Technische Universität Chemnitz

Lecture

Random Processes

Prof. Dr. vom Scheidt/ PD Dr. Wunderlich

Fakultät für Mathematik - Lehrstuhl Stochastik

February 8, 2001

Preface

The German original text of this lecture has been written by the first author. It is based on the book "The Theory of Stochastic Processes, Part 1" by Gihman and Skorohod . The lecture has been given several times by both authors at Chemnitz University of Technology. This English translation and some modifications are due to the second author.

Please send remarks and corrections to wunderlich@mathematik.tu-chemnitz.de.

Contents

1 Preliminaries

1.1 Basic concepts from probability theory

Definition 1.1 *The tupel* $\{\Omega, \mathcal{A}, \mathbf{P}\}$ *is called* **probability space** *with*

Ω*: the set of elementary events*

$\mathcal{A}$*: a collection of subsets of* Ω *forming a* σ*-algebra*

$\mathbf{P}$*: a probability measure on* $\mathcal{A}$ $(\mathbf{P}(\Omega) = 1)$.

The sets $A \in \mathcal{A}$ *are called* **random events**, $\mathbf{P}(A)$ *its probability.*

Theorem 1.1 properties of probability measures

$A, A_i \in \mathcal{A}; i = 1, 2, \ldots$

1. $\mathbf{P}(\emptyset) = 0$
2. $\mathbf{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbf{P}(A_i)$, *if* $A_i \bigcap A_j = \emptyset \;\; \forall\; i, j$ *with* $i \neq j$
3. $\mathbf{P}(A_2 \setminus A_1) = \mathbf{P}(A_2) - \mathbf{P}(A_1)$, *if* $A_1 \subset A_2$
4. $\mathbf{P}(\bar{A}) = 1 - \mathbf{P}(A)$
5. $\mathbf{P}(\bigcup_{i=1}^{\infty} A_i) = \lim_{i\to\infty} \mathbf{P}(A_i)$, *if* $A_i \subset A_{i+1} \;\; \forall\; i$
6. $\mathbf{P}(\bigcap_{i=1}^{\infty} A_i) = \lim_{i\to\infty} \mathbf{P}(A_i)$, *if* $A_i \supset A_{i+1} \;\; \forall\; i$

Definition 1.2 *A* **random variable (r.v.)** ξ *on a probability space* $\{\Omega, \mathcal{A}, \mathbf{P}\}$ *is a* $\mathcal{A}$*-measurable real function of the elementary events* $\omega : \xi(\omega)$.

Remark 1.1

1. Let $\mathcal{A}$ be a σ-algebra of Ω, $E \in \mathcal{A}$.
   A real valued function $f : E \to \mathbb{R}$ is called measurable on E if and only if $f^{-1}(G) \in \mathcal{A}\;\; \forall$ open subsets $G \subset \mathbb{R}$.
2. $f : E \to \mathbb{R}$ is measurable on E if and only if $\forall\; a \in \mathbb{R} : \;\; f^{-1}((-\infty, a)) \in \mathcal{A}$.

Definition 1.3 *The random variables* ξ *and* η *are called* **P - equivalent**, *if*

$$\mathbf{P}(\xi \neq \eta) = \mathbf{P}(\omega : \xi(\omega) \neq \eta(\omega)) = 0.$$

Notation: $\xi = \eta$ *(mod* $\mathbf{P}$*)*, $\xi = \eta$ **almost sure (a.s.)**, $\xi = \eta$ **with probability 1**.

The concept "‘almost sure"’ is used also in more general situations:
e.g. $\xi_n = f_n(\omega)$ converges a.s. to $\xi = f(\omega)$ if:

$$\lim_{n\to\infty} \xi_n(\omega) = \xi(\omega) \;\; \forall\; \omega \in \Omega \setminus N \text{ where } \mathbf{P}(N) = 0$$

$\Rightarrow$ notation: $\lim_{n\to\infty} \xi_n = \xi \mod \mathbf{P}$ or $\lim_{n\to\infty} \xi_n = \xi$ a.s.

Theorem 1.2 Application of properties of measurable functions

1. *Let* $h(t_1, \ldots, t_n)$ *be arbitrary Borel - functions of* n *variables* $t_1, \ldots, t_n$ *, and let* $\xi_1, \ldots, \xi_n$ *be r.v.* $\Rightarrow$ $h(\xi_1, \ldots, \xi_n)$ *is also a random variable.*

2. *Let $(\xi_n)_{n\geq 1}$ be a sequence of r.v.*

$$\Rightarrow \quad \sup_n \xi_n, \quad \inf_n \xi_n, \quad \overline{\lim_{n\to\infty}}\ \xi_n, \quad \underline{\lim_{n\to\infty}}\ \xi_n \text{ are also r.v.}$$

3. *Let $\xi_n = \eta_n$ a.s., $n = 1, 2, \dots$; and $h(t_1, \dots, t_n)$ be a Borel - function*

$$\Rightarrow \quad h(\xi_1, \dots, \xi_n) \ = \ h(\eta_1, \dots, \eta_n) \ a.s.$$

$$\Rightarrow \quad \sup_n \ \xi_n \ = \ \sup_n \ \eta_n \ a.s., \qquad \inf_n \ \xi_n \ = \ \inf_n \ \eta_n \ a.s.$$

$$\Rightarrow \quad \overline{\lim_n}\ \xi_n \ = \ \overline{\lim_n}\ \eta_n \ a.s., \qquad \underline{\lim_n}\ \xi_n \ = \ \underline{\lim_n}\ \eta_n \ a.s.$$

4. *Let $(\xi_n)_{n\geq 1}$ be a sequence of r.v.*

$$\Rightarrow \quad A = \{\omega \in \Omega : \ \lim_n \xi_n(\omega) \ \ \exists\} \text{ is } \mathcal{A}\text{-measurable}$$

$$\left(\text{see } A = \bigcap_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{m_1, m_2 > n} \{\omega : \ |\xi_{m_1} - \xi_{m_2}| \ < \ \tfrac{1}{k}\}\right)$$

Remark 1.2 $h(t_1, \dots, t_n)$ is called Borel - function, if it is measurable with respect to the σ-algebra of $\mathbb{R}^n$, which is generated by the open subsets of $\mathbb{R}^n$.

