*	μ				
1		*	μ			
2			*	μ		......
3				*	μ	
...					...	

Solution.

The element $p_{ij}(t)$ of the transition matrix can be explicitly given. If $i \le j$, then

$$p_{ij}(t) = \Pr\left(\xi_{s+t} = j \mid \xi_s = i \right)$$

$$= \Pr\left(j - i \text{ cars pass during a time interval of length } t \right) = \frac{(\mu t)^{j-i}}{(j-i)!} e^{-\mu t},$$

and $p_{ij}(t) = 0$ otherwise. Taking the derivative of $p_{ij}(t)$, and then substituting 0, we get that

$$a_{ij} = p'_{ij}(0) = -\mu \qquad \text{if } j = i,$$

. $$a_{ij} = p'_{ij}(0) = \mu \qquad \text{if } j = i + 1,$$

$$a_{ij} = p'_{ij}(0) = 0 \qquad \text{otherwise,}$$

which shows that the infinitesimal generator is what is given above. □

2. Example: BACTERIA SPLITTING

Assume that any member of a bacterium population, independently of the others, splits into two after its birth at a random time which is exponentially distributed with parameter μ. Let ξ_t mean the number of bacteria at time t. Prove that the infinitesimal generator of the process is

	1	2	3	4	5	
1	*	μ				
2		*	2μ			
3			*	3μ		......
4				*	4μ	
...					...	

Solution.

The element $p_{i,i+1}(t)$ of the transition matrix can be explicitly calculated. If $\gamma_1, \dots, \gamma_i, \gamma_{i+1}, \gamma_{i+2}$ denote independent, exponentially distributed random variables with parameter μ, then

$$p_{i,i+1}(t) = \Pr\left(\xi_{s+t} = i+1 \mid \xi_s = i \right)$$

$$= \Pr\left(\xi_t = i+1 \mid \xi_0 = i \right)$$

$$= \Pr\left(\begin{matrix} \text{exactly one bacteria splits between the time instants } 0 \text{ and } t \\ \text{and none of the two new ones splits before time instant } t \end{matrix} \;\middle|\; \begin{matrix} \text{there are } i \text{ bacteria} \\ \text{at time instant } 0 \end{matrix} \right)$$

$$= i \cdot \Pr\left(\gamma_1 > t, \dots, \gamma_{i-1} > t, \gamma_i < t, \gamma_i + \gamma_{i+1} > t, \gamma_i + \gamma_{i+2} > t \right)$$

$$= i \cdot \Pr(\gamma_1 > t) \dots \Pr(\gamma_{i-1} > t) \Pr\left(\gamma_i < t, \gamma_{i+1} > t - \gamma_i, \gamma_{i+2} > t - \gamma_i \right)$$

$$= i \cdot \left(e^{-\mu t} \right)^{(i-1)} \int_0^t \Pr(\gamma_{i+1} > t - x, \gamma_{i+2} > t - x \mid \gamma_i = x)\, \mu e^{-\mu x}\, dx$$

$$= i \cdot e^{-(i-1)\mu t} \int_0^t \left(e^{-\mu(t-x)} \right)^2 \mu e^{-\mu x}\, dx$$

$$= i \cdot e^{-(i-1)\mu t} \int_0^t e^{-2\mu(t-x)} \mu e^{-\mu x}\, dx$$

$$= i \cdot e^{-(i+1)\mu t} \int_0^t \mu e^{\mu x}\, dx$$

$$= i \cdot e^{-(i+1)\mu t} \left(e^{\mu t} - 1 \right) = i \cdot \left(e^{-i\mu t} - e^{-(i+1)\mu t} \right).$$

Taking the derivative, and then substituting 0, we get that

$$a_{i,i+1} = p'_{i,i+1}(0) = i\,\mu .$$

One can calculate that

$$a_{i,i} = p'_{i,i}(0) = -i\,\mu .$$

Intuitively clear, and it can be calculated, too, that, if j differs from i and $i+1$, then

$$a_{i,j} = p'_{i,j}(0) = 0 .$$

Remark 1.

Using the first order approximations of the transition probabilities suggested by the approximate equalities marked by *** above, we get that

$$p_{i,i+1}(t)$$

$$\approx i \cdot \Pr\Big(\gamma_1 > t, \ \dots, \gamma_{i-1} > t, \gamma_i < t \Big)$$

$$= i \cdot \Big(e^{-\mu t} \Big)^{(i-1)} \Big(1 - e^{-\mu t} \Big)$$

$$= i \cdot \Big(e^{-(i-1)\mu t} - e^{-i\mu t} \Big).$$

Be convinced that the derivative of the last expression at 0 is $(-\mu t)$, which is really the same as the derivative of $p_{i,i+1}(t)$ at 0.

