Gaussian vectors

Definition (Gaussian distribution): A real random variable Z is said to have the *standard Gaussian distribution* if its probability distribution is absolutely continuous with respect to the Lebesgue measure and its density is given by

$$f_{\mathcal{N}(0,1)}(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}.$$

(In this case, we write $Z \sim \mathcal{N}(0,1)$.) Let $m, n \in \mathbb{N}$. An n-dimensional random vector X is said to have the *n-dimensional Gaussian distribution* if there exists $\mu \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n \times m}$ such that $X = \mu + AZ$ where Z is an m-dimensional random vector whose components are independent and have the standard Gaussian distribution. (In this case, we write $X \sim \mathcal{N}_n(\mu, \Sigma)$ where $\Sigma = AA^\top$.)

Corollaries: Let Z and X be random vectors as in the above definition. Then

1. It holds that $\mathbb{E}X = \mu$ (vector of means) and $\operatorname{var} X = \Sigma$ (covariance matrix).
2. For $k \in \mathbb{N}$ and matrix $B \in \mathbb{R}^{k \times n}$, it holds that $BX \sim \mathcal{N}_k(B\mu, B\Sigma B^\top)$.
3. If Σ is non-singular, then the distribution of X is absolutely continuous with respect to the n-dimensional Lebesgue measure and its density is given by

$$f_{\mathcal{N}_n(\mu,\Sigma)}(x) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det \Sigma}} \mathrm{e}^{-\frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu)}, \quad x \in \mathbb{R}^n.$$

Stochastic processes

Recall the definition of a distribution function.

Definition (Distribution function): Let $k \in \mathbb{N}$. A function $F : \mathbb{R}^k \to [0,1]$ is a *distribution function* if it has the following properties:

1. For every $x \in \mathbb{R}^k$ there is a sequence $\{x^n\}_{n\in\mathbb{N}} \subset \mathbb{R}^k$ such that $x^n > x$ (coordinatewise) for every $n \in \mathbb{N}$, $\lim x^n = x$, and
$$\lim_{n\to\infty} F(x^n) = F(x).$$
2. There exists a sequence $\{y^n\}_{n\in\mathbb{N}} \subset \mathbb{R}^k$ such that $\lim y^n = (\infty, \dots, \infty)^\top$ and
$$\lim_{n\to\infty} F(y^n) = 1.$$
3. For every $i \in \{1, \dots, k\}$ and every $x = (x_1, \dots, x_k)^\top \in \mathbb{R}^k$ we have that
$$\lim_{y\to-\infty} F(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_k) = 0.$$
4. For every $x = (x_1, \dots, x_k)^\top \in \mathbb{R}^k$ and $y = (y_1, \dots, y_k)^\top \in \mathbb{R}^k$, $x < y$, it holds that
$$\sum_{\substack{\delta_1 \in \{x_1, y_1\} \\ \vdots \\ \delta_k \in \{x_k, y_k\}}} (-1)^{\mathrm{card}(\{i : \delta_i = x_i\})} F(\delta_1, \dots, \delta_k) \ge 0.$$

Let $T \subset \mathbb{R}$, $T \neq \emptyset$. A collection of (real-valued) random variables $\{X_t, t \in T\}$ is called a (real-valued) *stochastic process*. Given such a process, the system

$$\{\mathsf{F}^X_{t_1,\dots,t_n}, n \in \mathbb{N}, t_1, \dots, t_n \in T\}$$

of its marginal distribution functions defined by

$$\mathsf{F}^X_{t_1,\dots,t_n}(x_1, \dots, x_n) = \mathbb{P}(X_{t_1} \leq x_1, \dots, X_{t_n} \leq x_n), \qquad x_1, \dots, x_n \in \mathbb{R},$$

for $n \in \mathbb{N}$ and $t_1, t_2, \dots, t_n \in T$ can be considered. This system is symmetric and consistent in the following manner: For any $n \in \mathbb{N}$, $t_1, \dots, t_n \in T$ and $x_1, \dots, x_n \in \mathbb{R}$, we have that