Definition 1.4 *A r.v. is called* **discrete**, *if there are at most countable many values:*

$$\xi = \sum_k c_k I_{A_k}(\omega) \quad A_k \in \mathcal{A}, \quad \bigcup_k A_k = \Omega, \quad A_i \cap A_j = \emptyset \ \ \forall \ i \neq j$$

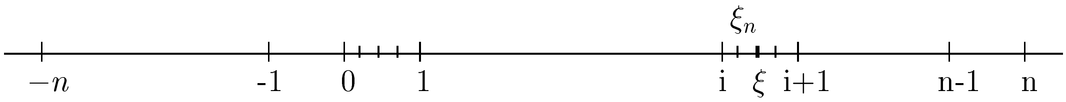
Theorem 1.3 *Let ξ be an arbitrary r.v., then there exists a sequence of discrete r.v. $(\xi_n)_{n\geq 1}$ with finite number of values, such that*

$$\lim_{n\to\infty} \xi_n(\omega) \ = \ \xi(\omega) \quad \forall \ \ \omega \in \Omega.$$

Proof:

$$\xi_n = \sum_{i=-n}^{n-1} \sum_{k=1}^{n} (i + \tfrac{k-1}{n}) I_{A_{ik}} \quad \text{with} \quad A_{ik} = \{\omega : \ i + \tfrac{k-1}{n} \leq \xi < i + \tfrac{k}{n}\}$$

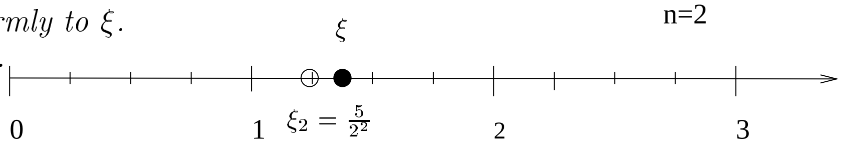
$$\Rightarrow \quad \text{for} \quad |\xi(\omega)| < n \text{ let } \omega \in A_{ik} \quad \Rightarrow \quad |\xi - \xi_n| = |\xi - (i + \tfrac{k-1}{n})| < \tfrac{1}{n}$$



$\Rightarrow$ assertion

Theorem 1.4 *Let ξ be a nonneagtive r.v.. Then there exists a monotone increasing sequence of discrete r.v.'s (with countable many values), which converges uniformly to ξ.*

Proof:



$$\xi_n := \sum_{i=0}^{\infty} \tfrac{i}{2^n} I_{A_{in}} \quad \text{with} \quad A_{in} := \{\omega : \ \tfrac{i}{2^n} \leq \xi < \tfrac{i+1}{2^n}\} \qquad \Rightarrow \quad |\xi - \xi_n| < \tfrac{1}{2^n} \ \ \forall \ \omega$$

Definition 1.5 *A* **random element(r.e.)** *ξ with values in a measurable space $(\mathcal{R}, \mathcal{B})$ is a measurable mapping of $\{\Omega, \mathcal{A}, \mathbf{P}\}$ into $\{\mathcal{R}, \mathcal{B}\}$, i.e.*

$$\xi^{-1}(B) = \{\omega : \ \xi(\omega) \in B\} \ \in \ \mathcal{A} \ \ \forall \ B \imath \mathcal{B}$$

- If $\mathcal{R}$ is a metric space, then $\mathcal{B}$ denotes the σ-algebra of Borel sets (if not denoted otherwise).
- If $\mathcal{R}$ is a vector space, then ξ is called **random vector** .

Let $\{\xi_k,\ k = 1,\ldots,n\}$ be r.e.'s on $\{\Omega, \mathcal{A}, \mathbf{P}\}$, with values of ξ_k in $\{\mathcal{R}_k, \mathcal{B}_k\}$ then $\{\xi_k,\ k = 1,\ldots,n\}$ can be considered as a r.e. η with values in a measurable space $\{\mathcal{S}, \mathcal{C}\}$

with $\mathcal{S} = \prod_{k=1}^{n} \mathcal{R}_k$ (product of spaces) $\mathcal{R}_1,\ldots,\mathcal{R}_n$)

and $\mathcal{C} = \prod_{k=1}^{n} \mathcal{B}_k$ (product of σ-algebras $\mathcal{B}_1,\ldots,\mathcal{B}_n$).

The r.e. η is called **direct product** of the r.e.'s $\xi_1,\ldots,\xi_n$.

Remark 1.3

1. $\prod_{k=1}^{n} \mathcal{R}_k := \{y :\ y = (x_1,\ldots,x_n),\ \ x_i \in \mathcal{R}_i\ \ \forall\ i\}$
2. projection: $\pi_i :\ \prod_{k=1}^{n} \mathcal{R}_k \to \mathcal{R}_i :\quad \pi_i(y) = x_i$
3. $\prod_{k=1}^{n} \mathcal{B}_k := \sigma(\prod_{k=1}^{n} B_k,\ B_k \in \mathcal{B}_k)$

   σ-algebra, which is generated by the sets $\prod_{k=1}^{n} B_k$ with $B_k \in \mathcal{B}_k$.

