Crossings Problems in Random Processes Theory and Their Applications in Aviation

Crossings Problems in Random Processes Theory and Their Applications in Aviation

By

Sergei L. Semakov

Cambridge Scholars Publishing

Crossings Problems in Random Processes Theory and Their Applications in Aviation

By Sergei L. Semakov

This book first published 2019

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN (10): 1-5275-2795-6
ISBN (13): 978-1-5275-2795-9

To My Mother

Contents

Introduction

This monograph presents the author's results that have only been previously partially published in the form of journal articles.[1] These results relate to the special section on the theory of random processes and their application in solving aviation problems.[2] This section is known in literature as outliers of random processes, or crossings problems.[3]

The book consists of three chapters. In the first chapter, the basic concepts in the theory of probability and random processes are stated at an elementary level in order to prepare the reader for the second and the third chapters. Specialists in the theory of probability and random processes can skip the first chapter. In this chapter, we have introduced the concepts of probabilistic space, random variables, random processes, correlation function, and the spectral density of random process, and other important concepts. We have defined types of convergence in a probabilistic space and we have formulated the corresponding characteristics in the smoothness of random processes. Stationary, Gaussian, and Markov processes are considered. Questions of the integration of random processes are discussed.

---

[1] See, for example, the following articles:

(a) S.L. Semakov, First Arrival of a Stochastic Process at the Boundary, *Autom. Remote Control*, 1988, vol. 49, no. 6, pp. 757-764;

(b) S.L. Semakov, The Probability of the First Hiting of a Level by a Component of a Multidimensional Process on a Prescribed Interval under Restrictions of the Remaining Components, *Theor. Prob. App.*, 1989, vol. 34, no. 2, pp. 357-361;

(c) S.L. Semakov, The Application of the Known Solution of a Problem of Attaining the Boundaries by Non-Markovian Process to Estimation of Probability of Safe Airplane Landing, *J. Comput. Syst. Sci.*, 1996. vol. 35, no. 2, pp. 302-308;

(d) S.L. Semakov, Estimating the Probability That a Multidimensional Random Process Reaches the Boundary of Region, *Autom. Remote Control*, 2015, vol. 76, no. 4, pp. 613-626.

[2] The formulations of the main results are listed at the end of the book in the "Conclusion".

[3] See, for example, the book: H. Cramer and M.R. Leadbetter, *Stationary and Related Stochastic Processes*, New York-London-Sydney: Wiley, 1967.

The behavior of any real system is a process which is, to a greater or lesser degree, probabilistic. As a rule, it is impossible to specify exactly what external influences and internal mechanisms in the interaction of the system components will be decisive in the future. As a consequence, we cannot accurately predict the behavior of the system. We can only talk about a probability that the system will come to a particular state in the future. If we pose the problem of the probabilistic description of all possible future states of the system, then this problem will be very difficult. Fortunately, for research purposes, it is often enough to get answers to simpler questions; for example, the question "How long will on average the system operate in a given mode?" or the question "What is the probability that the process of functioning of the system will come out of a given mode to a specific point in time?" Problems of this type are called problems about outliers of random processes, or problems about the crossing of a level. In the second chapter, we have stated some of the most important fundamental results, which are related to crossings problems. Known problems about attaining boundaries in a random process will be discussed in detail. We have stated the well-known solution of diffusion Markov processes for these problems and presented the author's results for arbitrary continuous processes.

In the third chapter, we have applied the mathematical results of the second chapter to an investigation on the safety of airplane landings. We have not considered simple examples and model problems, whose purpose is only to illustrate the theory, but real concrete problems posed by practice. We have shown how the results of the second chapter can be used to calculate the probability of an airplane's safe landing. A safe landing is defined as an event when the airplane touches down the landing surface on a given segment for the first time, and, at this moment, the coordinates of the airplane (elevation angle, banking angle, vertical velocity and so on) remain inside admissible ranges. These restrictions are established by flight standards and their violation leads to an accident. The scheme for this calculation is described and the implementation of this scheme is given for overland and ship-landing options. The results of the numerical calculations are discussed in detail.

The first chapter, and part of the second chapter are used to carrying out of lessons with students of the Moscow Institute of Physics and Technology and the Moscow Aviation Institute. The author's results in the second chapter will be of interest for mathematicians who study crossings problems. The third chapter is addressed to specialists in the field of aviation, as well as engineers and scientists who are interested in the application of random processes theory and use its methods. Some results of the second and third chapters were presented in the author's report at the 57th IEEE Conference on Decision and Control (Miami Beach, FL, USA, December 17-19, 2018).[1]

This monograph was recommended for publication by the Division of Mathematical Sciences at the Russian Academy of Sciences.

[1] S.L. Semakov and I.S. Semakov, Estimating the Probability That a Random Process First Reaches the Boundary of a Region on a Given Time Interval, *Proceedings of the 57th IEEE Conference on Decision and Control (CDC)*, Miami Beach, USA, 2018, pp. 256-261.

Chapter 1

Main Classes and Characteristics of Random Processes

1.1 Intuitive Prerequisites of the Theory

Let us consider the practical situation connected with an experiment E when each outcome of E is defined by a casual mechanism and the influence cannot be predicted in advance. In any event we are interested in A, which is connected with E; this is in the sense that the event A can either occur or not as a result of E. In many life situations there is a necessity to predict the possible degree of realization of the event A as a result of carrying out E. This degree of possibility is characterized by the number $P\{A\}$. This number is called the probability of A and is defined by the frequency of occurrence of A in numerous repetitions of the experiment E or, more precisely,

$$P\{A\} = \lim_{N\to\infty} \frac{N_A}{N},$$

where N is the number of experiments carried out, $N_A = N_A(N)$ is the number of experiments within which the event A was observed. We will put aside the question about the existence of the above limit and the possibility of its practical calculation.

