CONTENT

1. Random variable and random process. Classes of random processes (Overview)

1.1. **Random variable.** Denote by $\Omega = \{\omega\}$ a sample space. A function[1] $\xi(\omega)$, $\omega \in \Omega$:

$$\Omega \xrightarrow{\xi} \mathbb{R}$$

is called random variable. Typically, a description of the random variable $\xi(\omega)$ is given in term of its distribution function

$$F(x) = P(\omega : \xi(\omega) \le x),\ x \in \mathbb{R},$$

where P, named probabilistic measure, is a function transforming subsets of Ω to the interval $[0,1]$:

$$\Omega \xrightarrow{P} [0,1]$$

and is such that $P(\Omega) = 1$, $P(\varnothing) = 0$.

1.2. **Random vector.** A vector

$$\xi(\omega) = \big(\xi_1(\omega), \dots, \xi_n(\omega)\big).$$

with random variables as its entries is called the random vector. Also, the random vector is characterized by distribution function:

$$F(x_1, \cdots, x_n) = P(\omega : \xi_1(\omega) \le x_1, \cdots, \xi_n \le x_n).$$

[1] more exactly, measurable function w.r.t. some σ-algebra

1.3. **Random processes.** A random vector

$$\xi(\omega) = \big(\xi_1(\omega), \dots, \xi_n(\omega), \dots\big)$$

might be interpreted as random sequence (process), if $k = 1, \dots, n, \dots$ are understood as time moments. Formally, in a case of random sequence the distribution function of countable values of arguments is considered

$$F(x_1, \cdots, x_n, \cdots,) = P(\omega : \xi_1(\omega) \le x_1, \cdots, \xi_n \le x_n, \cdots,).$$

For a continuous time case, a family of random variables $\xi_t(\omega)$ parametrized by $t \ge 0$ or $-\infty < t < \infty$, is called continuous time random process. For fixed t_0, $\xi_{t_0}(\omega)$ is the random variable while for fixed ω_0, a function $\xi_t(\omega_0)$ of the argument t is called trajectory (or a path) of the random process $\xi_t(\omega)$. This trajectory might be continuous or discontinuous function. If all paths of a random process are continuous, we say shortly "continuous process". Under consideration of continuous time process, so called, finite dimensional distributions are introduces:

$$F_{t_1, \cdots, t_n}(x_1, ..., x_n) = P(\omega : \xi_{t_1}(\omega) \le x_1, \cdots, \xi_{t_n}(\omega) \le x_n)$$

for every $n \ge 1$ and $t_1, \cdots, t_n$.

Remark 1.1. For sake of a notational simplicity, henceforth the index "ω" will be omitted.

Brownian process $(B_t)_{t\ge 0}$ is the typical example of random process with continuous paths. Its mathematical description was done by N. Wiener and in the sequel notation $(W_t)_{t\ge 0}-$"Wiener process" will be used as well. A short description of Wiener process is the following:

1. $W_0 = 0$ and paths are continuous functions;
2. the expectation $\boldsymbol{E} W_t \equiv 0$;
3. the correlation function $W_t W_s = \min(t, s)$;
4. any finite dimensional distributions are Gaussian.

Poisson process $(\pi_t)_{t\ge 0}$ is the typical example of random process with discontinuous paths. Its short description is the following:

1. $\pi_0 = 0$ and paths are piece-wise constant functions having jumps of the unite size;
2. the expectation $\boldsymbol{E} \pi_t \equiv t$;
2. for non-overlapping time intervals increments of $(\pi_t)_{t\ge 0}$ are independent random variables having the Poisson distribution.