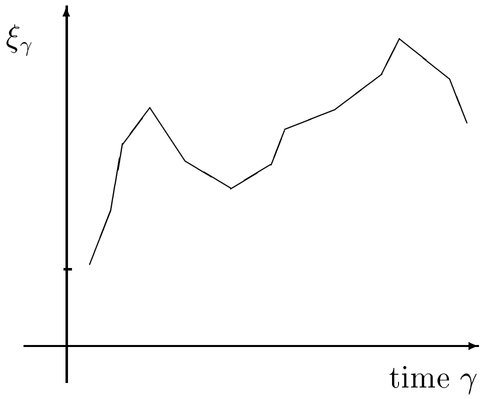
Generalization: $\{\xi_\gamma, \gamma \in \Gamma\}$ family of r.e.'s with ξ_γ in $\{\mathcal{R}_\gamma, \mathcal{B}_\gamma\}$

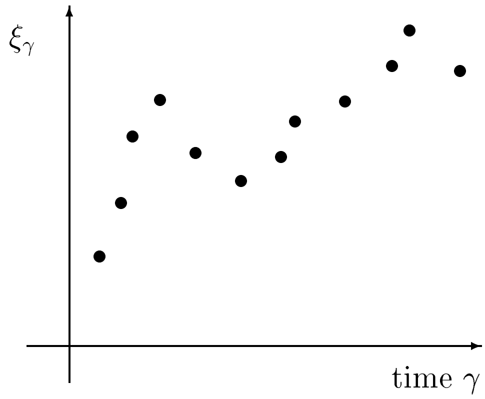
- $\mathcal{S} = \prod_{\gamma\in\Gamma} \mathcal{R}_\gamma = \{y = y(\gamma) :\ \gamma \to \xi_\gamma,\ \xi_\gamma \in \mathcal{R}_\gamma,\ \gamma \in \Gamma\}$
  is the space of functions defined on Γ with values in $\mathcal{R}_\gamma$ for $\gamma \in \Gamma$;
- cylinder sets in $\mathcal{S}$: $n \in \mathrm{N}$ arbitrary;
  $C = \{y \in \mathcal{S} :\ y(\gamma_k) \in B_{\gamma_k};\ \ B_{\gamma_k} \in \mathcal{B}_{\gamma_k};\ k = 1,2,\ldots,n\}$,
  $C = C_{\gamma_1\ldots\gamma_n}(B_{\gamma_1} \times \cdots \times B_{\gamma_n})$
  is cylinder set with basis $B_{\gamma_1} \times \cdots \times B_{\gamma_n}$ and coordinates $\gamma_1,\ldots,\gamma_n$;
- $\mathcal{B} = \prod_{\gamma\in\Gamma} \mathcal{B}_\gamma$:= minimal σ-algebra generated by the cylinder sets;
  is called product of σ-algebras $\mathcal{B}_\gamma,\ \gamma \in \Gamma$

$\Rightarrow$ $g :\ \omega \to g(\omega)$ is defined by $g(\omega) = g(\omega,\gamma) = \xi_\gamma(\omega)$;
is a measurable mapping $\{\Omega, \mathcal{A}\} \to \{\mathcal{S}, \mathcal{B}\}$

Special case: $\mathcal{R}_\gamma = \mathcal{R}\ \ \forall\ \gamma \in \Gamma \quad \Rightarrow \quad$ notation: $\mathcal{S} = \mathcal{R}^\Gamma$
$\mathcal{S}$: space of all functions with values in $\mathcal{R}$ defined on Γ
$g(\omega) :\ \ \Omega \to \mathcal{R}^\Gamma$, i.e. every ω is mapped into a function of $\mathcal{R}$
$\Rightarrow$ **random functions**
i.e. the family of r.v.'s $\{\xi_\gamma,\ \gamma \in \Gamma\}$ is a random function

Example 1.1 • stock price



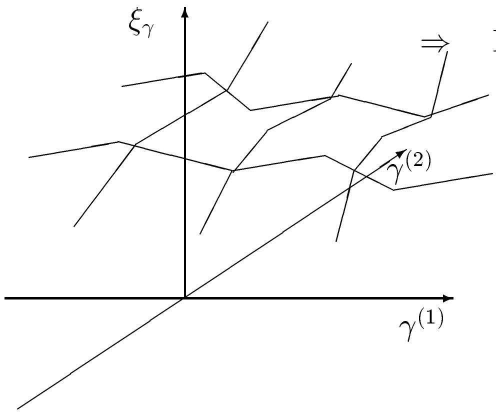


$\Gamma = [0, \infty), \quad \mathcal{R} = (0, \infty)$

daily values $\Rightarrow$ $\Gamma \subset \mathbb{N} \cup \{0\}$
(time series, random sequence)

- velocity of wind v : $\Rightarrow$ $v = (v_x, v_y, v_z)^T$
  - $v = v(x, y, z, t)$ $\Rightarrow$ $\Gamma \subset \mathbb{R}^3 \times [0, \infty),\ \mathcal{R} \subset \mathbb{R}^3$
  - $v = v(t)$, (fixed point (x, x, z)) $\Rightarrow$ $\Gamma \subset [0, \infty)$;

- random surface $\Rightarrow$ $\Gamma \subset \mathbb{R}^2,\ \mathcal{R} \subset \mathbb{R}$



Let $\xi(\omega)$ be a r.e. on $\{\Omega, \mathcal{A}, \mathbf{P}\}$ with values in $\{\mathcal{S}, \mathcal{B}\}$

