1.1. Introduction to stochastic risk analysis

1.1.1. About the subject

The concept of risk is diverse enough and is used in many areas of human activity. The object of interest in this book is the theory of collective risk. Swedish mathematicians Cramér and Lundberg established stochastic models of insurance based on this theory.

Stochastic risk analysis is a rather broad name for this volume. We will consider mathematical problems concerning the Cramér-Lundberg insurance model and some of its generalizations. The feature of this model is a random process, representing the dynamics of the capital of a company. These dynamics consists of alternations of slow accumulation (that may be not monotonous, but continuous) and fast waste with the characteristic of negative jumps.

All mathematical studies on the given subject continue to be relevant nowadays thanks to the absence of a compact analytical description of such a process. The stochastic analysis of risks which is the subject of interest has special aspects. For a long time, the most interesting problem within the framework of the considered model was ruin, which is understood as the capital of a company reaching a certain low level. Such problems are usually more difficult than those of the value of process at fixed times.
1.1.2. About the ruin model

Let us consider the dynamics of the capital of an insurance company. It is supposed that the company serves several clients, which bring in insurance premiums, i.e. regular payments, filling up the cash desk of the insurance company. Insurance premiums are intended to compensate company losses resulting from single payments of great sums on claims of clients at unexpected incident times (the so-called insured events). They also compensate expenditures on maintenance, which are required for the normal operation of a company. The insurance company’s activity is characterized by a random process which, as a rule, is not stationary. The company begins business with some initial capital. The majority of such undertakings come to ruin and only a few of them prosper. Usually they are the richest from the very beginning. Such statistical regularities can already be found in elementary mathematical models of dynamics of insurance capital.

The elementary mathematical model of dynamics of capital, the Cramér-Lundberg model, is constructed as follows. It uses a random process \( R_t \) \((t \geq 0)\)

\[
R_t = u + pt - \sum_{n=1}^{N_t} U_n, \tag{1.1}
\]

where \( u \geq 0 \) is the initial capital of the company, \( p > 0 \) is the growth rate of an insurance premium and \( pt \) is the insurance premium at time \( t \). \((U_n)_{n=1}^\infty\) is a sequence of suit sizes which the insurance company must pay immediately. It is a sequence of independent and identically distributed (i.i.d.) positive random variables. We will denote a cumulative distribution function of \( U_1 \) (i.e. of all remaining) as \( B(x) \equiv P(U_1 \leq x) \) \((x \geq 0)\). The function \((N_t) \ (t \geq 0)\) is a homogeneous Poisson process, independent of the sequence of suit sizes, having time moments of discontinuity at points \((\sigma_n)_{n=1}^\infty\). Here, \( 0 \equiv \sigma_0 < \sigma_1 < \sigma_2 < \ldots \); values \( T_n = \sigma_n - \sigma_{n-1} \) \((n \geq 1)\) are i.i.d. random variables with a common exponential distribution with a certain parameter \( \beta > 0 \).
Figure 1.1 shows the characteristics of the trajectories of the process.

This is a homogeneous process with independent increments (hence, it is a homogeneous Markov process). Furthermore, we will assume that process trajectories are continuous from the right at any point of discontinuity.

Let $\tau_0$ be a moment of ruin of the company. This means that at this moment, the company reaches into the negative half-plane for the first time (see Figure 1.1). If this event does not occur, this moment is set as equal to infinity.

The first non-trivial mathematical results in risk theory were connected with the function:

$$\psi(u) = P_u(\tau_0 < \infty) \quad (u \geq 0),$$

i.e. a probability of ruin on an infinite interval for a process with the initial value $u$. Interest is also represented by the function $\psi(u, t) = P_u(\tau_0 \leq t)$. It is called the ruin function on “finite horizon”.

Nowadays many interesting outcomes have been reported for the Cramér-Lundberg model and its generalizations. In this volume, the basic results of such models are presented. In addition, we consider its
generalizations, such as insurance premium inflow and distribution of suit sizes.

This is concentrated on the mathematical aspects of a problem. Full proofs (within reason) of all formulas, and volume theorems of the basic course are presented. They are based on the results of probability theory which are assumed to be known. Some of the information on probability theory is shortly presented at the start. In the last chapter some management problems in insurance business are considered.

1.2. Basic methods

1.2.1. Some concepts of probability theory

1.2.1.1. Random variables

The basis of construction of probability models is an abstract probability space \((\Omega, \mathcal{F}, P)\), where \(\Omega\) is a set of elementary events; \(\mathcal{F}\) is a sigma-algebra of subsets of the set \(\Omega\), representing the set of those random events, for which it makes sense to define the probability within the given problem; \(P\) is a probability measure on set \(\Omega\), i.e. non-negative denumerably additive function on \(\mathcal{F}\). For any event \(A \in \mathcal{F}\), the probability, \(P(A)\), satisfies the condition \(0 \leq P(A) \leq 1\). For any sequence of non-overlapping sets \((A_n)_{n=1}^{\infty} (A_n \in \mathcal{F})\) the following equality holds:

\[
P \left( \bigcup_{n=1}^{\infty} A_n \right) = \sum_{n=1}^{\infty} P(A_n),
\]

and \(P(\Omega) = 1\). Random events \(A_1\) and \(A_2\) are called independent if \(P(A_1, A_2) \equiv P(A_1 \cap A_2) = P(A_1)P(A_2)\). This definition is generalized on any final number of events. Events of infinite system of random events are called mutually independent if any of its final subsystem consists of independent events.

A random variable is a measurable function \(\xi(\omega) (\omega \in \Omega)\) with real values. It means that for any real \(x\), the set \(\{\omega : \xi(\omega) \leq x\}\) is a random
event and hence, probability of it exists, designated as $F_\xi(x)$. Thus, the cumulative distribution function, $F_\xi$, is defined as follows:

$$F_\xi(x) = P(\xi \leq x) \quad (-\infty < x < \infty).$$

It is obvious that this function does not decrease when $x$ increases. In this volume, we will deal with absolutely continuous distributions and discreet distributions (sometimes with their mixtures).

For an absolutely continuous distribution, there exists its distribution density $f_\xi(x) = dF_\xi(x)/dx$ for all $x \in (-\infty, \infty)$ such that

$$\int_{-\infty}^{\infty} f_\xi(x) \, dx = 1.$$

For discreet distributions, there exists a sequence of points (atoms) $(x_n)_{n=1}^{\infty}$ for which non-negative probabilities $p(x_n) = P(\xi = x_n)$ are defined as:

$$\sum_{n=1}^{\infty} p(x_n) = 1.$$

The random variable is called integer if it has a discreet distribution with atoms in the integer points of a numerical axis, denoted by $\mathbb{Z}$.