Remark 2.

If somebody determines the value of the expressions

$$i \cdot \Pr\Big(\gamma_1 > t, \ \dots, \gamma_{i-1} > t, \gamma_i < t, \gamma_i + \gamma_{i+1} > t, \gamma_i + \gamma_{i+2} > t \Big),$$

$$i \cdot \Pr\Big(\gamma_1 > t, \ \dots, \gamma_{i-1} > t, \gamma_i < t \Big),$$

then it turns out, as above, that their derivatives at 0 are the same. Pretend now that you are not able to determine the first expression, only the second one, which is definitely simpler. In order to guarantee that the derivatives of the two expressions at 0 are the same, it is enough to show that their difference is of order $o(t)$, when $t \to 0$. The difference can be estimated relatively simply:

$$i \cdot \Pr\Big(\gamma_1 > t, \ \dots, \gamma_{i-1} > t, \gamma_i < t \Big)$$

$$- i \cdot \Pr\Big(\gamma_1 > t, \ \dots, \gamma_{i-1} > t, \gamma_i < t, \gamma_i + \gamma_{i+1} > t, \gamma_i + \gamma_{i+2} > t \Big)$$

$$= i \cdot \Pr\Big(\gamma_1 > t, \ \dots, \gamma_{i-1} > t, \gamma_i < t, \overline{\gamma_i + \gamma_{i+1} > t, \gamma_i + \gamma_{i+2} > t} \Big)$$

$$= i \cdot \Big(e^{-\mu t} \Big)^{(i-1)} \Pr\Big(\gamma_i < t \text{ and } \overline{\gamma_i + \gamma_{i+1} > t \text{ and } \gamma_i + \gamma_{i+2} > t} \Big)$$

$$= i \cdot \Big(e^{-\mu t} \Big)^{(i-1)} \Pr\Big(\gamma_i < t \text{ and } (\gamma_i + \gamma_{i+1} < t \text{ or } \gamma_i + \gamma_{i+2} < t) \Big).$$

The probability in the last line can be estimated by

$$\Pr\Big(\gamma_i < t \text{ and } (\gamma_i + \gamma_{i+1} < t \text{ or } \gamma_i + \gamma_{i+2} < t) \Big)$$

$$\leq \Pr\Big(\gamma_i + \gamma_{i+1} < t \text{ or } \gamma_i + \gamma_{i+2} < t \Big)$$

$$\leq \Pr\Big(\gamma_i + \gamma_{i+1} < t \Big) + \Pr\Big(\gamma_i + \gamma_{i+2} < t \Big) = 2 \Pr\Big(\gamma_i + \gamma_{i+1} < t \Big)$$

$$= 2 \int_0^t \mu^2 \, x \, e^{-\mu x} \, dx \;\leq\; 2 \int_0^t \mu^2 \, x \; dx \;= 2 \, \mu^2 \, \tfrac{t^2}{2} \;= \mu^2 \, t^2 \, .$$

(At the beginning of the last line, we used the fact that the distribution of $\gamma_i + \gamma_{i+1}$ is a second order gamma distribution.) The estimation shows that the difference is really $o(t)$, when $t \to 0$. □

In most problems, a first order approximation of the transition probabilities is easier to calculate than the transition probabilities itself, and an estimation similar to the one made in Remark 2 may be intuitively obvious. This is why, using a first order approximation, the elements of the infinitesimal generator may be easier to determine than the elements of the transition matrix.

Note that if a first order approximation of the transition probability $p_{i,j}(t)$ is a linear expression:

$$p_{i,j}(t) \;\approx\; c_{i,j} \cdot t \, ,$$

then the element $a_{i,j}$ of the infinitesimal generator is equal to the coefficient in this linear expression

$$a_{i,j} \;=\; c_{i,j} \, .$$

If specifically, $p_{i,j}(t) \;\approx\; o(t)$, then $a_{i,j} \;=\; 0$. The most convenient way to determine the elements of the infinitesimal generator is to find linear first order approximations to the transition probabilities. In the examples later, we shall do so.

3. Example: QUEUEING WITH ONE SERVER

Assume that customers arrive to a teller's window in a bank at random time instants according to a μ-parametric homogeneous Poisson process. If the window is busy, they line up in a queue. Assume that the service time of each customer is exponentially distributed with parameter ν, independently of each other and the arrival process. Let ξ_t be the length of the queue. Prove that the infinitesimal generator of the process is

	0	1	2	3	4	
0	*	μ				
1	ν	*	μ			
2		ν	*	μ		$\ldots \nu\nu \ldots$
3			ν	*	μ	
...				...	...	