Definition 1.6 *The σ-algebra consisiting of the sets $\{\xi^{-1}(B),\ B \in \mathcal{B}\}$ is called the σ-algebra which is generated by the r.e. ξ and denoted by $\mathcal{A}_\xi$ $(\subset \mathcal{A})$.*

Remark 1.4 $\mathcal{C} = \{\xi^{-1}(B),\ B \in \mathcal{B}\}$ is a σ-Algebra, since

1. $A \in \mathcal{C} \quad \Rightarrow \quad A = \xi^{-1}(B) \quad \Rightarrow \quad \overline{A} = \xi^{-1}(\overline{B}) \quad \Rightarrow \quad \overline{A} \in \mathcal{C}$
2. $A_i \in \mathcal{C} \quad \Rightarrow \quad A_i = \xi^{-1}(B_i) \quad \Rightarrow \quad \bigcup\limits_i A_i = \xi^{-1}(\bigcup\limits_i B_i) \quad \Rightarrow \quad \bigcup\limits_i A_i \in \mathcal{C}$

Theorem 1.5 *Let ξ be a r.e. on $\{\Omega, \mathcal{A}, \mathbf{P}\}$ with values in $(\mathcal{S}, \mathcal{B})$ and let η be a $\mathcal{A}_\xi$ - measurable r.v.. Then there exists a measurable real function $g(x)$ defined on $\mathcal{S}$ such that $\eta = g(\xi)$.*

Proof: see basic course Theorem 3.5

$$\begin{array}{ccc} (\Omega,\mathcal{A}) & \stackrel{\xi(\omega)}{\longrightarrow} & (\mathcal{S},\mathcal{B}) \\ \bigcup & \stackrel{\xi^{-1}}{\swarrow} & \downarrow g \\ (\Omega,\mathcal{A}_\xi) & \stackrel{\eta}{\longrightarrow} & (\mathbb{R},\mathcal{C}) \end{array}$$

1.Let η be a discrete r.v. with values $a_n,\ n=1,2,\ldots$

$A_n=\{\omega:\ \eta=a_n\} \quad\Rightarrow\quad \eta^{-1}(a_n)=A_n\in\{\xi^{-1}(B),\ B\in\mathcal{B}\}$

i.e. $\exists\ B_n\in\mathcal{B}$ with $\xi^{-1}(B_n)=A_n$

$B'_n:=B_n\setminus\bigcup\limits_{i=1}^{n-1}B_i \quad\Rightarrow\quad B'_n\cap B'_m=\emptyset$ for $n\neq m$

$$\xi^{-1}(B'_n)=A_n\setminus\bigcup_{i=1}^{n-1}A_i=A_n$$

$$\xi^{-1}(\bigcup_{i=1}^{\infty}B'_n)=\bigcup_{i=1}^{\infty}A_n=\Omega \quad\Rightarrow\quad \xi(\Omega)\subset\bigcup_{i=1}^{\infty}B'_n$$

$\Rightarrow\quad g(x):=a_n$ for $x\in B'_n \quad\Rightarrow\quad \eta=g(\xi)$

2. general case: let η be an arbitrary r.v.

Theorem 1.3 $\Rightarrow\exists$ sequence of discrete σ_ξ-measurable r.v. η_n

with $\lim\limits_{n\to\infty}\eta_n=\eta\quad\forall\ \omega\in\Omega$

due to 1. $\exists\ g_n$ with $\eta_n=g_n(\xi)$ and $g_n(x)$ $\mathcal{B}$ - measurable

Theorem 1.2 4. $\Rightarrow\quad S:=\{x:\ \lim\limits_{n\to\infty}g_n(x)\ \exists\}\ \in\ \mathcal{B}$ (considered more general)

it holds $\lim\limits_{n\to\infty}\eta_n(\omega)=\lim\limits_{n\to\infty}g_n(\xi(\omega))=\eta(\omega)\quad\Rightarrow\quad\xi(\Omega)\subset S$

(see $\lim\limits_{n\to\infty}\eta_n=\eta\quad\forall\ \omega\in\Omega$)

$\Rightarrow$ for $x\in\xi(\Omega):\quad \lim\limits_{n\to\infty}g_n(x)=\lim\limits_{n\to\infty}g_n(\xi(\omega))=\lim\limits_{n\to\infty}\eta_n(\omega)=\eta(\omega)$

hence

$$g(x):=\begin{cases}\lim\limits_{n\to\infty}g_n(x) & \text{for } x\in S\\ 0 & \text{for } x\notin S\end{cases}\quad\Rightarrow\quad\eta=g(\xi)$$

(see $g(\xi(\omega))=\lim\limits_{n\to\infty}g_n(\xi(\omega))=\eta(\omega)$) $\Rightarrow$ assertion

Definition 1.7 *The* **expectation** *(or mean value) of the r.v. $\xi(\omega)$ is the integral of $\xi(\omega)$ with respect to the probability measure* $\mathbf{P}$:

$$\mathbf{E}\,\xi\ =\ \int_\Omega\xi(\omega)\mathbf{P}(d\omega)\ =\ \int_\Omega\xi d\mathbf{P}.$$

especially it holds for

abs. continuous ξ :	$\mathbf{E}\,\xi=\int\limits_{-\infty}^{\infty}xf_\xi(x)dx$	*(Lebesgue-integral, f_ξ-probability density)*
discrete ξ :	$\mathbf{E}\,\xi=\sum\limits_k x_k\mathbf{P}(\xi=x_k)$	*(see measure theory, section 2)*

We have introduced convergence with probability 1 (a.s. convergence) for sequences of r.v.'s (see Def. 1.3).