If $\mathbb{R}$ is the set of all real numbers, $\varphi$ is a measurable function on $\mathbb{R}$, and $\xi$ is a random variable, then superposition $\psi(\omega) \equiv \varphi(\xi(\omega))$ ($\omega \in \Omega$) is a random variable too. Various compositions of random variables are possible, which are also random variables. Two random variables $\xi_1$ and $\xi_2$ are called independent, if for any $x_1$ and $x_2$ events $\{\xi_1 \leq x_1\}$ and $\{\xi_2 \leq x_2\}$ are independent.

Expectation (average) $E\xi$ of a random variable $\xi$ is the integral of this function on $\Omega$ with respect to the probability measure $P$, i.e.:

$$E\xi = \int_{\Omega} \xi(\omega) \, P(d\omega) \equiv \int \xi \, dP$$
(an integral of Lebesgue). By a cumulative distribution function, this integral can be noted as an integral of Stieltjes:

\[ E\xi = \int_{-\infty}^{\infty} x \, dF\xi(x), \]

and for a random variable \( \xi \) with absolute continuous distribution, it can be represented as integral of Riemann:

\[ E\xi = \int_{-\infty}^{\infty} x f\xi(x) \, dx. \]

For a random variable \( \xi \) with a discreet distribution, it is possible to write an integral in the form of the sum:

\[ E\xi = \sum_{n=1}^{\infty} x_n p(x_n). \]

When evaluating an expectation, it is necessary to be careful in case the integral from the module of this random variable is equal to infinity. Sometimes it useful to distinguish three cases: an integral equal to plus infinity, an integral equal to minus infinity and an integral does not exist.

Let us note that it is possible to consider separately a cumulative distribution function out of connection with random variables generating them and probability spaces. However, for any non-decreasing, continuous from the right, function \( F \) such that \( F(x) \to 0 \) as \( x \to -\infty \) and \( F(x) \to 1 \) as \( x \to \infty \) (the cumulative distribution function of any random variable possesses these properties), it is possible to construct a probability space and with random variable on this space, which has \( F \) as its cumulative distribution function on this probability space. Therefore, speaking about a cumulative distribution function, we will always mean some random variable within this distribution. It allows us to use equivalent expressions such as “distribution moment”, “moment of a random variable”, “generating function of a distribution” and “generating function of a random variable”.
The following definitions are frequently used in probability theory. The moment of \( n \)th order of a random variable \( \xi \) is an integral \( E\xi^n \) (if it exists). The central moment of \( n \)th order of a random variable \( \xi \) is an integral \( E(\xi - E\xi)^n \) (if it exists). The variance (dispersion) \( D\xi \) of a random variable \( \xi \) is its central moment of second order.

The generating function of a random variable is the integral \( E\exp(\alpha\xi) \), considered as a function of \( \alpha \). Interest represents those generating functions which are finite for all \( \alpha \) in the neighborhood of zero. In this case, there is one-to-one correspondence between the set of distributions and the set of generating functions. This function has received the name because of its property “to make” the moments under the formula:

\[
E\xi^n = \frac{d^n E\exp(\alpha\xi)}{d\alpha^n} \bigg|_{\alpha=0}.
\]

A random \( n \)-dimensional vector is the ordered set of \( n \) random variables \( \xi = (\xi_1, \ldots, \xi_n) \). Distribution of this random vector (joint distribution of its random coordinates) is a probability measure on space \( \mathbb{R}^n \), defined by \( n \)-dimensional cumulative distribution function:

\[
F_\xi(x_1, \ldots, x_n) = P(\xi_1 \leq x_1, \ldots, \xi_n \leq x_n) \quad (x_i \in \mathbb{R}, i = 1, \ldots, n).
\]

As the generating function of a random vector is called function of \( n \) variables \( E\exp(\alpha, \xi) \), where \( \alpha = (\alpha_1, \ldots, \alpha_n) \) (\( \alpha_i \in \mathbb{R} \)) and \( (\alpha, \xi) = \sum_{i=1}^{n} \alpha_i \xi_i \). The mixed moment of order \( m \geq 2 \) of a random vector \( \xi \) is called \( E(\xi_1^{m_1} \cdots \xi_n^{m_n}) \), where \( m_i \geq 0, \sum_{i=1}^{n} m_i = m \). Covariance of random variables \( \xi_1 \) and \( \xi_2 \) is called central joint moment of the second order:

\[
\text{cov}(\xi_1, \xi_2) = E(\xi_1 - E\xi_1)(\xi_2 - E\xi_2).
\]

1.2.1.2. Random processes

In classical probability theory, random process on an interval \( T \subset \mathbb{R} \) is called a set of random variables \( \xi = (\xi_t)_{t \in T} \), i.e. function of two
arguments \((t, \omega)\) with values \(\xi_t(\omega) \in \mathbb{R} \ (t \in \mathbb{R}, \ \omega \in \Omega)\), satisfying measurability conditions. As random process, we can understand that an infinite-dimensional random vector, whose space is designated as \(\mathbb{R}^T\), is a set of all functions on an interval \(T\). Usually, it is assumed that a sigma-algebra of subsets of such set functions contains all so-called finite-dimensional cylindrical sets, i.e. sets of:

\[
\{ f \in \mathbb{R}^T : f_{t_1} \in A_1, \ldots, f_{t_n} \in A_n \} \quad (n \geq 1, t_i \in T, A_i \in \mathcal{B}(\mathbb{R}))
\]

where \(\mathcal{B}(\mathbb{R})\) is the Borel sigma-algebra of the subsets of \(\mathbb{R}\) (the sigma-algebra of subsets generated by all open intervals of a numerical straight line). For the problems connected with the first exit times, the minimal sigma-algebra \(\tilde{\mathcal{F}}\), containing all such cylindrical sets, is not sufficient. It is connected by that the set \(\mathbb{R}^T\) “is too great”. Functions belonging in this set are not connected by any relations considering an affinity of arguments \(t\), such as a continuity or one-sided continuity.