Random processes, considered as mathematical objects, are studied under different frameworks caused different mathematical theories for their study. We mention here *Linear Theory* when only linear transformations of the random process paths are available. In the context of that theory, the expectation and correlation function are main tools for analysis. An important class of random process successful served by the linear theory is the class of

Random Processes in the wide sense. Any process, X_n, from this class has the expectation $\boldsymbol{E}X_n \equiv$ const.$(= 0$ for notational convenience) and the correlation function $K(n-m) = \boldsymbol{E}X_nX_m$ of two arguments (n,m) being in reality the function of their difference $n-m$. Under $\sum\limits_n |K(n)| < \infty$, the Fourier transform defines, so called, spectral density (here $\imath = \sqrt{-1}$)

$$f(\lambda) = \sum_n e^{\imath\lambda n}K(n), \ \lambda \in \mathbb{R}$$

the main tool for *Wiener filter* theory.

An application of the linear theory is natural for

Gaussian random process since any finite dimensional distributions of which completely characterized by the expectation and correlation functions, so that under linear transformation, only these function are replaced by new ones while Gaussian structure of the distribution is preserved.

A class of random processes being beyond the range of linear theory but still in the framework of a stationary notion is a class of

Stationary random processes in the strict sense. Random processes from that class are not assumed to obey the second moment (so they are not obligatory stationary processes in the wide sense). Finite dimensional distributions of processes from that class possess the following characterization: finite dimensional distribution is preserved under any shift of the time parameter

$$P\Big(X_{n_1} \le x_1, \ldots, X_{n_k} \le x_k\Big) \equiv P\Big(X_{n_1+m} \le x_1, \ldots, X_{n_k+m} \le x_k\Big).$$

Processes with independent and identically distributed values forms its simple subclass:

$$P\Big(X_{n_1} \le x_1, \ldots, X_{n_k} \le x_k\Big) \equiv \prod_{j=1}^{k} P\Big(X_{n_1} \le x_j\Big)$$

If $\boldsymbol{E}X_n \equiv 0$ and $\boldsymbol{E}X_n^2 \equiv 1$ we obtain further subclass of

White noises. Every white noise is also the stationary process in the wide sense.

Beyond of the Linear theory, there are various applications, for instance, *Non linear Filtering*, where a crucial role plays a class of

Markov processes. In contrast to processes with independent values, when the joint distribution function $P\Big(X_1 \le x_1, \ldots, X_k \le x_k\Big)$ is split into the product of marginal distributions: $\prod_{j=1}^k P\Big(X_j \le x_j\Big)$, in

Markov case

$$
\begin{aligned}
&P\Big(X_1 \le x_1, \dots, X_k \le x_k\Big) \\
&\qquad = \int_{\mathbb{R}} \dots \int_{\mathbb{R}} P\Big(X_k \le x_k \big| X_{k-1} = x_{k-1}\Big) \\
&\qquad \times dP\Big(X_{k-1} \le x_{k-1} \big| X_{k-2} = x_{k-2}\Big) \\
&\qquad\qquad \times dP\Big(X_2 \le x_2 \big| X_1 = x_1\Big) dP\Big(X_1 \le x_1\Big).
\end{aligned}
$$

That formula provides simpler structure, if the distribution function obeys a density

$$
\frac{\partial^k P\Big(X_1 \le x_1, \dots, X_k \le x_k\Big)}{\partial x_1 \dots \partial x_k} = p(x_1, x_2, ..., x_k),
$$

and in terms of conditional densities is presented as:

$$
p(x_1, x_2, ..., x_k) = p_{k_1}(x_1) \prod_{i=2}^{n} p(x_i | x_{i-1}).
$$

In a modern part of *Theory of Random processes*, named *Stochastic Calculus* the main class of random processes is the class of
Semimartingales.

Appendix A. Random process in the wide sense

Random processes in the wide sense are extensively studied in many undergraduate courses. So, here we recall only the main properties of correlation function and spectral density and give some examples.

A.1. **Definitions.** Set time values $n = \cdots, -1, 0, 1, \cdots$ A random process (X_n) with

1. $\boldsymbol{E}X_n^2 \equiv \boldsymbol{E}X_0^2 < \infty$ ($EX_0^2 >$ is assumed);
2. $\boldsymbol{E}X_n \equiv 0$;
3. $K(n-m) = \boldsymbol{E}X_nX_m$

is said to be stationary in the wide sense. Here

- $K(0)$ is the variance
- $K(n)$ is the correlation function
- $\rho(n) = \frac{K(n)}{K(0)}$ is the correlation coefficient.