Definition 1.8 *The sequence of r.v.'s* $(\xi_n)_{n\geq1}$ *converges* **in probability** *to the r.v. ξ, if*

$$\lim_{n\to\infty}\mathbf{P}(|\xi_n-\xi|>\varepsilon)=0\qquad \textit{for every } \varepsilon>0.$$

notation: $\xi = \mathbf{P}\text{–}\lim\limits_{n\to\infty} \xi_n$

interpretation: For every $\varepsilon > 0$ the distribution of the "probability mass" of $\xi_n - \xi$ for increasing n concentrates in a ε-ball around zero.

Theorem 1.6 (relations between several kinds of convergence)
(see basic course, section 5.2)

1. *The convergence with probability 1 implies the convergence in probability. (In general, the converse statement is not true, but it holds:)*
   If a sequence is convergent in probability then it contains a subsequence that is convergent with probability 1.

2. *A sequence $(\xi_n)_{n\geq 1}$ of r.v. is convergent in probability if and only if* $\forall\ \varepsilon > 0,\ \delta > 0\ \exists\ n_0 = n_0(\varepsilon, \delta)$, *such that* $\forall\ n, n' > n_0$ *it holds:*
   $$\mathbf{P}(|\xi_{n'} - \xi_n| > \varepsilon) < \delta. \qquad \textit{(fundamental or Cauchy condition)}$$

3. *If* $\xi = \mathbf{P}\text{–}\lim \xi_n,\quad \eta = \mathbf{P}\text{–}\lim \xi_n$ *then* $\xi = \eta \bmod \mathbf{P}$. $\qquad (\xi = \eta$ *a.s.*$)$

4. *Let* $\eta_k = \mathbf{P}\text{–}\lim\limits_{n\to\infty} \xi_{kn}$ *for* $k = 1, 2, \ldots, m$ *and* $\varphi(t_1, \ldots, t_m)$ *a continuous function on* R^m *possibly with the exception of points of a Borel set* D $(D \subset R^m)$ *with* $\mathbf{P}((\eta_1, \ldots, \eta_m) \in D) = 0$.
   Then it holds $\quad \varphi(\eta_1, \ldots, \eta_m) = \mathbf{P}\text{–}\lim\limits_{n\to\infty} \varphi(\xi_{1n}, \ldots, \xi_{mn})$
   special cases:
   $$\begin{aligned} \mathbf{P}\text{–}\lim\,(\xi_{1n} + \xi_{2n}) &= \mathbf{P}\text{–}\lim \xi_{1n} + \mathbf{P}\text{–}\lim \xi_{2n} \\ \mathbf{P}\text{–}\lim\,(\xi_{1n} \cdot \xi_{2n}) &= \mathbf{P}\text{–}\lim \xi_{1n} \cdot \mathbf{P}\text{–}\lim \xi_{2n} \\ \mathbf{P}\text{–}\lim \frac{\xi_{1n}}{\xi_{2n}} &= \frac{\mathbf{P} - \lim \xi_{1n}}{\mathbf{P} - \lim \xi_{2n}}, \quad \textit{if} \quad \mathbf{P}\text{–}\lim \xi_{2n} \neq 0 \textit{ a.s.}. \end{aligned}$$

Theorem 1.7 (criterion for a.s. convergence)

1. *The sequence of r.v.'s $(\xi_n)_{n\geq 1}$ is convergent with probability 1 if there exist a sequence $(\varepsilon_n)_{n\geq 1}$ with $\varepsilon_n > 0$, $\sum\limits_{n=1}^{\infty} \varepsilon_n < \infty$ such that*
   $$\sum_{n=1}^{\infty} \mathbf{P}(|\xi_{n+1} - \xi_n| > \varepsilon_n) < \infty.$$

2. *The sequence of r.v.'s $(\xi_n)_{n\geq 1}$ converges with probability 1 to a r.v. ξ if there exists an $\varepsilon > 0$ such that*
   $$\sum_{n=1}^{\infty} \mathbf{P}(|\xi - \xi_n| > \varepsilon) < \infty.$$

Proof: see basic course (Corollary 5.2 - 5.4)

Definition 1.9 *Let for $p \geq 1$*

$$\mathcal{L}_p = \mathcal{L}_p(\Omega, \mathcal{A}, \mathbf{P}) = \{\xi \text{ r.v. on } \{\Omega, \mathcal{A}, \mathbf{P}\} : \mathbf{E}\,|\xi|^p < \infty\}$$

A sequence $(\xi_n)_{n\geq 1}$ is called $\mathcal{L}_p$**-convergent** *(or* **convergent in mean of order p***), if*

$$\lim_{n\to\infty} \mathbf{E}\,|\xi - \xi_n|^p = 0.$$

Remark 1.5 *For $p \geq 1$ $\mathcal{L}_p$ is a complete space (i.e. Cauchy sequences are convergent) and it can be equipped with the norm : $\|\xi\| = (\mathbf{E}\,|\xi|^p)^{\frac{1}{p}}$. $\Rightarrow$ $\mathcal{L}_p$ is a complete normed space (Banach space).*

Theorem 1.8 *The $\mathcal{L}_p$-convergence implies the convergene in probability.*

Proof: Markov (or generalized Chebychev) inequality implies

$$\mathbf{P}(|\xi_n - \xi| > \varepsilon) \leq \frac{\mathbf{E}\,|\xi - \xi_n|^p}{\varepsilon^p} \quad \text{for} \quad p \geq 1.$$

Consider the space $\mathcal{L}_2 = \mathcal{L}_2(\Omega, \mathcal{A}, \mathbf{P})$ of complex valued r.v.