For practical problems, it is preferable to use the other definition of the random process, namely not a set of random variables assuming the existence of the abstract probability spaces, but a random function as element of a certain set \(\Omega\), composed of all possible realizations within the given circle of problems. On this function space, a sigma-algebra of subsets and a probability measure on this sigma-algebra should be defined. For the majority of practical uses, it is enough to take as function space the set \(\mathcal{D}\) of all functions \(\xi : T \to \mathbb{R}\) continuous from the right and having a limit from the left at any point of an interval \(T \subset \mathbb{R}\). The set \(\mathcal{D}\) is a metric space with respect to the Skorokhod metric, which is a generalization of the uniform metric. A narrower set, that has numerous applications as a model of real processes, is the set \(\mathcal{C}\) of all continuous functions on \(T\) with locally uniform metric. In some cases it is useful to consider other subsets of space \(\mathcal{D}\), for example, all piece-wise constant function having a locally finite set of point of discontinuities. Sigma-algebra \(\mathcal{F}\) of subsets of \(\mathcal{D}\), generated by cylindrical sets with the one-dimensional foundation of an aspect \(\{ \xi \in \mathcal{D} : \xi(t) \in A \} \quad (t \in T, A \in \mathcal{B}(\mathbb{R}))\) that comprises all interesting subsets (events) connected with the moments of the first exit from open intervals belonging to a range of values of process.
Random process is determined if some probability measure on corresponding sigma-algebra of subsets of set of its trajectories is determined. In classical theory of random processes, a probability measure on $\mathcal{F}$ is determined if and only if there exists a consistent system of finite-dimensional distributions determined on cylindrical sets with finite-dimensional foundations [KOL 36]. To represent a measure on the sigma-algebra $\mathcal{F}$, Kolmogorov’s conditions for the coordination of distributions on the finite-dimensional cylindrical sets are not enough. In this case, some additional conditions are required. They, as a rule, are concerned with two-dimensional distributions $P(\xi(t_1) \in A_1, \xi(t_2) \in A_2)$ as $|t_1 - t_2| \to 0$. In problems of risk theory where, basically, Markov processes are used, these additional conditions are easily checked.

1.2.1.3. Shift operator

We will assume further that $T = [0, \infty) \equiv \mathbb{R}_+$. First, we define on set $\mathcal{D}$ an operator $X_t$ “value of process in a point $t$”: $X_t(\xi) \equiv \xi(t)$. We also use other labels for this operator, containing the information on concrete process, for example, $R_t$, $N_t$ and $S_t$. They are operators with meaning: values of concrete processes at a point $t$. By an operator $X_t$ we will represent the set $\{\xi \in \mathcal{D} : \xi(t_1) \in A_1, \ldots, \xi(t_n) \in A_n\}$ as $\{X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n\}$. Thus, finite-dimensional distribution is possible to note as probability $P(X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n)$. This rule of denotation when the argument in the subset exposition is omitted is also spread on other operators defined on $\mathcal{D}$.

A shift operator $\theta_t$ maps $\mathcal{D}$ on $\mathcal{D}$. It is possible to define function $\theta_t(\xi)$ ($t \geq 0$) by its values at points $s \geq 0$. These values are defined as:

$$(\theta_t(\xi))(s) = \xi(t + s) \quad (t, s \geq 0).$$

Using an operator $X_t$ this relation can be noted in an aspect $X_s(\theta_t(\xi)) = X_{t+s}(\xi)$ or, by lowering argument $\xi$, in an aspect $X_s(\theta_t) = X_{t+s}$. We also denote this relation (superposition) as $X_s \circ \theta_t = X_{t+s}$. Obviously, $\theta_s \circ \theta_t = \theta_{t+s}$. 
An important place in the considered risk models is taken by the operator $\sigma_\Delta$ “the moment of the first exit from set $\Delta$”, defined as $\sigma_\Delta(\xi) = \inf \{t \geq 0 : \xi(t) \not\in \Delta\}$, if the set in braces is not empty; otherwise, we suppose $\sigma_\Delta(\xi) = \infty$.

1.2.1.4. Conditional probabilities and conditional averages

From elementary probability theory, the concept of conditional probabilities $P(A \mid B)$ and a conditional average $E(f \mid B)$ concerning event $B$ are well-known, where $A$ and $B$ are events, $f$ is a random variable and $P(B) > 0$. The concept of conditional probability concerning a final partition of space on simple events $P(A \mid \mathcal{P})$ is not more complicated, where $\mathcal{P} = (B_1, \ldots, B_n)$ ($B_i \cap B_j = \emptyset$, $\bigcup_{k=1}^n B_k = \Omega$) and $P(B_i) > 0$. In this case, the conditional probability can be understood as function on partition elements: on a partition element $B_i$, its value is $P(A \mid B_i)$. This function accepts $n$ values. However, in this case, there is a problem as to how to calculate conditional probabilities with respect to some association of elements of the partition. It means to receive a function with a finite (no more $2^n$) number of values, measurable with respect to the algebra of subsets generated by this finite partition. In this way, we can attempt to apply an infinite partition in the right part of the conditional probability. Obviously, this generalization is not possible for non-denumerable partition, for example, set of pre-images of function $X_t$, i.e. $(X_t^{-1}(x))_{x \in \mathbb{R}}$. In this case, conditional probability is accepted to define a function on $\mathbb{R}$ with special properties, contained in the considered example with a final partition. That is, the conditional probability $P(A \mid X_t)$ is defined as a function of $\xi \in \mathcal{D}$, measurable with respect to sigma-algebra, generated by all events $\{X_t < x\}$ (we denote such a sigma-algebra as $\sigma(X_t)$), which for any $B \in \mathcal{B}($\mathbb{R}$)$ satisfies the required conditions:

$$P(A, X_t \in B) = \int_{X_t \in B} P(A \mid X_t)(\xi) dP \equiv E(P(A \mid X_t); X_t \in B).$$

This integral can be rewritten in other form, while using representation of conditional probability in an aspect:

$$P(A \mid X_t) = g_A \circ X_t,$$
where $g_A$ is a measurable function on $\mathbb{R}$, defined uniquely according to the known theorem from a course on probability theory [NEV 64]. Then, using a change of variables $x = X_t(\xi)$, we obtain the following representation:

$$P(A, X_t \in B) = \int_B g_A(x) p_t(dx),$$

where $p_t(S) = P(X_t \in S)$ ($S \in \mathcal{B}(\mathbb{R})$). The value of function $g_A(x)$ can be designated as $P(A | X_t = x)$. This intuitively clear expression cannot be understood literally in the spirit of elementary probability theory. In certain cases, it can be justified as a limit of conditional probabilities, where the right side of conditional probability is changed with the condition that $X_t$ belongs to a small neighborhood of a point $x$. Usually, function $g_A(x)$ may be identified using the value of function $P_x(A)$, where $A \rightarrow P_x(A)$ is a measure on $\mathcal{F}$ for each $x \in \mathbb{R}$ and $x \rightarrow P_x(A)$ is a $\mathcal{B}(\mathbb{R})$-measurable function for each $A \in \mathcal{F}$. Hence,

$$g_A \circ X_t = P_{X_t}(A).$$

1.2.1.5. Filtration

To define the Markov process, it is necessary to define the concepts of “past” and “future” of the process, which means to define conditional probability and average “future” relative to “past”. For this purpose, together with a sigma-algebra $\mathcal{F}$, the ordered increasing family of sigma-algebras $(\mathcal{F}_t)$ ($t \geq 0$) is considered. This family is called filtration if $\lim_{t \to \infty} \mathcal{F}_t \equiv \bigcup_{t=0}^{\infty} \mathcal{F}_t$. For example, such a family consists of sigma-algebras $\mathcal{F}_t$. The latter is generated by all one-dimensional cylindrical sets $\{X_s < x\}$, where $s \leq t$ and $x \in \mathbb{R}$. It is designated as $\sigma(X_s : s \leq t)$, which is called natural filtration. The sigma-algebra $\mathcal{F}_t$ contains all measurable events reflecting the past of the process until the moment $t$. In relation to it, any value $X_{t+s}$ ($s > 0$) is reasonably called “future”.