1. the equality
$$\mathsf{F}^X_{t_{\pi_1},\dots,t_{\pi_n}}(x_{\pi_1}, \dots, x_{\pi_n}) = \mathsf{F}^X_{t_1,\dots,t_n}(x_1, \dots, x_n)$$
holds for any permutation $\{\pi_1, \dots, \pi_n\}$ of $\{1, \dots, n\}$; and

2. we also have
$$\lim_{x_n \to \infty} \mathsf{F}^X_{t_1,\dots,t_{n-1},t_n}(x_1, \dots, x_n) = \mathsf{F}^X_{t_1,\dots,t_{n-1}}(x_1, \dots, x_{n-1}).$$

In fact, if for any $n \in \mathbb{N}$ and $t_1, \dots, t_n \in T$ we are given a distribution function $\mathsf{F}_{t_1,\dots,t_n}$, we say that the system $\{\mathsf{F}_{t_1,\dots,t_n}, n \in \mathbb{N}, t_1 \dots, t_n \in T\}$ is *consistent* if it has the two properties above. One can always construct a stochastic process from a given system of marginal distributions as long as the system is consistent.

Exercise 1.1: Let $T = \{1, 2\}$. Give an example of

- a consistent system of distribution functions,
- a system of distribution functions that does not satisfy consistency condition 1,
- a system of distribution functions that does not satisfy consistency condition 2.

Theorem 1.1 (Daniell-Kolmogorov): Let $\{\mathsf{F}_{t_1,\dots,t_n}, n \in \mathbb{N}, t_1 \dots, t_n \in T\}$ be a consistent system of distribution functions. Then there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a stochastic process $\{X_t, t \in T\}$ defined on it such that for every $n \in \mathbb{N}$ and $t_1, \dots, t_n \in T$, we have

$$\mathbb{P}(X_{t_1} \leq x_1 \dots, X_{t_n} \leq x_n) = \mathsf{F}_{t_1,\dots,t_n}(x_1, \dots, x_n), \quad x_1, \dots, x_n \in \mathbb{R}.$$

Gaussian processes

Definition (Gaussian process): A stochastic process $\{X_t, t \in T\}$ is called *Gaussian* if for all $n \in \mathbb{N}$ and $t_1, \dots, t_n \in T$, the random vector $(X_{t_1}, \dots, X_{t_n})^\top$ has n-dimensional Gaussian distribution.

Exercise 1.2: Let $\{X_t, t \in \mathbb{Z}\}$ be a stochastic process that consists of independent random variables that have the standard Gaussian distribution. Show that $\{X_t, t \in \mathbb{Z}\}$ is a Gaussian process. Let further $\rho \in \mathbb{R}$ be such that $|\rho| < 1$ and let $\{Y_t, t \in \mathbb{Z}\}$ be a stochastic process defined by

$$Y_t = X_t + \rho X_{t-1}, \quad t \in \mathbb{Z}.$$

Show that $\{Y_t, t \in \mathbb{Z}\}$ is also a Gaussian process.

Mean and autocovariance function

Definition (Mean of a process): Let $\{X_t, t \in T\}$ be a stochastic process with finite first moments, i.e. $\mathbb{E}|X_t| < \infty$ for every $t \in T$. Then the (possibly $\mathbb{C}$-valued) function μ defined by

$$\mu(t) := \mathbb{E}X_t, \quad t \in T,$$

is called the *mean value* of the process $\{X_t, t \in T\}$.

Definition (Autocovariance function): Let $\{X_t, t \in T\}$ be a stochastic process with finite second moments, i.e. $\mathbb{E}|X_t|^2 < \infty$ for all $t \in T$. Then the (possibly $\mathbb{C}$-valued) function defined by

$$R(s,t) = \mathbb{E}(X_s - \mathbb{E}X_s)(\overline{X_t - \mathbb{E}X_t}), \quad s,t \in T,$$

is called the *autocovariance function* of the process $\{X_t, t \in T\}$.