As a rule, certain quantitative characteristics can be connected with the experiment E. These characteristics accept numerical values that change in a random way from one experiment to another. These characteristics are called random values. For example, if the experiment E consists in n tossings of a coin, the number of dropped-out coats of arms, the maximum number of coats of arms which dropped out in a row, the number of coats of arms dropped out at even throwing, and so on will be random values. It is convenient to describe the problematic behavior of a random value X by means of the function

$$F(x)=P\{X<x\},$$

which is called a distribution function of X and is the probability of an event concluded in braces: i.e., the event consisting of the experiment value of X less the real number x. It turns out that knowledge of $F(x)$ provides a chance to receive the probability that a random value of X belongs to the preset subset of real numbers. This problem often arises in the solution to various practical tasks.

It is possible to consider a random vector value $(X_1, \ldots, X_n)$, where every X_i has a scalar random value. It is important from the practical point of view to be able to determine the probability that a casual point $(X_1, \ldots, X_n)$ in the n-dimensional space R^n as a result of E will get to the preset subset from R^n. For the big class of such subsets in the case of a scalar random value, this probability is unambiguously defined by the joint distribution function of the random values $X_1, \ldots, X_n$:

$$F(x_1, \ldots, x_n) = P\{X_1 < x_1, \ldots, X_n < x_n\},$$

which is the probability that each of X_i appeared less than the corresponding value of x_i as a result of E.

A vector random value $(X_1, \ldots, X_n)$ can be considered and defined as a family of random values $\{X_t\}$, where t runs some set of indexes T; in this case $T = \{1, \ldots, n\}$. If T represents a set of integers, then the family $\{X_t\}$ is called a random process with a discrete parameter. If T is an interval of a real axis, then a family of random values $\{X_t\}$ is called a random process with a continuous parameter. In this case, an outcome of the experiment is a set of values X_t, where, for example, $t \in [a, b]$; in other words, an outcome of this experiment gives a function of the variable t, where $t \in [a, b]$. The parameter t is called the argument of a random process. Most often the argument of a random process is time (from here and the chosen designation of t); however, this is optional. We will consider, for example, a process of change in the flying height H of a trip plane in its flight from point A to point B. This is dependent on the flying range l; $l \in [0, L]$, where L is a distance between points A and B. Here the experiment E is the flight of the plane from A to B. A result of E is the concrete (corresponding to this concrete flight) height change function of the variable l. It is clear that this type of function can change from flight to flight and depends on many random factors: weather conditions, the technical condition of the plane, the health of the pilot, indications of the land services of air traffic control, and other factors. At every fixed value of the l value, the H_l from experiment to experiment (from flight to flight) will change in an unpredicatable fashion and can, therefore, be considered a random value, and the family $\{H_l\}$ can be considered as a random process.

Therefore, in a definite sense, the random process combines the features of a random value and a function; at the fixed argument it turns into a random value, and at each implementation of an experiment it turns out to be a determined (not random) function of this argument. As a rule, we will further designate the random process as $X(t)$; specifying that if it is necessary then the possible range of change in the argument t, and the physical sense of magnitudes in X and t can be various.

1.2 Fundamental Concepts and Results Underlying the Construction of the Mathematical Theory

1. We consider the experiment E without being interested in its concrete type. An event A is connected with E and is called observable if we can say unambiguously that event A has occurred or that event A has not occurred as a result of E. An event is called a persistent event and denoted by Ω if this event occurs every time when carrying out experiment E. An event is called an impossible event and denoted by 0 if this event never occurs when carrying out experiment E. We consider that events, Ω and 0, are observable. Let A and B be observable events that are connected with carrying out experiment E. We define:

1) the additional event concerning A; this event is denoted by $\bar{A}$ and consists in the fact that A does not occur;
2) the sum, or crowd of events A and B; this event is denoted by $A + B$ or $A \cup B$ and consists in the fact that at least one of events A or B occurs;
3) the product, or intersection of events A and B; this event is denoted by AB or $A \cap B$ and consists in the fact that both events A and B occur.

We consider that if events A and B are observable, then events $\bar{A}$, $\bar{B}$, $A+B$, and AB are also observable. The family of all observable events connected with E forms a field of events $\mathcal{F}_0$: i.e., the class of events such that $\Omega \in \mathcal{F}_0$ and if $A, B \in \mathcal{F}_0$, then $\bar{A}, \bar{B}, A+B, AB \in \mathcal{F}_0$.

It can easily be checked out that events of field $\mathcal{F}_0$ satisfy the following relations:

$$A+A=AA=A,\ A+B=B+A,\ (A+B)+C=A+(B+C),$$

$$AB=BA,\ (AB)C=A(BC),\ A(B+C)=AB+AC,\ A+\bar{A}=\Omega,$$

$$A\bar{A}=0,\ A+\Omega=\Omega,\ A\Omega=A,\ A+0=A,\ A0=0.$$

Let $A', B', \ldots$ be sets of points ω of any space Ω'. The sets $\overline{A'}, \overline{B'}, A'+B', A'B'$ are defined in the elementary theory of sets. These definitions (which are assumed known for the reader) show that all ratios given above for events remain fair when $A, B, \ldots$ are sets. A family of sets from Ω' is called a field of sets in Ω' if this family includes the entire space Ω' and is closed relative to operations $\overline{A'}$, $A'+B'$, and $A'B'$.

It turns out that the following result takes place: for any field $\mathcal{F}_0$ of events satisfying the listed above ratios it is possible to find some space Ω' of points ω and field $\mathcal{F}'_0$ of ω-sets. This means that a biunique correspondence exists between events A taking from $\mathcal{F}_0$ and sets A' taking from $\mathcal{F}'_0$. If the event A corresponds to set A' and event B corresponds to set B', then event $\bar{A}$ corresponds to set $\overline{A'}$, event $A+B$ corresponds to set $A'+B'$, and event AB corresponds to set $A'B'$. Points ω correspond to some elementary events which can or cannot be observed separately (i.e., can enter or cannot enter $\mathcal{F}_0$ separately); a persistent event corresponds to the whole space Ω', and an impossible event corresponds to the empty set $\emptyset$.