Another feature of the considered example is a conditional probability (average) with respect to sigma-algebra $\mathcal{F}_t$. Under
conditional probability \( P(A \mid \mathcal{F}_t) \), it is understood that for such \( \mathcal{F}_t \)-measurable function (random variable) on \( \mathcal{D} \), for any \( B \in \mathcal{F}_t \) the equality is fulfilled:

\[
P(A, B) = \int_B P(A \mid \mathcal{F}_t)(\xi) \, dP \equiv E(P(A \mid \mathcal{F}_t); B).
\]

Conditional average \( E(f \mid \mathcal{F}_t) \) is similarly defined. For any random variable \( f \), the random variable \( E(f \mid \mathcal{F}_t) \) is \( \mathcal{F}_t \)-measurable function on \( \mathcal{D} \), for any \( B \in \mathcal{F}_t \) the equality is fulfilled:

\[
E(f; B) = \int_B E(f \mid \mathcal{F}_t)(\xi) \, dP \equiv E(E(f \mid \mathcal{F}_t); B).
\]

Let us note that the second definition is more general than the conditional probability of event \( A \) because it can be presented as a conditional average from an indicator function of the set \( A \). Let us note also that \( \Omega \in \mathcal{F}_t \) for any \( t \), and consequently

\[
Ef = E(f; \Omega) = E(E(f \mid \mathcal{F}_t); \Omega) = E(E(f \mid \mathcal{F}_t)).
\]

Existence and uniqueness (within set of a measure 0) of the conditional average is justified by the Radon-Nikodym theorem, which is one of the key theorems of the theory of measure [KOL 72].

### 1.2.1.6. Martingale

Random process \((X_t)_{t \geq 0}\), defined on a measurable space \((\mathcal{D}, \mathcal{F})\), supplied with filtration \((\mathcal{F}_t)_{\mathcal{F}_t \subset \mathcal{F}}\), is called martingale, if at any \( t \) value of process \( X_t \) measurable with respect to \( \mathcal{F}_t \), such that \( E|X_t| < \infty \) and at any \( s, t \geq 0 \) it is fulfilled \( E(X_{t+s} \mid \mathcal{F}_t) = X_t \) \( P \text{-a.s.} \). If for any \( s, t \geq 0 \) \( E(X_{t+s} \mid \mathcal{F}_t) \geq X_t \) \( P \text{-a.s.} \), then the process \( X(t) \) is called sub-martingale. Thus the martingale is a partial case of a sub-martingale. However, the martingale, unlike a sub-martingale, supposes many-dimensional generalizations. Some proofs of risk theory are based on the properties of martingales (sub-martingales).

Further, we will use the generalization of the sigma-algebra \( \mathcal{F}_t \) with a random \( t \) of special aspect, which depends on the filtration \((\mathcal{F}_t)\). We
consider a random variable \( \tau : \mathcal{D} \to \hat{\mathbb{R}}_+ \) such that for any \( t \geq 0 \), the event \( \{ \tau \leq t \} \) belongs to \( \mathcal{F}_t \). It is the Markov time. In this definition, \( \hat{\mathbb{R}}_+ \) denotes the enlarged positive half-line where the point “infinity” is supplemented. Therefore, we can admit infinity meanings for a Markov time. Let \( \tau \) be a Markov time. Then, we define a sigma-algebra:

\[
\mathcal{F}_\tau = \{ A \in \mathcal{F} : (\forall t > 0) A \cap \{ \tau \leq t \} \in \mathcal{F}_t \}.
\]

Intuitively, \( \mathcal{F}_\tau \) is a sigma-algebra of all events before the moment \( \tau \). Further, we will use the following properties of martingales (sub-martingales).

**Theorem 1.1.** (Theorem of Doob about Markov times) Let process \( (X_t) \) be a sub-martingale and \( \tau_1, \tau_2 \) be Markov times, for which \( E|X_{\tau_i}| < \infty \) \((i = 1, 2)\). Then, on set \( \{ \tau_1 \leq \tau_2 < \infty \} \)

\[
E(X_{\tau_2} | \mathcal{F}_{\tau_1}) \geq X_{\tau_1} \quad P\text{-a.s.}
\]

**Proof.** (see, for example, [LIP 86]).

Using evident property: if \( (X_t) \) is a martingale then \( (-X_t) \) is a martingale too, we receive a consequence: if \( (X_t) \) is a martingale, then on set \( \{ \tau_1 \leq \tau_2 < \infty \} \):

\[
E(X_{\tau_2} | \mathcal{F}_{\tau_1}) = X_{\tau_1} \quad P\text{-a.s.,}
\]

and for any finite Markov time \( E X_\tau = E X_0 \).

One of the most important properties of a martingale is the convergence of a martingale when its argument \( t \) tends to a limit. It is one of few processes for which such limit exists with probability 1.

**Theorem 1.2.** (Theorem of Doob about convergence of martingales). Let a process \( (X_t, \mathcal{F}_t) \) \((t \in [0, \infty))\) be a sub-martingale, for which \( \sup_{t \geq 0} E|X_t| < \infty \). Then, \( E|X_\infty| < \infty \) and with probability 1 there exists a limit:

\[
\lim_{t \to \infty} X_t = X_\infty.
\]
PROOF.— (see, for example, [LIP 86]).

It is clear that a martingale with the above properties satisfies the assertion of this theorem.

1.2.2. Markov processes

1.2.2.1. Definition of Markov process

Markov processes are defined in terms of the conditional probabilities (averages) considered above. The random process defined on measurable space \((\mathcal{D}, \mathcal{F})\), is called Markov, if for any \(t \geq 0, n \geq 1, s_i \geq 0, A_i \in \mathcal{B}(\mathbb{R}) (i = 1, \ldots, n)\) and \(B \in \mathcal{F}_t\) is fulfilled

\[
P\left(\bigcap_{k=1}^{n} X_{t+s_k} \in A_k, B\right) = E\left(P\left(\bigcap_{k=1}^{n} X_{t+s_k} \in A_k \mid X_t\right) ; B\right).
\]

[1.2]

Using definition of conditional probability, it follows:

\[
P\left(\bigcap_{k=1}^{n} X_{t+s_k} \in A_k, B\right) = E\left(P\left(\bigcap_{k=1}^{n} X_{t+s_k} \in A_k \mid \mathcal{F}_t\right) ; B\right).
\]

Because \(\sigma(X_t) \subset \mathcal{F}_t\) and \(B\) is an arbitrary set in \(\mathcal{F}_t\), it follows that for any \(t \geq 0, n \geq 1, s_i \geq 0, A_i \in \mathcal{B}(\mathbb{R}) (i = 1, \ldots, n)\) is fulfilled

\[
P\left(\bigcap_{k=1}^{n} (X_{t+s_k} \in A_k) \mid \mathcal{F}_t\right) = P\left(\bigcap_{k=1}^{n} (X_{t+s_k} \in A_k) \mid X_t\right)
\]

\(P\)-a.s. (almost sure, i.e. the set where these functions differ as \(P\)-measures zero).