Solution.

A first order approximation of the transition probability $p_{i,\,i+1}(t)$ is obvious:

$$p_{i,\,i+1}(t) \;=\; \Pr\Big(\xi_{s+t} \;= i + 1 \;\mid\; \xi_s \;=\; i \Big)$$

$\approx$ Pr $\Big($ during a time interval of length t, the customer at the teller's window is not served,

$$\text{and a new customer arrives} \Big)$$

$$= \Pr \Big(\text{during a time interval of length } t \text{, the customer at the teller's window is not served} \Big)$$

$$\cdot \Pr \Big(\text{during a time interval of length } t \text{, a new customer arrives} \Big)$$

$$= e^{-\nu t} \cdot (1 - e^{-\mu t}) \approx (1 - \nu t) \cdot \mu t \approx \mu t \,,$$

which shows that $a_{i,\,i+1} = \mu$.

A first order approximation of the transition probability $p_{i,\,i-1}(t)$ is obvious:

$$p_{i,\,i-1}(t) = \Pr \Big(\xi_{s+t} = i - 1 \mid \xi_s = i \Big)$$

$$\approx \Pr \Big(\text{during a time interval of length } t \text{, the customer at the teller's window is served,}$$

$$\text{and no new customer arrives} \Big)$$

$$= \Pr \Big(\text{during a time interval of length } t \text{, the customer at the teller's window is served} \Big)$$

$$\cdot \Pr \Big(\text{during a time interval of length } t \text{, no new customer arrives} \Big)$$

$$= (1 - e^{-\nu t}) \cdot e^{-\mu t} \approx \nu t \cdot (1 - \mu t) \approx \nu t \,,$$

which shows that $a_{i,\,i-1} = \nu$.

If $|j - i| > 2$, then

$$p_{i,\,j}(t) = \Pr \Big(\xi_{s+t} = j \mid \xi_s = i \Big)$$

$$\leq \Pr \Big(\text{during a time interval of length } t \text{, at least } 2 \text{ of the the customers are served,}$$

$$\text{or at last } 2 \text{ new customers arrive} \Big) = o(t)$$

because both the event 'during a time interval of length t , at least 2 of the the customers are served' and the event 'during a time interval of length t , at last 2 new customers arrive' have a probability equal to $o(t)$. Thus, if $|j - i| > 2$, then $a_{i,j} = 0$, which completes the solution. □

4. Example: TELEPHONE CENTER WITH INFINITE CAPACITY

Assume that telephone talks at a telephone center start at random time instants according to μ -parametric homogeneous Poisson process. Assume that capacity of the center is so large that each talk is accepted by the center. Assume that the length of each talk is exponentially distributed with parameter ν , independently of each other and the arrival process. Let ξ_t be the number of the busy lines. Prove that the infinitesimal generator of the process is

	0	1	2	3	4	
0	*	μ				
1	ν	*	μ			
2		2ν	*	μ		......
3			3ν	*	μ	
...				...	...	

Solution.

A first order approximation of the transition probability $p_{i,\,i+1}(t)$ is obvious:

$$p_{i,\,i+1}(\,t\,) \;=\; \Pr\Big(\;\xi_{s+t} \;= i+1 \;\mid\; \xi_s \;=\; i\;\Big)$$

$$\approx\; \Pr\Big(\text{ during a time interval of length } t\,,\text{ no call ends}$$

$$\text{and a new call starts}\Big)$$

$$=\; \Pr\Big(\text{ during a time interval of length } t\,,\text{ no call ends }\Big)$$

$$\cdot\; \Pr\Big(\text{ during a time interval of length } t\,,\text{ a new call starts }\Big)$$

$$=\; \left(e^{-\nu\,t}\right)^{i} \cdot\, (\,1 \;-\; e^{-\mu t}) \;\approx (\,1-\nu t\,)^{i} \,\cdot\, \mu\,t \;\approx\; \mu\,t\,,$$

which shows that $a_{i,\,i+1} \;= \mu$.