Theorem 2.2: The autocovariance function has the following properties:

- it is non-negative on the diagonal: $R(t,t) \geq 0$,
- it is Hermitian: $R(s,t) = \overline{R(t,s)}$,
- it satisfies the Cauchy-Schwarz inequality: $|R(s,t)| \leq \sqrt{R(s,s)}\sqrt{R(t,t)}$,
- it is positive semidefinite: for all $n \in \mathbb{N}$, $c_1, \ldots, c_n \in \mathbb{C}$ and $t_1, \ldots, t_n \in T$ it holds that

$$\sum_{j=1}^{n}\sum_{k=1}^{n} c_j \overline{c_k} R(t_j, t_k) \geq 0.$$

Stationarity

Definition (Properties of stochastic processes): A stochastic process $\{X_t, t \in T\}$ is called

- *centered* if $\mu(t) = 0$ for every $t \in T$,
- *a process of uncorrelated random variables* if the process has finite second moments and for its autocovariance function it holds that $R(s,t) = 0$ whenever $s,t \in T$ are such that $s \neq t$,
- *process with independent increments* if for all $n \in \mathbb{N}$ and $t_1, \ldots, t_n \in T$ such that $t_1 < \cdots < t_n$ the random variables $X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}$ are independent,
- *process with stationary increments* if for all $s,t \in T$ such that $s < t$ the distribution of the increment $X_t - X_s$ depends only on $t - s$.

Definition (Stationarity): A stochastic process $\{X_t, t \in T\}$ is called

- *strictly stationary* if for any $n \in \mathbb{N}$, $t_1, \ldots, t_n \in T$ and $h > 0$ such that $t_1 + h, \ldots, t_n + h \in T$ the distributions of the random vectors $(X_{t_1}, \ldots, X_{t_n})^\top$ and $(X_{t_1+h}, \ldots, X_{t_n+h})^\top$ are the same;
- *weakly stationary* if the process has
  - finite second moments,
  - constant mean value (i.e. if there is $\mu \in \mathbb{R}$ such that $\mu(t) = \mu$ holds for every $t \in T$), and
  - if its autocovariance function depends only on the difference of its arguments (i.e. if $R(s+h, t+h) = R(s,t)$ holds for every $h \in \mathbb{R}$, $s,t \in T$ such that $s+h$, $t+h \in T$),
- *covariance stationary* if the process has finite second moments and its autocovariance function depends only on the difference of its arguments.

Theorem 2.1: The following implications hold:

a) strictly stationary with finite second moments $\Rightarrow$ weakly stationary,
b) weakly stationary $\Rightarrow$ covariance stationary,
c) covariance stationary and constant mean $\Rightarrow$ weakly stationary,
d) weakly stationary and Gaussian $\Rightarrow$ strictly stationary,
e) process of uncorrelated random variables with constant variance $\Rightarrow$ covariance stationary,
f) centered process of uncorrelated random variables with constant variance $\Rightarrow$ weakly stationary.

Important examples of stochastic processes

Definition (Poisson process): A Poisson process $N = \{N_t, t \geq 0\}$ is a stochastic process with the following properties:

- $N_0 = 0$ almost surely,
- N has independent increments, i.e. for every $n \in \mathbb{N}$ and $0 < t_1 < t_2 < \ldots < t_n$, the random variables $N_{t_1}, N_{t_2} - N_{t_1}, \ldots, N_{t_n} - N_{t_{n-1}}$ are independent,
- N has Poisson stationary increments, i.e. there exists a finite positive constant λ such that for every $0 \leq s < t$, we have that $N_t - N_s$ has the Poisson distribution $\mathsf{Po}(\lambda(t-s))$.

Definition (Wiener process): A Wiener process $W = \{W_t, t \geq 0\}$ is a stochastic process with the following properties:

- $W_0 = 0$ almost surely,
- W has continuous trajectories, i.e. for almost every $\omega \in \Omega$, the map $t \mapsto W_t(\omega)$ is continuous,
- W has independent increments, i.e. for every $n \in \mathbb{N}$ and $0 < t_1 < t_2 < \ldots < t_n$, the random variables $W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_n} - W_{t_{n-1}}$ are independent,
- W has stationary Gaussian increments, i.e. there exists a finite positive constant σ^2 such that for every $0 \leq s < t$, we have that $W_t - W_s$ has the normal distribution $\mathsf{N}(0, \sigma^2(t-s))$.