A well-known Markov property: the conditional distribution of “future” at the fixed “past” depends only on the “present”.
Let us note that the shift operator \( \theta_t \), defined on set of trajectories, defines an inverse operator \( \theta_t^{-1} \), defined on set of all subsets of \( D \). Thus,

\[
\{ X_s \circ \theta_t \in A \} = \{ \xi \in D : X_s(\theta_t(\xi)) \in A \} = \\
= \{ \xi \in D : \theta_t(\xi) \in X_s^{-1}(A) \} = \{ \xi \in D : \xi \in \theta_t^{-1}X_s^{-1}(A) \} = \\
= \theta_t^{-1}X_s^{-1}(A) = \theta_t^{-1}\{ \xi \in D : X_s(\xi) \in A \} = \theta_t^{-1}\{ X_s \in A \}.
\]

From here,

\[
\{ X_{t+s_1} \in A_1, \ldots, X_{t+s_n} \in A_n \} = \bigcap_{k=1}^{n} \{ X_{t+s_k} \in A_k \} = \\
= \bigcap_{k=1}^{n} \{ X_{s_k} \circ \theta_t \in A_k \} = \bigcap_{k=1}^{n} \theta_t^{-1}\{ X_{s_k} \in A_k \} = \\
= \theta_t^{-1} \bigcap_{k=1}^{n} \{ X_{s_k} \in A_k \} = \theta_t^{-1} S,
\]

where \( S = \{ X_{s_1} \in A_1, \ldots, X_{s_n} \in A_n \} \) is a cylindrical set with finite-dimensional foundation. From the well-known theorem of extension of measure from algebra on a sigma-algebra generated by it (see [DYN 63]), a Markov behavior condition [1.2] can be rewritten in the following aspect:

\[
P(\theta_t^{-1}S, B) = E(P(\theta_t^{-1}S | X_t); B) \tag{1.3}
\]

for any set \( S \in \mathcal{F} \), whence the relation for conditional probabilities follows.

In terms of averages, the condition of a Markov behavior of process looks as follows:

\[
E(f(X_{t+s_1}, \ldots, X_{t+s_n}); B) = E(E(f(X_{t+s_1}, \ldots, X_{t+s_n}) | X_t); B).
\]

Using a shift operator, it is possible to note that for any measurable function \( f \) it holds:

\[
f(X_{t+s_1}, \ldots, X_{t+s_n}) = f(X_{s_1} \circ \theta_t \ldots, X_{s_n} \circ \theta_t) = f(X_{s_1}, \ldots, X_{s_n}) \circ \theta_t.
\]
From here, under the extension theorem, the Markov behavior condition can be rewritten in the following aspect:

$$E(g \circ \theta_t; B) = E(E(g \circ \theta_t | X_t); B), \quad [1.4]$$

where $g$ is arbitrary $\mathcal{F}$-measurable function on $\mathcal{D}$, whence the relation for conditional averages follows. Let us note that the condition [1.3] can be considered as a special case of conditions [1.4] where $f = I_S$. In this case, the following equality holds

$$E(I_S \circ \theta_t \cdot) = P(\theta_t^{-1}S \cdot).$$

1.2.2.2. Temporally homogeneous Markov process

A temporally homogeneous Markov process is usually defined in terms of transition functions.

A Markov transition function is called as a function $\tilde{P}_{s,t}(S | x)$, where $0 \leq t < s$ and

1) $S \rightarrow \tilde{P}_{s,t}(S | x)$ is a probability measure on $\mathcal{B}(\mathbb{R})$ for each $s, t$ and $x$;

2) $x \rightarrow \tilde{P}_{s,t}(S | x)$ is $\mathcal{B}(\mathbb{R})$-measurable function for each $s, t$ and $S$;

3) if $0 \leq t < s < u$, then

$$\tilde{P}_{u,t}(S | x) = \int_{-\infty}^{\infty} \tilde{P}_{s,t}(dy | x) \tilde{P}_{u,s}(S | y) \quad [1.5]$$

for all $x$ and $S$.

Relationship [1.5] is called the Chapman - Kolmogorov equation.

A Markov transition function $\tilde{P}_{s,t}(S | x)$ is said to be temporally homogeneous provided there exists a function $P_t(S | x)$ ($t > 0, x \in \mathbb{R}, S \in \mathcal{B}(\mathbb{R})$) such that $\tilde{P}_{s,t}(S | x) = P_{s-t}(S | x)$. For this
case, equation [1.5] becomes:

\[
P_{s+t}(S \mid x) = \int_{-\infty}^{\infty} P_s(dy \mid x) P_t(S \mid y)
\]  

[1.6]

We define the distribution of a temporally homogeneous Markov process to within the initial distribution as a consistent measurable family of measures \((P_x)\) on \(\mathcal{F}\), where \(P_x(X(0) = x) = 1\), and for any \(x \in \mathbb{R}, t > 0, B \in \mathcal{F}_t\) and \(S \in \mathcal{F}\) the following holds:

\[
P_x(\theta_t^{-1}(S); B) = E_x(P_{X_t}(S); B),
\]  

[1.7]

and for any measurable function \(f\):

\[
E_x(f \circ \theta_t; B) = E_x(E_{X_t}(f); B).
\]  

[1.8]

Finite-dimensional distributions of a temporally homogeneous Markov process is constructed from the temporally homogeneous transition functions according to the formula:

\[
P_x(X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n) =
\]

\[
= \int_{A_1} p_{t_1}(dx_1 \mid x) \int_{A_2} p_{t_2-t_1}(dx_2 \mid x_1) \times \cdots \times \int_{A_n} p_{t_n-t_{n-1}}(dx_n \mid x_{n-1})
\]

where \(p_t(dx_1 \mid x_0)\) is a transition kernel.