A first order approximation of the transition probability $p_{i,\,i-1}(t)$ is obvious:

$$p_{i,\,i-1}(\,t\,) \;=\; \Pr\Big(\;\xi_{s+t} \;= i-1 \;\mid\; \xi_s \;=\; i\;\Big)$$

$$\approx\; \Pr\Big(\text{ during a time interval of length } t\,,\text{ a call ends,}$$

$$\text{and no new call starts}\Big)$$

$$=\; \Pr\Big(\text{ during a time interval of length } t\,,\text{ a call ends }\Big)$$

$$\cdot\; \Pr\Big(\text{ during a time interval of length } t\,,\text{ no new call starts }\Big)$$

$$\approx\; i\,(\,1\; - e^{-\nu\,t}\,)\,\left(e^{-\nu\,t}\right)^{i-1} \cdot\; e^{-\mu t} \;\approx\; i\,\nu\,t\,,$$

which shows that $a_{i,\,i-1} \;= i\,\nu$.

Similarly to the previous solution, one can see that if $|\,j-i\,| > 2$, then $a_{i,j} = \;0$. ☐

3. Absolute Distributions

We shall derive a differential equation for the absolute distribution, too. Let the distribution of ξ_t be denoted by

$$D(t) = \Big(d_0(t), d_1(t), d_2(t), \ldots \Big),$$

that is, $d_j(t) = \Pr(\xi_t = j)$. Since the total probability formula yields that

$$d_j(t+h) = \sum_k d_k(t) \cdot p_{kj}(h),$$

the vector $D(t)$, obviously satisfies the equation

$$D(t+h) = D(t) \cdot P(h).$$

Taking formally the derivative with respect to h, we get that

$$D'(t+h) = D(t) \cdot P'(h).$$

Substituting h by 0, we get that

$$D'(t) = D(t) \cdot A.$$

This differential equation is called the **forward Kolmogorov – Feller equation for the absolute distribution vector**. This is a system of an infinite number of differential equations for the elements of the absolute distribution:

$$d_j'(t) = \sum_k d_k(t) \cdot a_{kj} \qquad (j = 0, 1, 2, \ldots).$$

The differential equation can be extended by the initial condition:

$$D(0) = D,$$

where D is the initial distribution of the process, that is,

$$d_j(0) = d_j \qquad (j = 0, 1, 2, \ldots).$$

5. Example: BACTERIA SPLITTING (continued, see Section 2)

a) Assuming that initially there is only one bacterium, prove that the distribution of ξ_t is geometrical distribution with parameter $e^{-\mu t}$, that is,

$$d_j(t) = \Pr(\xi_t = j) = e^{-\mu t}(1 - e^{-\mu t})^{j-1} \qquad (j = 1, 2, \ldots).$$

b) What is the distribution of ξ_t, if initially there are N bacteria? (Hint: think of the sum of independent, geometrically distributed random variables.)

Solution.

a) The forward Kolmogorov – Feller equation for the absolute distribution vector of this process is

$$[\, d_1'(t)\,,\;\; d_2'(t)\,,\;\; d_3'(t)\,,\;\; \dots \,] \;=\; [\, d_1(t)\,,\;\; d_2(t)\,,\;\; d_3(t)\,,\;\; \dots \,] \cdot \begin{bmatrix} -\mu & \mu & & & & \\ & -2\mu & 2\mu & & & \\ & & -3\mu & 3\mu & & \\ & & & -4\mu & 4\mu & \\ & & & & \dots & \dots \end{bmatrix},$$

which means the following system of differential equations:

$$\begin{aligned} d_1'(t) &= -\mu \cdot d_1(t)\,, \\ d_2'(t) &= \mu \cdot d_1(t) - 2\,\mu \cdot d_2(t)\,, \\ d_3'(t) &= 2\mu \cdot d_2(t) - 3\,\mu\, d_3(t)\,, \\ &\dots \end{aligned}$$

that is,

$$d_1'(t) = -\mu \cdot d_1(t)\,,$$

and

$$d_j'(t) = (j-1)\,\mu \cdot d_{j-1}(t) - j\,\mu\, d_j(t) \qquad (\, j = 2\,,3\,,\dots\,)\,.$$

Since initially there is only 1 bacterium, the initial condition for the system is

$$d_1(0) = 1\,, \qquad d_2(0) = 0\,, \qquad d_3(0) = 0\,, \qquad \dots \;.$$

This system, which consist of linear differential equations, is solvable step-by-step. The first equation contains only $d_1(t)$, so one can find its solution. As soon as $d_1(t)$ is given, the only unknown in the second equation is $d_2(t)$, so one can determine it. With $d_2(t)$ given, one can solve the third equation for $d_3(t)$, and so on. Instead of actually performing this step-by-step, which fits more a differential equation course, than this one, we declare and the reader is asked to check that the formulas

$$d_j(t) = e^{-\mu t}\,(1 - e^{-\mu t})^{j-1} \qquad (\, j = 1\,,2\,,\dots\,)$$

really solve both the differential equations and the initial conditions.