Exercises

Exercise 2.1: Let $X_t = a + bt + Y_t$, $t \in \mathbb{Z}$, where $a, b \in \mathbb{R}$, $b \neq 0$ and $\{Y_t, t \in \mathbb{Z}\}$ be a sequence of independent identically distributed random variables with zero mean and finite positive variance σ^2.

a) Determine the autocovariance function of the sequence $\{X_t, t \in \mathbb{Z}\}$ and discuss its stationarity.
b) For $q \in \mathbb{N}$ we define random variables V_t by the formula

$$V_t = \frac{1}{2q+1} \sum_{j=-q}^{q} X_{t+j}, \quad t \in \mathbb{Z}.$$

Determine the autocovariance function of the sequence $\{V_t, t \in \mathbb{Z}\}$ and discuss its stationarity.

Exercise 2.2: Let X be a random variable with a uniform distribution on the interval $(0, \pi)$. Consider the sequence of random variables $\{Y_t, t \in \mathbb{N}\}$ where $Y_t = \cos(tX)$. Discuss the properties of such a random sequence.

Exercise 2.3: Consider the stochastic process $X_t = \cos(t + B)$, $t \in \mathbb{R}$, where B is a random variable with a uniform distribution on the interval $(0, 2\pi)$. Check whether the process is weakly stationary.

Exercise 2.4: Let X be a random variable such that $\mathbb{E}X = 0$ and $\operatorname{var} X = \sigma^2 < \infty$. We define $X_t = (-1)^t X$, $t \in \mathbb{N}$. Discuss the properties of the sequence $\{X_t, t \in \mathbb{N}\}$.

Exercise 2.6: Let $\{N_t, t \geq 0\}$ be a Poisson process with intensity $\lambda > 0$ and let A be a real-valued random variable with zero mean and unit variance, independent of the process $\{N_t, t \geq 0\}$. We define $X_t = A(-1)^{N_t}, t \geq 0$. Determine the autocovariance function of $\{X_t, t \geq 0\}$.

Exercise 2.8: Let $\{W_t, t \geq 0\}$ be a Wiener process. We define the so-called *Ornstein-Uhlenbeck process* $\{U_t, t \geq 0\}$ by the formula $U_t = \mathrm{e}^{-\alpha t/2} W_{\exp\{\alpha t\}}, t \geq 0$, where $\alpha > 0$ is a parameter. Decide whether $\{U_t, t \geq 0\}$ is weakly (strictly) stationary and determine its autocovariance function.

Exercise 2.11: Let $\{X_t, t \in \mathbb{Z}\}$ be a sequence of independent identically distributed random variables. Prove that the process is strictly stationary. Is it also weakly stationary?

Exercise 2.12: Let $\{X_t, t \in \mathbb{Z}\}$ be a sequence of uncorrelated random variables with zero mean and finite positive variance (so-called *white noise*). Prove that it is weakly stationary. Is it also strictly stationary?

Exercise 2.13: Let $X_0 = 0$, $X_t = Y_1 + \cdots + Y_t$ for $t = 1, 2, \ldots$, where $Y_1, Y_2, \ldots$ are independent identically distributed discrete random variables with zero mean and finite positive variance. Show that $\{X_t, t \in \mathbb{N}_0\}$ is a Markov chain. Determine its autocovariance function. What can we say about the properties of such a random sequence?

Exercise 2.14: Let $\{X_t, t \in T\}$ a $\{Y_t, t \in T\}$ be uncorrelated weakly stationary processes, i.e. for all $s, t \in T$ the random variables X_s and Y_t are uncorrelated. Show that in such a case also the process $\{Z_t, t \in T\}$ with $Z_t = X_t + Y_t$ is weakly stationary.

Exercise 2.18: Determine the autocovariance function of the Wiener process $\{W_t, t \geq 0\}$. For $0 \leq t_1 < t_2 < \cdots < t_n$ determine the variance matrix of the random vector $(W_{t_1}, \ldots, W_{t_n})^\top$.