However, \textit{a priori} a set of transition functions submitting to coordination condition [1.6] do not necessarily define the probability measure on set of functions with given properties. In a class of Poisson processes, to verify the existence of a process with piece-wise constant trajectories requires a special proof.
1.2.3. Poisson process

1.2.3.1. Poisson distribution

The Poisson distribution is a discreet probability distribution on set of non-negative integers $\mathbb{Z}^+$ with values

$$p_n = \frac{\mu^n}{n!} e^{-\mu} \quad (n = 0, 1, 2, \ldots),$$

where $\mu > 0$ is the distribution parameter. Let us denote a class of Poisson distribution with a parameter $\mu$ as $\text{Pois}(\mu)$. Thus, from $\xi \in \text{Pois}(\mu)$ we understand that $\xi$ has Poisson distribution with parameter $\mu$.

It is known that expectation, variance and the third central moment of a Poisson distribution have the same meaning as the parameter of this distribution, i.e.:

$$E\xi = D\xi = E(\xi - E\xi)^3 = \mu.$$

A mode of the Poisson distribution is $n_{mod}$ such that $p_{n_{mod}} \geq p_n$ for each $n \in \mathbb{Z}^+$. This integer is determined by relations

$$p_{n+1}/p_n = \frac{\mu^{n+1}}{(n+1)!} \div \frac{\mu^n}{n!} = \frac{\mu}{n+1}.$$

1) If $\mu$ is an integer $n_1 + 1$, then $p_{n_1+1} = p_{n_1}$; for $n < n_1$, we have $p_{n+1}/p_n = \mu/(n+1) > \mu/(n_1 + 1) = 1$; this implies that in this case $p_n$ increases; analogously for $n > n_1 + 1$, $p_n$ decreases. Hence, there are two modes: $n_{mod}^{(1)} = n_1$ and $n_{mod}^{(2)} = n_1 + 1$.

2) Let $\mu$ be not an integer and $n_1 < \mu < n_1 + 1$; let us assume that $p_{n_1+1} \geq p_{n_1}$; this means that

$$\frac{\mu^{n_1+1}}{(n_1 + 1)!} \geq \frac{\mu^{n_1}}{n_1!};$$

which implies that $\mu \geq n_1 + 1$; from this contradiction, it follows that $p_{n_1+1} < p_{n_1}$; hence, $n_{mod} = n_1$ is a unique mode of this Poisson distribution.
Generating the function of a Poisson distribution (or corresponding random variable $\xi \in \text{Pois}(\mu)$) is a function of $\alpha \in \mathbb{R}$:

$$Ee^{\alpha\xi} = \sum_{n=0}^{\infty} e^{\alpha n} \frac{\mu^n}{n!} e^{-\mu} = \sum_{n=0}^{\infty} \frac{(\mu e^{\alpha})^n}{n!} e^{-\mu} = \exp(-\mu(1-e^{\alpha})).$$

[1.9]

Let $\xi_1$ and $\xi_2$ be independent Poisson variables with parameters $\mu_1$ and $\mu_2$, respectively. Then, the sum of these variables is a Poisson random variable with parameter $\mu_1 + \mu_2$. This can be proved easily by means of a generating function. Using independence, we have:

$$E \exp(\alpha(\xi_1 + \xi_2)) = E \exp(\alpha \xi_1) E \exp(\alpha \xi_2)$$

$$= \exp(- (\mu_1 + \mu_2)(1-e^{\alpha})).$$

This corresponds to the distribution $\text{Pois}(\mu_1 + \mu_2)$ as the equality is fair at any $\alpha \in \mathbb{R}$.

1.2.3.2. Poisson process

A non-decreasing integer random process $(N(t))$ ($t \geq 0$) with values from set $\mathbb{Z}_+$ is said to be a temporally homogeneous Poisson process if $N(0) = 0$ and if its increments on non-overlapping intervals are independent and have Poisson distributions. That is, there exists such a positive $\beta$, called the intensity of process, that $N(t) - N(s) \in \text{Pois}(\beta(t-s))$ ($0 \leq s < t$). For $N(t)$, we will also use a label $N_t$. This process has step-wise trajectories with unit jumps. By the additional definition, such a trajectory is right continuous at point of any jump.

The sequence of the moments of jumps of the process $(\sigma_n)$ ($n \geq 1$) completely characterizes a Poisson process. This sequence is called a point-wise Poisson process. Let us designate $T_n = \sigma_n - \sigma_{n-1}$ ($n \geq 1$, $\sigma_0 = 0$), where $(T_n)$ is a sequence of independent and identically distributed (i.i.d.) random variables with common exponential
distribution $P(T_1 > t) = e^{-\beta t}$. Using a shift operator on set of sample trajectories of a Poisson process, it is possible to note that

$$\sigma_{n+1} = \sigma_n + \sigma_1 \circ \theta_{\sigma_n}.$$

A generalization of the above process, the so-called inhomogeneous Poisson process $(N(t))$, is characterized by means of a non-constant function of intensity $\beta(t) \geq 0$, setting Poisson distributions as independent increments, for $0 \leq s < t$

$$P(N(t) - N(s) = n) = \frac{\mu(s,t)^n}{n!} e^{-\mu(s,t)} \quad (n \in \mathbb{Z}_+),$$

where

$$\mu(s, t) = \int_s^t \beta(u) \, du.$$ 

Gaps between unit jumps in this process are not independent and are not identically distributed.

1.2.3.3. Stochastic continuity

A random process $X(t)$ is called stochastically continuous from the right at a point $t_0$, if for any positive $\epsilon$ and $t > t_0$

$$\lim_{t \to t_0} P(|X(t) - X(t_0)| > \epsilon) = 0.$$

Similarly, a stochastic continuity from the left is defined. Bilateral stochastic continuity is both from the left and right. For a Poisson process with locally limited intensity $\beta(t)$, the stochastic continuity from the right at a point $t_0$ follows from inequality:

$$P(X(t) - X(t_0) > \epsilon) \leq P(X(t) - X(t_0) > 0) =$$

$$= 1 - \exp \left( - \int_{t_0}^t \lambda(s) \, ds \right) = \int_{t_0}^t \lambda(s) \, ds + o(t - t_0).$$

Stochastic continuity from the left is proved analogously.
1.2.3.4. **Composite Poisson process**

A random process \((X_t) \; (t \geq 0)\) is called a temporally homogeneous composite Poisson process if it is defined by means of temporally homogeneous Poisson process \((N(t)) \; (t \geq 0)\) and a sequence \((U_n) \; (n \geq 1)\) of i.i.d. random variables, where \((N(t))\) and \((U_n)\) are independent. By definition:

\[
X_t = \sum_{n=1}^{N_t} U_n \quad (t \geq 0).
\]

Let us designate \(B(x) = P(U_1 \leq x)\) as the cumulative distribution function of \(U_1\), \(\mu_B = EU_1\) the expectation of \(U_1\) and \(\mu_B^{(n)} = EU_1^n\) the \(n\)th moment of \(U_1\) \((n \geq 1)\). The sequence of jump times, \((\sigma_n) \; (n \geq 1)\), of the composite Poisson process coincides with sequence of jumps of the original Poisson process and hence it is possible to note that:

\[
X_t = \sum_{n=1}^{\infty} U_n I_{\{\sigma_n \leq t\}}.
\]

From here the formula for an average follows:

\[
EX_t = \mu_B \sum_{n=1}^{\infty} P(\sigma_n \leq t) = \mu_B EN_t = \mu_B \beta t.
\]

In the next chapters, this process will be considered in more detail.