(above results also hold complex valued r.v.)

$\mathcal{L}_2$ is a Hilbert space with scalar product $[\xi, \eta] = \mathbf{E}\,\xi\bar{\eta}$; $\|\xi\|^2 = [\xi, \xi]$

$\Rightarrow$ ξ und η are called **orthogonal** if $\mathbf{E}\,\xi\bar{\eta} = 0$

for $\mathbf{E}\,\xi = \mathbf{E}\,\eta = 0$: ξ, η orthogonal $\Rightarrow$ ξ, η uncorrelated

$\Rightarrow$$(\xi_n)_{n\geq 1}$ is $\mathcal{L}_2$-convergent to ξ if $\lim_{n\to\infty} \|\xi_n - \xi\|^2 = \lim_{n\to\infty} \mathbf{E}\,|\xi_n - \xi|^2 = 0$

notation: **convergence in mean square** $\underset{n\to\infty}{\text{l.i.m.}}\ \xi_n = \xi$ (limit in mean)

Definition 1.10 *Let $\{\xi_t,\ t \in T\}$ be a family of r.v. with $\xi_t \in \mathcal{L}_2$ and T an arbitrary set. Then*

$$B(t_1, t_2) := \mathbf{E}\,\xi_{t_1}\bar{\xi}_{t_2} \qquad t_1, t_2 \in T$$

is called **covariance***.*

Definition 1.11 *Let $\psi(t)$ be a nonnegative real function with $D(\psi) = T$ which may have arbitrarily small values. For a family $\{\xi_t,\ t \in T\}$ with $\xi_t \in \mathcal{L}_2$the r.v. $\eta \in \mathcal{L}_2$ is called* $\mathcal{L}_2$**-limit** *or* **mean-square limit** *for $\psi(t) \to 0$ if for all $\varepsilon > 0$ there exists a $\delta > 0$ such that*

$$\mathbf{E}\,|\xi_t - \eta|^2 \leq \varepsilon \quad \forall\, t \quad \text{with} \quad 0 \leq \psi(t) < \delta.$$

Example 1.2 *$T = \mathbb{R}^1$, $\psi(t) = |t - t_0|$;*
η is the mean-square limit of ξ_t for $t \to t_0$

Theorem 1.9 *There exists the mean-square limit η of the family $\{\xi_t, t \in T\}$ for $\psi(t) \to 0$ if and only if there exists the limit of the covariance $B(t_1, t_2) = \mathbf{E}\,\xi_{t_1}\bar{\xi}_{t_2}$ for $\psi(t_1) + \psi(t_2) \to 0$. In this case it holds $\mathbf{E}\,|\eta|^2 = \lim_{\psi(t)\to 0} B(t, t)$.*

Proof:

necessity: to prove: $\exists \operatorname*{l.i.m.}_{\psi(t)\to 0} \xi_t = \eta \quad \Rightarrow \quad \exists \lim_{\psi(t_1)+\psi(t_2)\to 0} B(t,t_2) \quad (= [\eta,\eta])$

$$\begin{array}{ccccl}
\|\eta-\xi_{t_1}\| & \cdot & \|\eta-\xi_{t_2}\| & \geq & |[\eta-\xi_{t_1},\eta-\xi_{t_2}]| \qquad \text{(Cauchy-inequality)} \\
 & & & = & |[\eta,\eta]-[\eta,\xi_{t_2}]-[\xi_{t_1},\eta]+[\xi_{t_1},\xi_{t_2}]| \\
 & & & = & |-[\eta,\eta]-[\eta,\xi_{t_2}-\eta]-[\xi_{t_1}-\eta,\eta]+[\xi_{t_1},\xi_{t_2}]| \\
\psi(t_1)\to 0 & & \psi(t_2)\to 0 & & \qquad\quad \psi(t_2)\to 0 \qquad \psi(t_1)\to 0 \\
\downarrow & & \downarrow & & \qquad\qquad \downarrow \qquad\qquad\quad \downarrow \\
0 & & 0 & & \qquad\qquad 0 \qquad\qquad\quad 0
\end{array}$$

$$\Rightarrow \quad [\xi_{t_1},\xi_{t_2}] \;=\; \mathbf{E}\,\xi_{t_1}\overline{\xi}_{t_2} \;=\; B(t_1,t_2) \overset{\psi(t_1)+\psi(t_2)\to 0}{\longrightarrow} [\eta,\eta]$$

(follows from continuity of scalar product)

sufficiency: it $\exists \lim_{\psi(t_1)+\psi(t_2)\to 0} B(t_1,t_2) = B_0 \geq 0$ since $B_0 = \lim_{\psi(t)\to 0} B(t,t) \geq 0$

$$\begin{aligned}
\mathbf{E}\,|\xi_{t_1}-\xi_{t_2}|^2 &= \mathbf{E}\,(\xi_{t_1}-\xi_{t_2})(\overline{\xi_{t_1}-\xi_{t_2}}) \\
&= \mathbf{E}\,\xi_{t_1}\overline{\xi_{t_1}} - \mathbf{E}\,\xi_{t_1}\overline{\xi_{t_2}} - \mathbf{E}\,\xi_{t_2}\overline{\xi_{t_1}} + \mathbf{E}\,\xi_{t_2}\overline{\xi_{t_2}} \\
&= B(t_1,t_1) - 2\mathrm{Re}\,B(t_1,t_2) + B(t_2,t_2) \overset{\psi(t_1)+\psi(t_2)\to 0}{\longrightarrow} 0
\end{aligned}$$

$$\overset{\mathcal{L}_2\text{complete}}{\longrightarrow} \quad \exists\, \eta \in \mathcal{L}_2 \quad \text{with} \quad \eta = \operatorname*{l.i.m.}_{\psi(t)\to 0} \xi_t$$

$$\begin{aligned}
\Rightarrow \quad \big|\|\eta\|^2 - \|\xi_t\|^2\big| &= \big|\|\eta\| - \|\xi_t\|\big|(\|\eta\|+\|\xi_t\|) \\
&\leq \|\eta-\xi_t\|\;\|\eta\| + \|\eta-\xi_t\|\|\xi_t\| \overset{\psi(t)\to 0}{\longrightarrow} 0
\end{aligned}$$