The formulated result allows us to use the technique of the theory of sets for an analysis of fields of events. Further strokes are omitted and the same designations $\mathcal{F}_0, A, B, \ldots$ are used for events and ω-sets. A set from $\mathcal{F}_0$ is called an observable event, or ω-set, and Ω is considered as a persistent event or as the whole space.

We now assume that the number $P_0\{A\}$ is placed according to each event A from $\mathcal{F}_0$. This number is called the probability of event A. We consider that $P_0\{A\}$ is a function, so that $P_0\{A\}$ is determined for all $A \in \mathcal{F}_0$ and the following conditions are satisfied:

1) $0 \le P_0\{A\} \le 1$;
2) $P_0\{\Omega\} = 1$;
3) if $A = A_1 + \ldots + A_n$, where $A_i \in \mathcal{F}_0$, $i = 1, \ldots, n$, $A_j A_k = \emptyset$ for $j \ne k$, then $P_0\{A\} = P_0\{A_1\} + \ldots + P_0\{A_n\}$.

The statement about the existence of the function $P_0\{A\}$ with these properties is accepted as an axiom, i.e., as a statement which does not demand proof.

Let $A_1, A_2, \ldots$ be any countable sequence of sets from $\mathcal{F}_0$. Sets of a type $A_1 + A_2 + \ldots$ and $A_1 A_2 \ldots$ (i.e., sets that are denumerable number of crowds and intersections of sets from $\mathcal{F}_0$) may not be elements of field $\mathcal{F}_0$. A field of ω-sets is called a borelevsky field or σ-field if this field includes all countable (finite or infinite) crowds and intersections of elements of this field. It turns out that for any field $\mathcal{F}_0$ of observable events there is a minimum σ-field containing $\mathcal{F}_0$. This minimum σ-field is denoted by $\mathcal{F}$. The field $\mathcal{F}$, as well as the field $\mathcal{F}_0$, may also contain unobservable events.

One of the main results is that there is only a unique extension of the function $P_0\{A\}$, which is defined for all sets $A \in \mathcal{F}_0$, to the function $P\{A\}$, which is defined for all sets $A \in \mathcal{F}$: if $A \in \mathcal{F}_0$, then $P\{A\} = P_0\{A\}$. The function $P\{A\}$ possesses all properties of the function $P_0\{A\}$. In particular, $0 \leq P\{A\} \leq 1$ for any $A \in \mathcal{F}$ and if $A = A_1 + A_2 + \ldots$, where all $A_i \in \mathcal{F}$ and $A_i A_j = \emptyset$ for $i \neq j$, then $P\{A\} = P\{A_1\} + P\{A_2\} + \ldots$.

The space Ω of points ω, the σ-field $\mathcal{F}$ of sets from Ω, and the probability $P\{A\}$ defined for sets A from $\mathcal{F}$ form the probabilistic space $(\Omega, \mathcal{F}, P)$. Sets from $\mathcal{F}$ are called measurable.

2. Let $X = X(\omega)$ be a function defined for all ω. If for any $x \in (-\infty, \infty)$ the set $\{\omega\colon X(\omega) < x\}$ is an element of σ-field $\mathcal{F}$, then we say that function X is measurable and we call this function a random value.

If $X = X(\omega)$ is a random value, then the probability $F(x) = P\{\omega\colon X(\omega) < x\}$ (or briefly $F(x) = P\{X < x\}$) represents the nondecreasing function of the variable x. It is easy to prove that this function is continuous at the left, $\lim\limits_{x \to -\infty} F(x) = 0$, and $\lim\limits_{x \to \infty} F(x) = 1$. The function $F(x)$ is called a distribution function of the random value X.

If function $f(x)$ exists such that

$$F(x) = \int_{-\infty}^{x} f(t)dt,$$

then $f(x)$ is called a distribution density of random value X. For a system of random values $X_1, \ldots, X_n$, a function of joint distribution of these random values is defined as

$$F(x_1,\ldots,x_n) = P\{\omega\colon X_1(\omega) < x_1,\ldots,X_n(\omega) < x_n\}.$$

The density of joint distribution of these random values (if this density exists) is defined as a function $f(x_1,\ldots,x_n)$ such that

$$F(x_1,\ldots,x_n) = \int\limits_{-\infty}^{x_1}\ldots\int\limits_{-\infty}^{x_n} f(t_1,\ldots,t_n)dt_1\ldots,dt_n.$$

Let $(\Omega,\mathcal{F},P)$ be some probabilistic space and T be a set of values from the parameter t. The function $X(t,\omega)$, where $t\in T$ and $\omega\in\Omega$, is called a random process on a probabilistic space $(\Omega,\mathcal{F},P)$ if the following condition is satisfied: for each fixed $t=\tilde{t}$ from T, the $X(\tilde{t},\omega)$ is some random value on this probabilistic space $(\Omega,\mathcal{F},P)$. The records $X(t)$ or X_t instead of $X(t,\omega)$ are used for a brief designation of this random process.

Let $X(t)$ be a random process. For each fixed $t = t_1$, the random value $X(t_1)=X(t_1,\omega)$ has the distribution function $F(x_1,t_1)=P\{\omega\colon X(t_1,\omega)<x_1\}$. Let $t_1,\ldots,t_n$ be any finite set of values t. The random values $X(t_1),\ldots,X(t_n)$ have the function of joint distribution

$$F(x_1,\ldots,x_n;t_1,\ldots,t_n) = P\{\omega\colon X(t_1,\omega) < x_1,\ldots,X(t_n,\omega) < x_n\}.$$

The family of all such joint distributions for $n = 1,2,\ldots$ and for all possible values $t_j \in T$, where $j=1,2,\ldots,n$, is called a family of finite-dimensional distributions of the process $X(t)$. As it will become clear from further sections, many properties of random processes are defined by the properties of their finite-dimensional distributions.

Two random processes $X(t)$, $t\in T$, and $Y(t)$, $t\in T$, are called equivalent if for each fixed $t\in T$ the random values $X(t)$ and $Y(t)$ are equivalent random values, i.e., $P\{\omega\colon X(t,\omega)=Y(t,\omega)\} = 1$. It is easy to prove that families of finite-dimensional distributions for equivalent processes coincide.