We emphasize a few properties of the correlation functions K.

1. K is nonnegative definite function: for any numbers $a_1, ..., a_m$ and $\ell_1, ..., \ell_m$
$$
\sum_{i,j=1}^{m} a_i a_j K(\ell_i - \ell_j) \ge 0. \tag{1.1}
$$
2. $K(-n) = K(n)$ and $|K(n)| \le K(0)$.

If

$$\sum_{n=-\infty}^{\infty} |K(n)| < \infty, \tag{1.2}$$

the Fourier transform of $K(n)$

$$f(\lambda) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} e^{-\imath\lambda\ell} K(\ell), \ \lambda \in [-\pi, \pi] \tag{1.3}$$

is well defined and determine the spectral density function. The spectral density $f(\lambda)$ possesses the following properties.

1. the is the spectral density is real (not complex) nonnegative function with
$$\int_{-\pi}^{\pi} f(\lambda) d\lambda < \infty;$$
2. $\int_{-\pi}^{\pi} e^{\imath\lambda n} f(\lambda) d\lambda = K(n)$;

An increasing function $F(\lambda)$ with $dF(\lambda) = f(\lambda)d\lambda$ is called the spectral measure. If (1.2) is not valid, an existence of the spectral density is problematic. Nevertheless, the spectral measure $F(\lambda)$ always exists and any correlation function K obey a spectral decomposition with appropriate F (Herglotz theorem):

$$K(n) = \int_{-\pi}^{\pi} e^{\boldsymbol{i}\lambda n} dF(\lambda). \tag{1.4}$$

A.2. **Examples.**

1. White noise. $f(\lambda) = \frac{1}{2\pi}$.
2. Moving average.

$$X_n = \sum_{k=-\infty}^{\infty} a_k \varepsilon_{n-k} \tag{1.5}$$

is the linear transformation of the white noise (ε_k), where a_k are numbers such that $\sum\limits_{k=-\infty}^{\infty} |a_k|^2 < \infty$.

3. One side moving average. $X_n = \sum_{k=0}^{\infty} a_k \varepsilon_{n-k}$.
4. Moving average of order $p > 0$. $X_n = \sum_{k=0}^{p} a_k \varepsilon_{n-k}$.
5. Autoregression.

$$X_n + b_1 X_{n-1} + ... + b_q X_{n-q} = \varepsilon_n, \tag{1.6}$$

where parameters $b_1, ..., b_q$ are such that the roots of the polynomial

$$Q(z) = 1 + b_1 z^{-1} + ... + b_q z^{-q}$$

lie within of the unit circle. The spectral density

$$f(\lambda) = \frac{1}{2\pi} \frac{1}{|Q(e^{\imath\lambda})|^2}.$$

6. Autoregression and moving average.

$$X_n + b_1 X_{n-1} + ... + b_q X_{n-q} = a_0 \varepsilon_n + a_1 \varepsilon_{n-1} + ... + a_p \varepsilon_{n-p}, \qquad (1.7)$$

where the roots of the polynomial $Q(z) = 1 + b_1 z^{-1} + ... + b_q z^{-q}$ lie within the unit circle. The spectral density is defined in terms of polynomials $Q(z)$ and $P(z) = a_0 + a_1 z^{-1} + ... + a_p z^{-p} p$ as:

$$f(\lambda) = \frac{1}{2\pi} \frac{\left|P(e^{\iota\lambda})\right|^2}{\left|Q(e^{\iota\lambda}\right|^2}, \qquad (1.8)$$

The spectral density of the type given in (1.8) is called *rational spectral density*.

Remark A.1. A linear transformation of stationary sequence not obligatory preserves stationarity.