b) We can interpret the population generated by N bacteria as the union of N populations generated by the initial bacteria independently of each other. Since the sum of independent, geometrically distributed random variables is distributed according to negative binomial distribution, the answer to the question is negative binomial distribution of order N with parameter $e^{-\mu t}$. □

4. Stationarity

The process is stationary if $D(t)$ does not depend on t, that is, $D'(t) = 0$. The constant value of $D(t)$ may only be the initial distribution D, that is, $D(t) = D$. Since $D(t)$ satisfies the differential equation

$$D'(t) = D(t) \cdot A\,,$$

the stationary absolute distribution satisfies the equation

$$0 = D \cdot A ,$$

where 0 here means the zero row vector. This equation represents a system of an infinite number of homogeneous linear equations:

$$0 = \sum_k d_k \cdot a_{kj} \qquad (j = 0, 1, 2, \dots) .$$

The system can be extended by the equation

$$\sum_k d_k = 1 ,$$

expressing the fact that the initial distribution is normalized.

6. Example: STATIONARY DISTRIBUTION FOR BIRTH AND DEATH PROCESSES

Prove that the stationary absolute distribution $(d_0 , d_1 , d_2 , \dots)$ of a birth and death process with the infinitesimal generator

	0	1	2	3	...
0	$*$	μ_0			
1	ν_1	$*$	μ_1		
2		ν_2	$*$	μ_2	
3			ν_3	$*$	
...					

satisfies the relations $d_0 \mu_0 = d_1 \nu_1 , \quad d_1 \mu_1 = d_2 \nu_2 , \quad d_2 \mu_2 = d_3 \nu_3 , \quad \dots$.

Solution.

$0 = d \cdot A$ gives the system of equations

$$\begin{aligned} 0 &= -d_0 \mu_0 + d_1 \nu_1 \\ 0 &= d_0 \mu_0 - d_1 (\nu_1 + \mu_1) + d_2 \nu_2 \\ 0 &= d_1 \mu_1 - d_2 (\nu_2 + \mu_2) + d_3 \nu_3 \\ &\dots \end{aligned}$$

The first equation yields that

$$d_0 \mu_0 = d_1 \nu_1 .$$

Opening the parenthesis in the second, and then using that $d_0 \mu_0 = d_1 \nu_1$, we get that

$$d_1 \mu_1 = d_2 \nu_2 .$$

Opening the parenthesis in the third, and then using that $d_1 \mu_1 = d_2 \nu_2$, we get that

$$d_2 \mu_2 = d_3 \nu_3 .$$

And so on, we get the desired equalities one after the other..

Remark.

Note that if the sum

$$1 + \frac{\mu_0}{\nu_1} + \frac{\mu_0}{\nu_1}\frac{\mu_1}{\nu_2} + \frac{\mu_0}{\nu_1}\frac{\mu_1}{\nu_2}\frac{\mu_2}{\nu_3} + \dots$$

is finite, then

$$d_1 = \frac{\mu_0}{\nu_1} \cdot d_0 \quad , \quad d_2 = \frac{\mu_0}{\nu_1}\frac{\mu_1}{\nu_2} \cdot d_0 \quad , \quad d_3 = \frac{\mu_0}{\nu_1}\frac{\mu_1}{\nu_2}\frac{\mu_2}{\nu_3} \cdot d_0 \quad , \quad \dots ,$$

where

$$d_0 = \left(1 + \frac{\mu_0}{\nu_1} + \frac{\mu_0}{\nu_1}\frac{\mu_1}{\nu_2} + \frac{\mu_0}{\nu_1}\frac{\mu_1}{\nu_2}\frac{\mu_2}{\nu_3} + \dots \right)^{-1} .$$ □

7. Example: QUEUEING WITH ONE SERVER, STATIONARY DISTRIBUTION (continued, see Section 2)

Prove that, if $\mu < \nu$, then the geometrical distribution on the set of non-negative integers with parameter $\frac{\nu-\mu}{\nu}$ is a stationary absolute distribution for the process.

Solution.

For the process of queueing with one server $\mu_j = \mu$ and $\nu_j = \nu$ for all j , thus, according to the previous remark

$$d_1 = \left(\frac{\mu}{\nu} \right) \cdot d_0 \quad , \quad d_2 = \left(\frac{\mu}{\nu} \right)^2 \cdot d_0 \quad , \quad d_3 = \left(\frac{\mu}{\nu} \right)^3 d_0 \quad , \quad \dots ,$$

which means that the stationary absolute distribution constitutes a geometrical sequence with quotient $\frac{\mu}{\nu}$. Because of the assumption $\mu < \nu$, the sum of the terms in this sequence is finite. Since the quotient of the geometrical sequence is $\frac{\mu}{\nu}$, the stationary distribution is geometrical distribution with parameter $1 - \frac{\mu}{\nu} = \frac{\nu-\mu}{\nu}$ on the set of non-negative integers.