From the definition of the random process, it follows that $X(t,\omega)$ becomes a function of variable $t\in T$ for every fixed elementary event ω. In other words, a nonrandom function of the variable t corresponds to each possible outcome of the experiment. Each such function $x(t)$ is called a realization or a trajectory or a sample function of process $X(t)$.

1.3 Mathematical Expectation, Variance, and Correlation Function of a Random Process

Let $X = X(\omega)$ be a discrete random value defined on a probabilistic space $(\Omega, \mathcal{F}, P)$. Possible values of X are given by countable (finite or infinite) numerical sequence $x_1, x_2, \dots$. We will assume that a set $A_i \in \mathcal{F}$ is formed by those and only those ω for the $X(\omega) = x_i$. If the series $\sum\limits_i x_i P\{A_i\}$ converges absolutely, then its sum is called a mathematical expectation of random value X and is denoted by $M\{X\}$:

$$M\{X\} = \sum_i x_i P\{A_i\}.$$

If X is a continuous random value having a distribution density $f(x)$, then by definition

$$M\{X\} = \int_{-\infty}^{\infty} x f(x) dx$$

given that this integral converges absolutely. In the case of divergence of this integral (or in the case of divergence of the above series if X is a discrete random value), we say that the corresponding random value has no mathematical expectation.

By definition, the mathematical expectation of random process $X(t)$ is a nonrandom function

$$m(t) = M\{X(t)\}.$$

The right part of this equality is a mathematical expectation of the random value $X(t)$. We interpret this random value as the random process cross-section corresponding to the argument t.

By definition, the variance of random process $X(t)$ is a nonrandom function

$$D(t) = M\{(X(t) - m(t))^2\}.$$

For each fixed t, the number $D(t)$ gives the variance of random value $X(t)$.

Random values $X_1, \dots, X_n$ are called independent if

$$F(x_1, \dots, x_n) = F_1(x_1) \dots F_n(x_n)$$

or (if $f(x_1, \dots, x_n)$ exists)

$$f(x_1, \dots, x_n) = f_1(x_1) \dots f_n(x_n),$$

where $F_i(x)$ is a distribution function of random value X_i, $i=1,\dots,n$, and $f_i(x)$ is a distribution density of random value X_i, $i=1,\dots,n$. To characterize a degree of dependence between various process cross-sections, i.e., between random values $X(t_1)$ and $X(t_2)$, we define a nonrandom function of two variables as

$$K(t_1,t_2) = M\{(X(t_1)-m(t_1))(X(t_2)-m(t_2))\}.$$

This function is called a correlation function of the random process $X(t)$. If we introduce a centered random process as

$$\overset{\circ}{X}(t) = X(t) - m(t),$$

then we obtain

$$K(t_1,t_2) = M\{\overset{\circ}{X}(t_1)\,\overset{\circ}{X}(t_2)\}.$$

Notice that $K(t,t) = D(t)$.

Suppose $D(t)\neq 0\quad \forall t \in T$. Then we can introduce the function

$$k(t_1,t_2) = \frac{K(t_1,t_2)}{\sigma(t_1)\sigma(t_2)},$$

where $\sigma(t_1)=\sqrt{D(t_1)}$ and $\sigma(t_2)=\sqrt{D(t_2)}$ are so-called mean square deviations. The function $k(t_1,t_2)$ is called a normalized correlation function of the random process $X(t)$. A convenience from the introduction of $k(t_1,t_2)$ consists of the fact that $k(t_1,t_2)$ is a nondimensional quantity and $|k(t_1,t_2)| \le 1$ for any values t_1 and t_2.

It is easy to prove that $k(t_1,t_2)=0$ if $X(t_1)$ and $X(t_2)$ are independent random values. But an independence by the process of cross-sections $X(t_1)$ and $X(t_2)$ does not follow from the condition $k(t_1,t_2)=0$.

Now we will consider some examples.

1. Let $X(t)$ be a random process of following type

$$X(t) = U\cos\omega t + V\sin\omega t,$$

where $\omega > 0$ is a constant number, and U and V are two independent random values with mathematical expectations $M\{U\} = M\{V\} = 0$ and variances $D\{U\}=D\{V\}=D$. Then for each t

$$m(t) = \cos\omega t M\{U\} + \sin\omega t M\{V\} = 0,$$

so

$$K(t_1,t_2) = M\{(U\cos\omega t_1 + V\sin\omega t_1)(U\cos\omega t_2 + V\sin\omega t_2)\} =$$

$$= \cos\omega t_1 \cos\omega t_2 M\{U^2\} + \sin\omega t_1 \sin\omega t_2 M\{V^2\} +$$
$$+ (\cos\omega t_1 \sin\omega t_2 + \sin\omega t_1 \cos\omega t_2) M\{UV\}.$$

Using the independence of U from V, we have

$$M\{UV\} = M\{U\} \cdot M\{V\} = 0.$$

Therefore,

$$K(t_1, t_2) = D\cos\omega t_1 \cos\omega t_2 + D\sin\omega t_1 \sin\omega t_2 = D\cos\omega(t_1 - t_2).$$

2. We will now generalize the previous example. We will consider the random process

$$X(t) = U\cos\Omega t + V\sin\Omega t,$$

where U, V, and Ω are independent random values, $M\{U\}=M\{V\}=0$, $M\{U^2\} = M\{V^2\} = D$, and a random value Ω is characterized by a distribution density $f(\omega)$. We find

$$m(t) = M\{U\cos\Omega t\} + M\{V\sin\Omega t\}.$$

Taking into account an independence of U from Ω and the independence of V from Ω, we obtain

$$m(t) = M\{U\}M\{\cos\Omega t\} + M\{V\}M\{\sin\Omega t\} = 0.$$

To simplify the calculation of $K(t_1, t_2)$ we note that from the previous example, it follows that the conditional correlation function $K(t_1, t_2|\Omega = \omega)$ is equal to

$$K(t_1, t_2|\Omega = \omega) = M\{X(t_1)X(t_2)|\Omega = \omega\} = D\cos\omega(t_1 - t_2).$$

To find the correlation function $K(t_1, t_2)$ it is necessary to multiply this expression using an element of probability $f(\omega)d\omega$ and to integrate all possible values of frequency ω.[1] Thus,

$$K(t_1, t_2) = D\int_0^\infty f(\omega)\cos\omega(t_1 - t_2)d\omega.$$

[1] The simplification for the calculation of $K(t_1, t_2)$, of course, needs a justification. Lowering the level of proof, we note that this simplification follows on from the mathematical expectation properties and from the total probability formula.