1.2.4. **Gamma process**

A gamma function is called an integral depending on parameter \(m > 0\),

\[
\Gamma(m) = \int_0^{\infty} x^{m-1} e^{-x} \, dx.
\]

For a case where \(m = 1, 2, 3, \ldots\) we have \(\Gamma(m) = (m - 1)!\).
The non-negative random variable \( X \) has a gamma distribution if its distribution density resembles

\[
f_X(x) = \frac{\delta}{\Gamma(\gamma)} (x\delta)^{\gamma-1} e^{-x\delta} \quad (x > 0),
\]

where \( \delta \) is a scale parameter and \( \gamma \) is a form parameter of the distribution. We designate such a class of random variables as \( \text{Gam}(\gamma, \delta) \).

At \( \gamma = 1 \), the gamma distribution coincides with a exponential distribution with parameter \( \delta \). For the integer \( \gamma = n \), where \( n \geq 2 \), the gamma distribution is called the Erlang distribution. It is an \( n \)-fold convolution of exponential distributions.

Let us obtain a Laplace transformation of a gamma distribution density:

\[
\int_0^\infty e^{-\lambda x} \frac{\delta}{\Gamma(\gamma)} (x\delta)^{\gamma-1} e^{-x\delta} \, dx = \\
= \left( \frac{\delta}{\lambda + \delta} \right)^\gamma \int_0^\infty e^{-(\lambda+\delta)x} \frac{\gamma^{-1}}{\Gamma(\gamma)} (x(\lambda+\delta))^{\gamma-1} \, dx = \left( \frac{\delta}{\lambda + \delta} \right)^\gamma.
\]

From here it follows that the sum of two independent random variables \( X_1 \in \text{Gam}(\gamma_1, \delta) \) and \( X_2 \in \text{Gam}(\gamma_2, \delta) \) is a random variable from a class \( \text{Gam}(\gamma_1 + \gamma_2, \delta) \).

A process \( X(t) \) \( (t \geq 0) \), possessing the following properties:

1) \( X(0) = 0 \), process trajectories do not decrease and are continuous from the right;
2) the process has independent increments;
3) at any \( s \geq 0 \) and \( t > s \) the increment \( X(t) - X(s) \) belongs to class \( \text{Gam}(\gamma(t-s), \delta) \);

is called a gamma process with parameters \( \gamma \) and \( \delta \).
Let us prove that a homogeneous gamma process is stochastically continuous in any point of the area of the representation. Designate $\gamma_1 = \gamma (t - t_0)$. We have at $t > t_0$ and $0 < \gamma_1 < 1$ (without loss of generality):

$$P(|X(t) - X(t_0)| > \epsilon) = P(X(t) - X(t_0) > \epsilon) = P(X(t-t_0) > \epsilon) =$$

$$= \int_\epsilon^\infty \frac{\delta}{\Gamma(\gamma_1)} \left(\delta x\right)^{\gamma_1-1} e^{-\delta x} \, dx = \frac{1}{\Gamma(\gamma_1)} \int_{\epsilon\delta}^\infty y^{\gamma_1-1} e^{-y} \, dy \leq$$

$$\leq \frac{1}{\Gamma(\gamma_1)} \left(\epsilon\delta\right)^{\gamma_1-1} \int_{\epsilon\delta}^\infty e^{-y} \, dy \leq \frac{1}{\Gamma(\gamma_1)} \left(\epsilon\delta\right)^{\gamma_1-1}. $$

From our definition of the function $\Gamma$ follows that $\Gamma(z) \to \infty$ at $z \downarrow 0$. Hence, the inequality right member aspires to zero. The stochastic continuity of the gamma process is proved. □

Finite-dimensional distributions of a gamma process possess good analytical properties. The practical application of the gamma process model is hindered a little by the property of its sample trajectories because these trajectories may have ruptures such as in short intervals. It is possible to tell that trajectories “consist only of positive jumps”. The inverse gamma process is more convenient for physical interpretation.

**1.2.5. Inverse gamma process**

The process $X(t)$ ($t \geq 0$) is known as the inverse gamma process if it possesses the following properties:

1) $X(0) = 0$, process trajectories do not decrease and are continuous with probability 1;

2) process $X(t)$ possesses a Markov property with respect to the time of the first exit from any interval $[0, u]$;

3) the inverse process for the process $X(t)$, i.e. the process:

$$Y(u) \equiv \inf\{t : X(t) \geq u\} \quad (u > 0),$$

is a gamma process with some parameters $\gamma$ and $\delta$. 
Sample trajectories of this process are exotic enough. Almost everywhere (concerning a Lebesgue measure) for domain of definition, these trajectories are constant (they have zero derivatives). The increase of these trajectories on an interval \([0, \infty)\) is ensured with the presence of a non-enumerable set of points of growth (like in a Cantor curve), filling any interval from the moment of first reaching level \(u_1\) (the random moment \(Y(u_1)\)) until moment \(Y(u_2)\), where \(0 \leq u_1 < u_2\). Intervals of constancy of a sample trajectory of \(X(t)\) correspond to jumps of the trajectory \(Y(u)\). It is important to note that the beginning time of each interval of constancy of the process \(X(t)\) is not a Markov time. Thus, it is an example of a real stopping time, which is not a Markov time.

In modern terminology, processes such as the inverse gamma process are known as continuous semi-Markov processes (see [HAR 07]).

1.2.6. Renewal process

Renewal theory is commonly used in risk theory. For example, a renewal process can serve as a model when entering sequences of claim times into an insurance company (instead of using a Poisson process). Renewal equations arise during the analysis of probability of ruin. The asymptotics of a solution of such an equation allows to express probability of ruin in case of a high initial capital.

1.2.6.1. Renewal process

A simple temporally homogeneous renewal process is said to be a non-decreasing integer random process \((N(t)) \ (t \geq 0)\). It is assumed that \(N(0) = 0\), where a process has jumps of unit magnitude, and where distances in time \((T_n) \ (n \geq 1)\) between the neighboring jumps (the renew times) are i.i.d. positive random variables, and at any jump time a sample trajectory of the process is continuous from the right. Such a process is determined by a distribution function of \(T_1\), such as \(F(x) = P(T_1 \leq x) \ (x \geq 0)\). Magnitude \(T_k\) is interpreted as distance in time between the \((k - 1)\)th and \(k\)th process jumps, thus \(\sigma_n = \sum_{k=1}^{n} T_k\) is a time of the \(n\)th renew.
A temporally homogeneous Poisson process is a partial case of renewal process. In the Poisson case, $F(x) \equiv P(T_1 \leq x) = 1 - e^{-\beta x}$ $(x \geq 0)$ for some $\beta > 0$. Also for a Poisson process, we will sometimes use the notation $N_t$ instead of $N(t)$.