Solving the equation $1 - p = \frac{\mu}{\nu}$ for p , we get that the parameter of the geometrical distribution is $\frac{\nu-\mu}{\nu}$.□

5. Equation for the Expected Value

Let the expected value of ξ_t be denoted by $m(t)$:

$$m(t) = d_0(t)\cdot 0 + d_1(t)\cdot 1 + d_2(t)\cdot 2 + \dots = D(t)\cdot v_1 ,$$

where v_1 is the vector

$$v_1 = \begin{bmatrix} 0 \\ 1 \\ 2 \\ \vdots \end{bmatrix} .$$

Formal differentiation yields that

$$m'(t) = D'(t)\cdot v_1 ,$$

that is,

$$m'(t) = D(t)\cdot A\cdot v_1 .$$

If the infinitesimal generator A of the process is so special that the vector $A\cdot v_1$ turns out to be equal to a linear combination of the vectors v_1 and

$$v_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \end{bmatrix} ,$$

that is,

$$A\cdot v_1 = c_0\cdot v_0 + c_1\cdot v_1 ,$$

where c_0 and c_1 are constant numbers, then

$$m'(t) = D(t)\cdot A\cdot v_1 = D(t)\cdot(c_0\cdot v_0 + c_1\cdot v_1) =$$

$$= c_0\cdot D(t)\cdot v_0 + c_1\cdot D(t)\cdot v_1 = c_0 + c_1\cdot m(t) .$$

We used here that $D(t)\cdot v_0 = 1$, because the distribution $D(t)$ is normalized. The linear differential equation

$$m'(t) = c_0 + c_1\cdot m(t)$$

with constant coefficients can be extended by the initial condition

$$m(0) = m_0 ,$$

where m_0 is the expected value of the initial distribution. The solution to this differential equation with the initial condition, as it is known from a differential equation course, is

$$m(t) = m_0 e^{c_1 t} + \frac{c_0}{c_1}(e^{c_1 t} - 1).$$

8. Example: TELEPHONE CENTER WITH INFINITE CAPACITY (continued, see Section 2)

Prove that the expected value function of the process with the initial condition $m(0) = N$ is

$$m(t) = \frac{\mu}{\nu}(1 - e^{-\nu t}) + N e^{-\nu t}.$$

Solution.

Let us examine whether for the infinitesimal generator A of this process the vector $A \cdot v_1$ is equal to a linear expression of the vectors v_0 and v_1 :

$$A \cdot v_1 = \begin{bmatrix} -\mu & \mu & & & & \\ \nu & -(\mu+\nu) & \mu & & & \\ & 2\nu & -(\mu+2\nu) & \mu & & \\ & & 3\nu & -(\mu+3\nu) & \mu & \\ & & & & \cdots & \cdots \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \\ \vdots \end{bmatrix}$$

$$= \begin{bmatrix} \mu \\ \mu-\nu \\ \mu-2\nu \\ \mu-3\nu \\ \vdots \end{bmatrix} = \mu \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ \vdots \end{bmatrix} - \nu \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \\ \vdots \end{bmatrix} = \mu \cdot v_0 - \nu \cdot v_1 ,$$

that is, $c_0 = \mu$ and $c_1 = -\nu$. Simple replacement yields the given formula for $m(t)$.

9. Exercise.

Prove that the stationary absolute distribution of the process is Poisson distribution with parameter $\frac{\mu}{\nu}$. Can you give an explanation why it is natural that, when $t \to \infty$, then the limit of the function $m(t)$ is the same as the parameter of the stationary absolute distribution? □

6. PROBLEMS

1. GENERATING A JUMP PROCESS

Assume that you have a calculator which generates random numbers uniformly distributed between 0 and 1 . Using a calculator, give the physical interpretation of the jump process whose infinitesimal generator is

$$A = \begin{bmatrix} * & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & * & \frac{1}{4} \\ \frac{3}{4} & \frac{1}{4} & * \end{bmatrix} .$$

2. QUEUEING WITH ONE SERVER (continued, see Sections 2 and 4)

Assume that customers arrive to a teller's window in a bank at random time instants according to μ -parametric homogeneous Poisson process. If the window is busy, they line up in a queue. Assume that the service time of each customer is exponentially distributed with parameter ν , independently of each other and the arrival process.

a) If $\mu < \nu$, then what is the probability that, at a large time instant, a customer arriving to the window, does not have to wait in queue?

b) What is the distribution of the waiting time of a customer who enters the queue when there are j customers in the queue in front of him.