For example, if the random value Ω has the Cauchy distribution, i.e.,

$$f(\omega) = \begin{cases} \frac{2\lambda}{\pi}\frac{1}{\lambda^2+\omega^2} & \text{if } \omega \geq 0, \\ 0 & \text{if } \omega < 0, \end{cases}$$

where λ is some positive number, then

$$K(t_1, t_2) = \frac{2D\lambda}{\pi}\int\limits_0^\infty \frac{\cos\omega(t_1 - t_2)}{\lambda^2+\omega^2}d\omega = D\exp\{-\lambda|t_1 - t_2|\}.$$

3. We will now consider one more example. Let λ be a constant positive number, and $t_1, t_2, \ldots$ be a random sequence of points on axis t such that

1) the probability $P_n(T)$ that a time interval duration T contains exactly n points is equal to
$$P_n(T) = \frac{(\lambda T)^n}{n!}\exp\{-\lambda T\}$$
and does not depend on the provision of this interval on a time-base;
2) if the intervals do not intersect, the corresponding numbers of points are independent random values.

In this case, let us say that this sequence of points $t_1, t_2, \ldots$ forms the Poisson stream with a constant density λ. We suppose that a random process $X(t)$ is defined by its realizations as follows:

$$x(t) = \begin{cases} 0 & \text{at} \quad -\infty < t < t_1, \\ x_1 & \text{at} \quad t_1 \leq t < t_2, \\ x_2 & \text{at} \quad t_2 \leq t < t_3, \\ x_3 & \text{at} \quad t_3 \leq t < t_4, \\ \text{and so on,} & \end{cases}$$

where numbers $x_1, x_2, x_3, \ldots$ are realizations of independent random values $X_1, X_2, X_3, \ldots$ with zero mathematical expectations and with identical variances D.

It is clear that $M\{X(t)\} \equiv 0$ because $M\{X_1\} = M\{X_2\} = M\{X_3\} = \ldots = 0$. We find the correlation function $K(t, t')$:

$$K(t, t') = M\{X(t)X(t')\} =$$
$$= M\{X(t)X(t')|A\}P\{A\} + M\{X(t)X(t')|B\}P\{B\},$$

where the event A means that the interval (t, t') contains at least one of the points $t_1, t_2, t_3, \dots$, and the event B means that the interval (t, t') does not contain points $t_1, t_2, t_3, \dots$, i.e., $B = \bar{A}$. Taking into account the independence of random values $X_1, X_2, X_3, \dots$, we have $M\{X_i X_j\} = M\{X_i\}M\{X_j\} = 0$ for $i \neq j$, so

$$M\{X(t)X(t')|A\} = 0.$$

If there was the event B, the value $X(t)$ coincides with the value $X(t')$ and, therefore,

$$M\{X(t)X(t')|B\} = M\{X^2(t)\} = D.$$

Since the probability $P\{B\}$ is $P_n(T)$ when $n = 0$ and $T = |t - t'|$, we obtain

$$K(t, t') = D \exp\{-\lambda |t - t'|\}.$$

Note that the correlation function from example 3 is equal to the correlation function from example 2. At the same time, the realizations of random processes have a different nature in these examples: realizations are sinusoids in example 2, and realizations are step functions in example 3. Therefore, the same correlation function can correspond to random processes having a various nature of realizations.

1.4 Types of Convergence in a Probabilistic Space and Characteristics of Smoothness in a Random Process

Several types of convergence are considered in the theory of random processes and respectively various definitions are introduced for a continuity and a differentiability of random processes.

Let $X_1, X_2, \dots$ be a sequence of random values defined on some probabilistic space; let X be one more random value defined on this space. We define the following three main types of convergence of sequence $X_n, n{=}1, 2, \dots$, to X as $n \to \infty$.

1. *Convergence with probability* 1 (other names are "almost everywhere convergence" and "almost sure convergence"). A sequence of random values $X_n, n = 1, 2, \dots$, converges to a random value X with probability 1 if

$$P\{\omega\colon \lim_{n\to\infty} X_n(\omega) = X(\omega)\} = 1.$$

This requirement is shorter when it is written down as follows: $P\{X_n \to X\} = 1$, or $X_n \xrightarrow{\text{a.e.}} X$, or $X_n \xrightarrow{\text{a.s.}} X$, where reductions a.e. or a.s. mean, respectively, almost everywhere or almost sure.

2. *Convergence in the mean-square.* A sequence of random values $X_n, n = 1, 2, \dots$, converges to a random value X in the mean-square if

$$\lim_{n\to\infty} M\{(X_n - X)^2\} = 0.$$

Write $X_n \xrightarrow{\text{m.s.}} X$.

3. *Convergence on probability.* A sequence of random values $X_n, n = 1, 2, \dots$, converges to a random value X on probability if for any given $\varepsilon > 0$

$$\lim_{n\to\infty} \delta_n = 0, \quad \text{where} \quad \delta_n = P\{|X_n - X| > \varepsilon\}.$$

Write $X_n \xrightarrow{\text{P}} X$.

A common feature of all three definitions can be formulated as follows: for enough big n the random value X_n and the limit random value X are close in a certain probabilistic sense.