(see $|\|f\|-\|g\|| \leq \|f-g\|$)

$$\Rightarrow \quad \|\eta\|^2 \;=\; \mathbf{E}\,|\eta|^2 \;=\; \lim_{\psi(t)\to 0} B(t,t) \quad (\text{see } \|\xi_t\| = [\xi_t,\xi_t] = \mathbf{E}\,\xi_t\bar{\xi}_t = B(t,t))$$

$\Rightarrow$ assertion

extension: $\mathcal{L}_2^m = \mathcal{L}_2^m(\Omega,\mathcal{A},\mathbf{P})$

denotes the Hilbert space of random vectors with values in the complex space $\mathbb{C}^m$

$\xi \in \mathcal{L}_2^m$ if $\quad \xi = (\xi^1(\omega),\ldots,\xi^m(\omega)) \in \mathrm{C}^m$ and $\mathbf{E}\sum_{i=1}^m |\xi^i|^2 < \infty$

$\xi,\eta \in \mathcal{L}_2^m \quad \Rightarrow$ scalar product $[\xi,\eta] = \mathbf{E}\,(\xi,\eta)$

and norm $\|\xi\|^2 = [\xi,\xi] = \mathbf{E}\sum_{i=1}^m |\xi^i|^2$

(ξ_t^i - i-th component of ξ_t)

where $(\xi,\eta) = \sum_{i=1}^m \xi^i\overline{\eta^i}$, is the scalar product of C^m,

$$B(t_1,t_2) := [\xi_{t_1},\xi_{t_2}] = \mathbf{E}\sum_{i=1}^m \xi_{t_1}^i\,\overline{\xi_{t_2}^i}$$

Theorem 1.9 can be extended to spaces $\mathcal{L}_2^m$

Definition 1.12 *Let ξ be a r.e. with values in a measurable space $\{\mathcal{S},\mathcal{B}\}$. The measure μ on $\mathcal{B}$ with*

$$\mu(B) \;=\; \mathbf{P}(\omega:\; \xi \in B), \quad B \in \mathcal{B},$$

is called the **distribution** *of the r.e.* ξ*.*

Theorem 1.10 *For an arbitrary* $\mathcal{B}$*-measurable function* $f(x)$ *it holds*

$$\mathbf{E}\, f(\xi) = \int\limits_{\Omega} f(\xi(\omega))\mathbf{P}(d\omega)) = \int\limits_{\mathcal{S}} f(x)\mu(dx)$$

if the above integrals exist.

Proof: basic course Theorem 3.25.

Remark 1.6 *The above theorem states that arbitrary stochastic characteristics* $\mathbf{E}\, f(\xi)$ *can be computed as integrals over* $\mathcal{S}$ *using the distribution of* ξ*.*

<u>restriction to distributions in R^m</u> (can be generalized to general metric spaces)

$\mathcal{B}^m$: σ-algeba of Borel sets of R^m

$$a = (a^1, a^2, \ldots, a^m) \in R^m, \qquad b = (b^1, b^2, \ldots, b^m) \in R^m$$
$$a < b \ \ (\text{resp } a \le b) \quad \longleftrightarrow \quad a^i < b^i \ \ (\text{resp. } a^i \le b^i), \quad i = 1, 2, \ldots, m$$
$$I_a = \{x : \ x \le a\}, \qquad I(a, b] = \{x : \ a < x \le b\}$$

Definition 1.13 *Let* ξ *be a random vector in* $\{R^m, \mathcal{B}^m\}$*. The function*

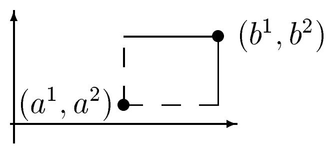
$$F(x) \ = \ \mu(I_x) \ = \ \mathbf{P}(\xi \le x) \quad \textit{with} \quad x \in R^m$$

is called the **distribution function** *of* ξ*.*

computation of $\mathbf{P}(\xi \in I(a, b]) = \mu(I(a, b])$ using the distribution function:

let G be an arbitrary function on R^m, $\alpha, \beta \in \mathbb{R}$

$$\Delta^{(k)}_{(\alpha,\beta]} G(x) := G(x^1, \ldots, x^{k-1}, \beta, x^{k+1}, \ldots, x^m)$$
$$-G(x^1, \ldots, x^{k-1}, \alpha, x^{k+1}, \ldots, x^m)$$



$$\Rightarrow \quad \Delta^{(k)}_{(\alpha,\beta]} F(x) =$$
$$\mathbf{P}(\xi^1 \le x^1, \ldots, \xi^{k-1} \le x^{k-1}, \alpha < \xi^k \le \beta, \xi^{k+1} \le x^{k+1}, \ldots, \xi^m \le x^m)$$

$$\Rightarrow \quad \mu(I(a, b]) = \Delta^1_{(a^1,b^1]}\ \Delta^2_{(a^2,b^2]}\ \cdots\ \Delta^m_{(a^m,b^m]}\ F(x) \qquad (*)$$

Using the subsequent properties of distribution functions it is also possible to consider open intervals $I(a, b)$ and closed intervals $I[a, b]$.