3. HOW MUCH TIME TO WAIT IN A STATIONARY QUEUE WITH ONE SERVER?

Assume that a customer arrives to a teller's window in a bank when the distribution of the length of the queue is the stationary distribution. Observe the amount of time the customer has to wait until he has been served and can leave the shop. Prove that

a) its expected value is $\frac{1}{\nu-\mu}$,

b) its distribution is exponential distribution with parameter $\nu - \mu$.

Hint: The number of people in the queue, our customer included, has a geometrical distribution with parameter $\frac{\nu-\mu}{\nu}$. The expected value of each serving time has an exponential distribution with parameter ν . In part a) use the total expected value theorem. In part b) use the discrete-continuous version of the total probability formula for the geometrical and gamma distributions.

4. TRANSITION PROBABILITIES FOR THE POISSON-PROCESS

Find an explicate formula of the transition probability $p_{ij}(t)$ for the λ -parametric Poisson-process.

5. TWO-STATE JUMP PROCESS WITH EQUAL PARAMETERS

Consider a jump process that has only two states. The parameter of both resting times is λ . Find an explicit formula for the transition matrix $P(t)$.

Hint: Notice that, for a Poisson process, the probability of an even or odd number of occurrences during a time interval can be given with a sum which can be given by a closed formula.

6. TWO-STATE JUMP PROCESS WITH DIFFERENT PARAMETERS

Consider a jump process that has only two states. The parameters of its resting times are μ and ν , respectively.

a) Determine the infinitesimal generator, then find an explicit formula for the transition matrix $P(t)$.

b) Accepting an initial distribution find an explicit formula for the absolute distribution $D(t)$. Find the initial distribution for which the process is stationary.

c) Find the limit of $P(t)$ and $D(t)$ when $t \to \infty$.

7. STATIONARY DISTRIBUTION OF A THREE-STATE JUMP PROCESS

Consider a jump process that has only three states. The parameters of its resting times are 1 , 2 , 3 , respectively. From state 1 a jump is possible only into state 3 , from state 2 , jumps to state 1 and state 2 are equally probable, from state 3 a jump is possible only to states 1 and 2 , to state 2 is twice as much probable as to state 1 . Find the stationary absolute distribution of the process. Solution: (0.3; 0.4; 0.3) .

8. QUEUEING WITH A MAXIMALIZED QUEUE LENGTH

Assume that customers arrive to a teller's window in a bank at random time instants according to μ -parametric homogeneous Poisson process. If the window is busy, they line up in a queue. The maximum length of the queue is N , that is, if the queue length is N and a customer arrives, then he or she goes away, and the length of the queue does not increase. Assume that the service time of each customer is exponentially distributed with parameter ν , independently of each other and the arrival process. Let ξ_t be the length of the queue.

a) Prove that the infinitesimal generator of the process is

	0	1	2	3	4	...	N
0	*	μ					
1	ν	*	μ				
2		ν	*	μ			
3			ν	*	μ		
4				ν	*	μ	
...						...	
N						ν	*

b) Find the stationary absolute distribution of the process.

c) What is the probability that a person who wants to enter the queue at a large time instant is driven away because the queue is full?

9. QUEUEING WITH MORE SERVERS

Assume that customers arrive at random time instants according to μ -parametric homogeneous Poisson process to a bank, and form one queue, if necessary, for two windows. Assume that the service time of each customer is exponentially distributed with parameter ν , independently of each other and the arrival process. Let ξ_t be the number of customers in the bank.

a) Find the infinitesimal generator of the process.

b) Modify the process by assuming that there are K windows, and find the infinitesimal generator of the modified process.

c) Add the restriction that the maximal number of customers in the bank may be N $(N > K)$. If there are N customers, then the arriving new customers go away. Find the infinitesimal generator of this process, too.

10. QUEUEING WITH AN UNLIMITED NUMBER OF SERVERS (continued)

Modify the previous process by assuming that there are infinitely many servers (i.e. windows), and find the infinitesimal generator under these circumstances.