We show that a convergence on probability is the weakest type of convergence, i.e., both a convergence in the mean-square and a convergence with probability 1 involve a convergence on probability. The first of these statements follows from Chebyshev's inequality: if for a random value Y there exists $M\{|Y|^k\}$, where $k > 0$, then for any fixed $\varepsilon > 0$

$$P\{|Y| \geq \varepsilon\} \leq \frac{M\{|Y|^k\}}{\varepsilon^k}.$$

We prove this inequality, for example, when Y is a continuous random value with a distribution density $f(y)$. In this case,

$$M\{|Y|^k\} = \int\limits_{-\infty}^{\infty} |y|^k f(y)dy \geq \int\limits_{-\infty}^{-\varepsilon} |y|^k f(y)dy + \int\limits_{\varepsilon}^{\infty} |y|^k f(y)dy \geq$$

$$\geq \int\limits_{-\infty}^{-\varepsilon} \varepsilon^k f(y)dy + \int\limits_{\varepsilon}^{\infty} \varepsilon^k f(y)dy =$$

$$= \varepsilon^k \left(\int\limits_{-\infty}^{-\varepsilon} f(y)dy + \int\limits_{\varepsilon}^{\infty} f(y)dy \right) = \varepsilon^k P\{|Y| \geq \varepsilon\},$$

and Chebyshev's inequality is proved. Now having put $k = 2$ and $Y = X_n - X$, we obtain

$$P\{|X_n - X| \geq \varepsilon\} \leq \frac{M\{(X_n - X)^2\}}{\varepsilon^2}.$$

Therefore, if $M\{(X_n - X)^2\} \to 0$ as $n \to \infty$, then $P\{|X_n - X| > > \varepsilon\} \to 0$ as $n \to \infty$, i.e., a convergence on probability follows from a convergence in the mean-square.

Now we prove that a convergence with probability 1 also involves a convergence on probability. We will choose any $\varepsilon > 0$ and we will consider the event

$$A_\varepsilon = \{\omega : \quad \exists N = N(\omega) \quad \forall n \geq N \quad |X_n(\omega) - X(\omega)| < \varepsilon\}.$$

Clearly, if a convergence of sequence of random values $X_n, n = 1, 2, \ldots,$ to a random value X takes place with probability 1, then $P\{A_\varepsilon\} = 1$. Opposite to the event A_ε, the event $\overline{A_\varepsilon}$ is

$$\overline{A_\varepsilon} = \{\omega : \quad \forall N \quad \exists n \geq N \quad |X_n(\omega) - X(\omega)| \geq \varepsilon\}; \quad P\{\overline{A_\varepsilon}\} = 0.$$

We will now introduce the event

$$B_{\varepsilon,N} = \{\omega : \quad \exists n \geq N \quad |X_n(\omega) - X(\omega)| \geq \varepsilon\}.$$

Then

$$\overline{A_\varepsilon} = B_{\varepsilon,1} B_{\varepsilon,2} B_{\varepsilon,3} \ldots = \prod_{k=1}^{\infty} B_{\varepsilon,k},$$

and

$$B_{\varepsilon,1} \supset B_{\varepsilon,2} \supset \ldots \supset B_{\varepsilon,k} \supset B_{\varepsilon,k+1} \supset \ldots \quad .$$

In this situation, obviously, the following presentation takes place:

$$B_{\varepsilon,1} = \overline{A_\varepsilon} + B_{\varepsilon,1}\overline{B_{\varepsilon,2}} + B_{\varepsilon,2}\overline{B_{\varepsilon,3}} + \ldots + B_{\varepsilon,k}\overline{B_{\varepsilon,k+1}} + \ldots =$$

$$= \overline{A_\varepsilon} + \sum_{k=1}^{\infty} B_{\varepsilon,k}\overline{B_{\varepsilon,k+1}}.$$

Since any two summands in the right part of this equality are non-joint events, we have

$$P\{B_{\varepsilon,1}\} = P\{\overline{A_\varepsilon}\} + \sum_{k=1}^{\infty} P\{B_{\varepsilon,k}\overline{B_{\varepsilon,k+1}}\}.$$

By definition, put $S_N = \sum\limits_{k=1}^{N-1} P\{B_{\varepsilon,k}\overline{B_{\varepsilon,k+1}}\}$. Then

$$P\{B_{\varepsilon,1}\} = P\{\overline{A_\varepsilon}\} + \lim_{N\to\infty} S_N.$$

Since $B_{\varepsilon,k+1} \subset B_{\varepsilon,k}$, we get

$$P\{B_{\varepsilon,k}\overline{B_{\varepsilon,k+1}}\} = P\{B_{\varepsilon,k}\} - P\{B_{\varepsilon,k+1}\} \quad \text{for every} \quad k = 1, 2, \dots .$$

Therefore, $S_N = P\{B_{\varepsilon,1}\} - P\{B_{\varepsilon,N}\}$ and

$$P\{B_{\varepsilon,1}\} = P\{\overline{A_\varepsilon}\} + (P\{B_{\varepsilon,1}\} - \lim_{N\to\infty} P\{B_{\varepsilon,N}\}).$$

From this it follows that

$$\lim_{N\to\infty} P\{B_{\varepsilon,N}\} = P\{\overline{A_\varepsilon}\} = 0.$$

Now we introduce the event

$$C_{\varepsilon,N} = \{\omega : \quad |X_N(\omega) - X(\omega)| \geq \varepsilon\}.$$

Then $C_{\varepsilon,N} \subset B_{\varepsilon,N}$ and

$$0 \leq P\{C_{\varepsilon,N}\} \leq P\{B_{\varepsilon,N}\}.$$

Passing in the last double inequality to a limit as $N \to \infty$, we obtain $\lim_{N\to\infty} P\{C_{\varepsilon,N}\} = \lim_{N\to\infty} P\{|X_N - X| \geq \varepsilon\} = 0$. This completes the proof. Thus, a convergence with probability 1 as well as convergence in the mean-square involves a convergence on probability.

Various determinations of convergence in probabilistic space lead to various understandings of the continuity and differentiability of a random process. According to the determinations of convergence, as well as the continuity and differentiability of a random process, it is possible to understand it as follows: a) a continuity and a differentiability with probability 1, b) a continuity and a differentiability in the mean-square, c) a continuity and a differentiability on probability. We give, for example, definitions of continuity and differentiability in the mean-square (m.s.).