1.2.6.2. Renewal equation

Outcomes of renewal theory are used in risk theory mainly in connection with a solution of so-called renew equations. First, we consider the so-called renewal function.

Renewal function $H(t)$ $(t \geq 0)$ is expressed as:

$$H(t) \equiv 1 + E N_t = 1 + \sum_{n=1}^{\infty} P(\sigma_n \leq t) = \sum_{n=0}^{\infty} F^{(n)}(x),$$

where $F^{(n)}$ is $n$-fold convolution of distribution functions $F$:

$$F^{(n)}(x) = \int_{0}^{x} F^{(n-1)}(x-y) dF(y) \quad (n \geq 1).$$

$F^{(0)}(x) = I_{[0,\infty)}(x)$, “zero convolution”. It corresponds to the sum $n$ of i.i.d. random variables; we will also use the notation $H_t$ in addition to $H(t)$. Using a permutability of summation with convolution operation, we obtain the equation:

$$H(t) = 1 + \int_{0}^{t} H(t-x) dF(x).$$

For the given cumulative distribution function $F$ on interval $[0, \infty)$, and a known function $y(t)$ $(t \geq 0)$, the equation

$$Z(t) = y(t) + \int_{0}^{t} Z(t-x) dF(x) \quad [1.10]$$

is a renewal equation concerning unknown function $Z(t)$. The solution of the renewal equation always exists and is unique:

$$Z(t) = \int_{0}^{t} y(t-x) dH(x).$$
It is easy to prove this by substituting the right-hand-side of equation [1.10], where the function expressed, by the whole equation (i.e. iterating the equation).

Analytical expression for function $H_t$ is known only in exceptional cases. For example, if $F$ is an exponential distribution function with parameter $\beta$, then $H_t = 1 + \beta t$, in the case of a Poisson process. The basic outcome of the theory is connected with an asymptotics of the renewal function and a limit of a solution of the equation [1.10].

**Theorem 1.3.** – Elementary renewal theorem

$$\frac{H_t}{t} \to \frac{1}{ET_1} \quad (t \to \infty).$$

**Theorem 1.4.** – Blackwell theorem

$$H_{t+s} - H_t \to \frac{s}{ET_1} \quad (t \to \infty).$$

**Theorem 1.5.** – Smith theorem

For any function $y(t)$ immediately integrable by Riemann:

$$\int_0^t y(t - x) \, dH(x) \to \frac{1}{ET_1} \int_0^\infty y(t) \, dt \quad (t \to \infty).$$

**Proof.** – (for all the three theorems, see Feller [FEL 66]). An immediately integrable function by Riemann on an interval $(a, b)$ is called a function $f$ for which there exist identical limits

$$\lim_{n \to \infty} \sum_{k=1}^n \frac{b-a}{n} m_k,$$

$$m_k = \min \left\{ f(x) : \frac{b-a}{n} (k-1) \leq x \leq \frac{b-a}{n} k \right\},$$

$$\lim_{n \to \infty} \sum_{k=1}^n \frac{b-a}{n} M_k,$$

$$M_k = \max \left\{ f(x) : \frac{b-a}{n} (k-1) \leq x \leq \frac{b-a}{n} k \right\}.$$
A function for which this condition is fulfilled for all its restrictions on final intervals is an immediately integrable function by Riemann on an infinite interval. An example of such function is any monotone function integrable by Riemann.

1.2.6.3. Direct and inverse renewal times

In risk theory, properties of the so-called direct and inverse renewal times are used. They are as follows:

\[ \zeta(t) \equiv \sigma_{N_t+1} - t, \quad \eta(t) \equiv t - \sigma_{N_t}. \]

We have:

\[
P(\eta(t) > x, \zeta(t) > y) = \sum_{n=1}^{\infty} \int_{0}^{t} P(t - \sigma_{n-1} > x, \sigma_{n} - t > y, \sigma_{1} \in ds) =
\]

\[
= I_{t>x} P(\sigma_{1} - t > y) + \sum_{n=2}^{\infty} \int_{0}^{t} P(t - \sigma_{n-1} > x, \sigma_{n} - t > y, \sigma_{1} \in ds) .
\]

Using representation \( \sigma_k = \sigma_1 + \sigma_{k-1} \circ \theta_{\sigma_1} \) \((k \geq 2)\) and property of a renewal process concerning the time \( \sigma_1 \), we know that this expression is:

\[
I_{t>x} P(\sigma_{1} - t > y) + \\
+ \sum_{n=2}^{\infty} \int_{0}^{t} P(t - s - \sigma_{n-2} \circ \theta_{\sigma_1} > x, s + \sigma_{n-1} \circ \theta_{\sigma_1} - t > y, \sigma_{1} \in ds) =
\]

\[
= I_{t>x} P(\sigma_{1} - t > y) + \\
+ \sum_{n=2}^{\infty} \int_{0}^{t} P(t - s - \sigma_{n-2} > x, s + \sigma_{n-1} - t > y) P(\sigma_{1} \in ds).
\]

Designating \( P(\eta(t) > x, \zeta(t) > y) = Z(t) \), \( I_{t>x} P(\sigma_{1} - t > y) = y(t) \), we come to the equation [1.10] for which limit of a solution as \( t \to \infty \) is:

\[
\frac{1}{\mu} \int_{0}^{\infty} I_{t>x} P(T_1 - t > y) dt = \frac{1}{\mu} \int_{x}^{\infty} P(T_1 > y+t) dt = \frac{1}{\mu} \int_{x+y}^{\infty} F(t) dt,
\]
where $\mu = ET_1$ and $\bar{F}(t) = 1 - F(t)$. From here, both variables $\eta(t)$ and $\zeta(t)$ have the same limit distribution:

$$\tilde{F}(x) = \frac{1}{\mu} \int_{0}^{x} \bar{F}(t) \, dt.$$ 

In risk theory, the following properties of the variable $\zeta(t)$ are useful.

**THEOREM 1.6.– Property of direct renewal time**

For renewal process, the following limits are true:

a) $\frac{\zeta(t)}{t} \xrightarrow{a.s.} 0 \quad (t \to \infty)$;

b) $\frac{E\zeta(t)}{t} \to 0 \quad (t \to \infty)$.

**PROOF.–** (see, for example, Asmussen [ASM 00]).