11. TELEPHONE CENTER WITH FINITE CAPACITY

Assume that telephone talks at a telephone center start at random time instants according to μ - parametric homogeneous Poisson process. Assume that capacity of the center is N, which means that is all the N lines are busy, then any more talks are lost. Assume that the length of each talk is exponentially distributed with parameter ν, independently of each other and the arrival process. Let ξ_t be the number of the busy lines.

a) Prove that the infinitesimal generator of the process is

	0	1	2	3	4	...	N
0	*	μ					
1	ν	*	μ				
2		2ν	*	μ			
3			3ν	*	μ		
4				4ν	*	μ	
...						...	
N						$N\nu$	*

b) Find the stationary absolute distribution of the process.

c) What is the probability that a talk needed at a large time instant is driven away because the queue is full?

12. BACTERIA SPLITTING OR DYING, WITH IMMIGRATION

Assume that any member of a bacterium population, independently of the others, dies after its birth at a random time which is exponentially distributed with parameter ν. When a bacterium dies, then independently of others, with probability p, it simply disappears from the world of living creatures, or with probability q, splitting into to gives a birth to two newborn ones $(p + q = 1)$. Moreover, at random time instants according to a homogeneous Poisson distribution with parameter μ additional external bacteria join the population. Let ξ_t mean the number of bacteria at time t.

a) Prove that the infinitesimal generator of the process is

	0	1	2	3	4	
0	*	μ				
1	$p\nu$	*	$q\nu + \mu$			$\cdots\cdots$
2		$2p\nu$	*	$2q\nu + \mu$		
3			$3p\nu$	*	$3q\nu + \mu$	
$\cdots$				$\cdots$	$\cdots$	

b) Prove that the expected value function $m(t)$ satisfies the differential equation

$$m'(t) = \mu + (q - p)\nu \cdot m(t) .$$

c) Prove that the solution of this differential equation with the initial condition $m(0) = N$ is

$$m(t) = \mu t + N \qquad \text{if } p = q ,$$

$$m(t) = \frac{\mu}{(q-p)\nu}\left(e^{(q-p)\nu t} - 1\right) + N e^{(q-p)\nu t} \qquad \text{if } p \neq q .$$

d) Find a differential equation for the second moment function $m_2(t)$.

Hint:

Let the 2nd moment of ξ_t be denoted by $m_2(t)$:

$$m_2(t) = d_0 \cdot 0^2 + d_1 \cdot 1^2 + d_2 \cdot 2^2 + \ldots = D(t) \cdot v_2 ,$$

where

$$v_2 = \begin{bmatrix} 0^2 \\ 1^2 \\ 2^2 \\ \vdots \end{bmatrix} .$$

Formal differentiation yields that

$$m_2'(t) = D'(t) \cdot v_2 = D(t) \cdot A \cdot v_2 .$$

Show that the vector Av_2 is equal to a linear combination of the vectors v_0 , v_1 , v_2 , that is,

$$A \cdot v_2 = c_{20} v_0 + c_{21} \cdot v_1 + c_{22} \cdot v_2 ,$$

where c_{20} , c_{21} , c_{22} are constant numbers which we can easily determine. Then use the same ideas as in Section 5 to derive the linear differential equation

$$m_2'(t) = c_{20} + c_{21} \cdot m_1(t) + c_{22} \cdot m_2(t) ,$$

where $m_1(t) = m(t)$ is the expected value function of the process.

13. MACHINES TO BE REPAIRED

Assume there are 10 machines and 3 repairmen at a work place. Each machine works continuously until it breaks. The life time of each machine is exponential with parameter 7 . As soon as a repairmen is available, then he starts the reparation, which lasts for a random time length having exponential distribution with parameter 13 . Consider the number of machines working as a random function of time.

a) Prove that the process is a birth and death process.

b) Find the infinitesimal generator of the process.

c) Find the stationary absolute distribution of the process.

14. ABSORPTION

Assume that any member of a bacterium population, after its birth at a random time which is exponentially distributed with parameter μ , independently of the others, dies with probability p or splits into two with probability q $(p+q=1)$. Let ξ_t mean the number of bacteria at time t .

a) Prove that the infinitesimal generator of the process is

	0	1	2	3	4	
0	*	0				
1	$p\mu$	*	$q\mu$			
2		$2p\mu$	*	$2q\mu$		......
3			$3p\mu$	*	$3q\mu$	
...				...	...	

b) Realize that the state 0 is an absorbing state.

c) Define a discrete time Markov chain: let its initial state and let its n -th jump be the same as the n -th jump of the continuous time process. What is the one step transition matrix of the discrete process?

d) Realize that the state 0 is an absorbing state for the discrete time process, too, and for any initial condition, the probability of the absorption is the same for the two processes.

e) Compare the problem of absorption to the bankruptcy problem when one of the players has infinitely much money. Using the result of the bankruptcy problem, find the probability of absorption on condition that i is the initial state ($i>0$) .