If a random process $X(t)$ satisfies the condition

$$\lim_{t\to t_0} M\{(X(t) - X(t_0))^2\} = 0,$$

then we say that $X(t)$ is continuous in the mean-square at point $t = t_0$. If this condition is satisfied for all points of some interval (a, b), then process $X(t)$ is called continuous in the mean-square on the interval (a, b).

If a random process $X(t)$ and a random value Y satisfy the condition

$$\lim_{t\to t_0} M\left\{\left(\frac{X(t)-X(t_0)}{t-t_0} - Y\right)^2\right\} = 0,$$

then process $X(t)$ is called differentiable in the mean-square at point t_0, and random value Y is called a derivative in the mean-square of process $X(t)$ at point t_0, and in this case Y is denoted by $X'(t_0)$.

Let $K(t, u)$ be a correlation function of process $X(t)$. We find the sufficient condition of continuity in the mean-square of process $X(t)$ in terms of correlation function. We have

$$M\{(X(t_0+h) - X(t_0))^2\} = M\{X(t_0+h)X(t_0+h)\} -$$

$$-2M\{X(t_0+h)X(t_0)\} + M\{X(t_0)X(t_0)\}.$$

Let $m(t) = M\{X(t)\}$. Since for any t' and t''

$$K(t', t'') = M\{X(t')X(t'')\} - m(t')m(t''),$$

we obtain

$$M\{(X(t_0+h) - X(t_0))^2\} = K(t_0+h, t_0+h) + m(t_0+h)m(t_0+h) \ -$$

$$- \ 2K(t_0+h, t_0) - 2m(t_0+h)m(t_0) + K(t_0, t_0) + m(t_0)m(t_0).$$

From this it follows that if $m(t)$ is continuous at point t_0 and $K(t, u)$ is continuous at point $t = u = t_0$, then the process $X(t)$ is continuous in the mean-square at point t_0. It is possible to show that this condition is also necessary: a continuity of mathematical expectation $m(t)$ at point t_0 and a continuity of the correlation function $K(t, u)$ at point $t = u = t_0$ follow from a continuity in the mean-square of the process $X(t)$ at point t_0.

We note that a continuity in the mean-square of process $X(t)$ does not mean a continuity with probability 1. For example, sample functions of process from example 3 from the previous section are step functions and, therefore, are discontinuous functions with probability 1. Nevertheless, this process is continuous in the mean-square at any point because its mathematical expectation $m(t)=0$ and correlation function $K(t,u) = D\exp\{-\lambda|t-u|\}$ are continuous.

A necessary and sufficient condition for differentiability in the mean-square of the process $X(t)$ at point t_0 can also be formulated in terms of mathematical expectation $m(t)$ and correlation function $K(t,u)$ of process $X(t)$. It is possible to show that such condition is an existence of derivatives

$$\left.\frac{dm(t)}{dt}\right|_{t=t_0} \quad \text{and} \quad \left.\frac{\partial K^2(t,u)}{\partial t \partial u}\right|_{t=u=t_0}.$$

As well as in the case of continuity in the mean-square, an existence of derivative in the mean-square does not mean that the sample functions of the process are differentiated in the usual sense.

At last, we will formulate one more result. Let the random process $X(t)$ have a mathematical expectation $m(t)$ and a correlation function $K(t_1,t_2)$. If the process $X(t)$ is differentiated in the mean-square and its derivative in the mean-square is equal to $X'(t)$, then formulas

$$m_1(t) = \frac{dm(t)}{dt} \quad \text{and} \quad K_1(t_1,t_2) = \frac{\partial^2 K(t_1,t_2)}{\partial t_1 \partial t_2}$$

define a mathematical expectation $m_1(t)$ and a correlation function $K_1(t_1,t_2)$ of the process $X'(t)$. Formally this result can be obtained if we change a sequence of performance for operations of mathematical expectation and differentiability:

$$m_1(t) = M\{X'(t)\} = M\left\{\frac{d}{dt}X(t)\right\} = \frac{d}{dt}M\{X(t)\} = \frac{dm(t)}{dt},$$

$$K_1(t_1,t_2) = M\{(X'(t_1)-m_1(t_1))(X'(t_2)-m_1(t_2))\} =$$
$$= M\left\{\left.\frac{d}{dt}(X(t)-m(t))\right|_{t=t_1}\left.\frac{d}{dt}(X(t)-m(t))\right|_{t=t_2}\right\} =$$
$$= M\left\{\frac{\partial^2}{\partial t_1 \partial t_2}(X(t_1)-m(t_1))(X(t_2)-m(t_2))\right\} =$$
$$= \frac{\partial^2}{\partial t_1 \partial t_2}M\{(X(t_1)-m(t_1))(X(t_2)-m(t_2))\} = \frac{\partial^2 K(t_1,t_2)}{\partial t_1 \partial t_2}.$$

If we are interested in a mutual correlation function $K_{01}(t_1, t_2)$ of processes $X(t)$ and $X'(t)$, where by definition

$$K_{01}(t_1, t_2) = M\{(X(t_1) - m(t_1))(X'(t_2) - m_1(t_2))\},$$

then the same formala leads to the result

$$K_{01}(t_1, t_2) = \frac{\partial K(t_1, t_2)}{\partial t_2}.$$

This result is also fair and can be proved mathematically.

1.5 Stationary Random Processes

In many appendices it is necessary to study functions determined by casual factors when the behavior of these factors are more or less constant throughout an observation period. Considering these functions, $x(t)$, as realizations of a random process, $X(t)$, we ask what is the definition of the process property which characterizes this situation. Such property is defined as the property of invariancy from all the finite-dimensional distributions of process relative to shifts of time t: i.e., when for any n an joint n-dimensional distribution of random values

$$X(t_1 + \tau), \ldots, X(t_n + \tau)$$

does not depend on τ for any finite sequence of points $t_1, \ldots, t_n$. Such a process is called a stationary procces.

Let us consider the case when densities of distributions exist. Let $f(x_1, \ldots, x_n; t_1, \ldots, t_n)$ be a density of joint distribution of random values $X(t_1), \ldots, X(t_n)$, and let $f(x_1, \ldots, x_n; t_1 + \tau, \ldots, t_n + \tau)$ be a density of joint distribution of random values $X(t_1+\tau), \ldots, X(t_n+\tau)$. If process $X(t)$ is stationary, then

$$f(x_1, \ldots, x_n; t_1, \ldots, t_n) = f(x_1, \ldots, x_n; t_1 + \tau, \ldots, t_n + \tau)$$

for any τ. In particular, for $n = 1$

$$f(x_1; t_1) = f(x_1; 0),$$

i.e., a distribution density of random value $X(t_1)$ does not depend on t_1. In this case,

$$m(t) = M\{X(t)\} = M\{X(0)\} = const$$

if a mathematical expectation of process $X(t)$ exists.

Furthermore, for $n = 2$

$$f(x_1, x_2; t_1, t_2) = f(x_1, x_2; 0, t_2 - t_1),$$

i.e., a distribution density of system of random values $X(t_1), X(t_2)$ depends only on the difference $t_2 - t_1$. Therefore, all characteristics of system $\{X(t_1), X(t_2)\}$, in particular, the correlation function $K(t_1, t_2)$, do not separately depend on t_1 and t_2. These characteristics are completely defined (if, of course, they exist) by the difference $t_2 - t_1$. In particular, $K(t_1, t_2) = K(\tau)$, where $\tau = t_2 - t_1$.

A stationary random process is also called strictly stationary, or stationary in a narrow sense. Besides, there is the concept of stationarity in a wide sense: a random process is called wide-sense stationary if its mathematical expectation is constant and its correlation function depends only on the difference of arguments. For example, all processes considered in examples from section 1.3 are known as wide-sense stationary. As shown above, a wide-sense stationarity follows from a strict stationarity, if both the mathematical expectation and the correlation function of the random process exist.

We claim that the function $K(\tau)$ is even. Indeed, if $K(t_1, t_2) = K(t_2 - t_1) = K(\tau)$, then $K(t_2, t_1) = K(t_1 - t_2) = K(-\tau)$. It is obvious that $K(t_1, t_2) = K(t_2, t_1)$. Therefore, $K(\tau) = K(-\tau)$. Thus for a stationary process it is possible to write $K(t_1, t_2) = K(\tau)$, where $\tau = |t_2 - t_1|$.

If a random process is stationary, the conditions of continuity and differentiability in the mean-square become simpler. A continuity in function $K(\tau)$ at $\tau = 0$ is required for a continuity in the mean-square and an existence of the second derivative $K''(0)$ is required for differentiability in the mean-square. For example, the process from example 1 of section 1.3 has the correlation function $K(\tau) = D\cos\omega\tau$ and this process is continuous and differentiate in the mean-square. The processes from examples 2 and 3 have the correlation function $K(\tau) = D\exp\{-\lambda|\tau|\}$ and these processes are continuous in the mean-square, but are not differentiated in the mean-square.

Stationary random processes are encountered in practice quite often. By means of such processes, it is possible to model, for example, a random noise in radio sets; fluctuations of tension in a lighting network; pitching in a ship; and fluctuations in the height of airplane at cruising horizontal flight. As a rule, the change of phase coordinates of any stochastic system is described by a stationary random process, if this system functions in steady mode.

Let us consider an example from a bank activity. We will assume that the flow of deposits to the bank is described by the Poisson process with a density of λ (see example 3 from section 1.3). The value λ is the mathematical expectation of a number of deposits that arrive during a unit interval of time. We denote by $F(x)$ the distribution function of the duration of the contribution. Let r be the percentage rate on deposits. We use the formula of continuous percents: the capital M after t units of time is equal to $M \exp\{rt\}$. For simplicity, all contributions have the same size m. We find the total capital at moment t, if this capital is equal to zero at the initial moment $t_0 = 0$. For this purpose, we will consider a partition of the interval (t_0, t) by intermediate moments s_i:

$$t_0 = s_0 < s_1 < s_2 < \ldots < s_{i-1} < s_i < \ldots < s_{n-1} < s_n = t.$$

Let n be rather great and all $\Delta s_i = s_i - s_{i-1}$ be rather small. The average contribution made during the period $[s_{i-1}, s_i)$ is equal to $m\lambda\Delta s_i$. At the moment t this contribution will be more than $m\lambda\Delta s_i \exp\{r(t - s_i)\}$ and less than $m\lambda\Delta s_i \exp\{r(t - s_{i-1})\}$. There exists a point $\xi_i \in (s_{i-1}, s_i)$ such that this contribution is equal to $m\lambda\Delta s_i \exp\{r(t - \xi_i)\}$. Let X_i be the capital that remained on the account at the moment t from the contribution $m\lambda\Delta s_i \exp\{r(t-\xi_i)\}$, which arrived during the interval $[s_{i-1}, s_i)$. Then the total capital on the account at the moment t is

$$X(t) = \sum_{i=1}^{n} X_i,$$

and the mathematical expectation of this capital is

$$M\{X(t)\} = \sum_{i=1}^{n} M\{X_i\}.$$

For the calculation $M\{X_i\}$ at small Δs_i, we consider approximately that the entire contribution $m\lambda\Delta s_i \exp\{r(t - \xi_i)\}$ is made at moment ξ_i and X_i is equal to $m\lambda\Delta s_i \exp\{r(t - \xi_i)\}$ with the probability p_i or X_i is equal to 0 with probability $1 - p_i$, where p_i is the probability that a duration of contribution made at moment ξ_i is equal to not less than $(t - \xi_i)$ time units:

$$p_i = P\{\text{duration} \geq t - \xi_i\} =$$

$$= 1 - P\{\text{duration} < t - \xi_i\} = 1 - F(t - \xi_i).$$