Theorem 1.11 *(necessary and sufficient conditions for distribution functions)*
1. A distribution function $F(x) = \mu(I_x) = \mathbf{P}(\xi \le x)$ *posseses the following properties*

1. $0 \le F(x) \le 1$
2. $F(x) \le F(y)$ *for* $x \le y$
3. $\mu\big((a, b]\big) \ = \ \mu(I(a, b])$ *is given by* $(*)$
4. $F(x + 0) = F(x)$*, i.e.* $F(x)$ *continuous on the right*

5. $F(x) \to 0$, *if* $x^k \to -\infty$ *for at least one* k

6. $F(+\infty, +\infty, \dots, +\infty) = 1$

2. For an arbitrary functiuon $F(x)$ *on* $\mathbb{R}^m$ *satisfying conditions 1.-6. there exists a unique measure* μ *on* $\mathcal{B}^m$ *with a distribution function which coincides with* $F(x)$.

Proof: see basic course Theorem 2.6 and 2.7

Definition 1.14 *A sequence of finite measures* $(\mu_n)_{n\geq 1}$ *on* $\mathcal{B}^m$ *is called* **weakly convergent** *to a measure* μ *on* $\mathcal{B}^m$ *if for all bounded and continuous functions* $f(x)$ *it holds*

$$\int_{R^m} f(x)\mu_n(dx) \quad \overset{n\to\infty}{\longrightarrow} \quad \int_{R^m} f(x)\mu(dx).$$

Definition 1.15 *A family of measures is called* **weakly compact**, *if every sequence of measures from that family contains a weakly convergent subsequence.*

Theorem 1.12 *A family of measures* $(\mu_n)_{n\geq 1}$ *on* $\{R^m, \mathcal{B}^m\}$ *is weakly compact if and only if*

1. $\mu_n(R^m) \ \leq \ C$
2. $\forall \ \ \varepsilon > 0 \ \ \exists$ *an interval* $I(a,b]$ *with* $\sup\limits_n \mu_n(R^m \setminus I(a,b]) < \varepsilon$.

Characteristic functions

Definition 1.16 *Let* ξ *be a random vector with values in* R^m *and* μ *be its distribution. The function*

$$J(u) \ = \ \mathbf{E}\, e^{i(u,\xi)} \ = \int_{R^m} e^{i(u,x)}\mu(dx), \qquad u = (u_1, \dots, u_m)$$

is called **characteristic function** *of* ξ *(or of* μ*).*

(see scalar product in $R^m : (u,x) = \sum\limits_{j=1}^m u^j x^j$)

Theorem 1.13 *Characteristic functions possess the following properties:*

1. $J(0) = 1, \quad |J(u)| \leq 1,$
2. $J(u)$ *is uniformly continous* $\forall \ \ u \in R^m$,
3. $J(u)$ *is positive semi-definite, i.e.* $\forall\, n \in N, \ \ \forall\, z_k \in C, \ \ \forall\, u_k \in R^m$ *it holds* $\sum\limits_{k,l=1}^n J(u_k - u_l) z_k \bar{z}_l \geq 0.$

Every function satisfying properties 1.-3. is a characteristic function of a certain distribution.

Proof: see basic course Theorem 6.2, 6.6

Question: Is the distribution uniquely determined by its characteristic function ?

Theorem 1.14 *Let μ_1, μ_2 be measures on $\{R^m, \mathcal{B}^m\}$. Then*

$$\int\limits_{R^m} e^{i(u,x)}\mu_1(dx) = \int\limits_{R^m} e^{i(u,x)}\mu_2(dx) \quad \forall\ u \in R^m \qquad \textit{implies} \quad \mu_1 = \mu_2.$$

Proof: basic course, Theorem 6.3.

Relation between weak convergence of distributions and convergence of characteristic functions:

Theorem 1.15 *Let $(\mu_n)_{n\geq 1}$ be sequence of distributions on $\mathbb{R}^m$, $m \in \mathbb{N}$ and let $(J_n)_{n\geq 1}$ be the corresponding sequence of characteristic functions.*

If $\lim\limits_{n\to\infty} J_n(u)$ exists $\forall\ u \in R^m$ and $K(u) = \lim\limits_{n\to\infty} J_n(u)$ is continuous at $u = 0$ then $K(u)$ is the characterstic function of a distribution μ on $\mathbb{R}^m$ and $(\mu_n)_{n\geq 1}$ is weakly convergent to μ.

Proof: $m = 1$: see basic course Theorem 8.4, $m > 1$ using Theorem 1.12

Theorem 1.16 Further properties of characteristic functions

1. *Let ξ_1, ξ_2 be independent random vectors with values in $\mathbb{R}^m$, $\xi_3 := \xi_1 + \xi_2$ and let $J_i(u)$ denote the characteristic functions of ξ_i, $i = 1, 2, 3$*
   $\Rightarrow \quad J_3(u) = J_1(u) \cdot J_2(u)$

2. *Let ξ_1, ξ_2 be random v. with values in $\mathbb{R}^{m_1}, \mathbb{R}^{m_2}$, $\xi_3 := (\xi_1, \xi_2) \in \mathbb{R}^{m_1+m_2}$.*
   ξ_1, ξ_2 are independent if and only if
   $$J_3(u, v) = J_1(u) \cdot J_2(v) \qquad (u \in R^{m_1},\ v \in R^{m_2})$$
   where $J_3(u, v) = \mathbf{E}\, e^{i((u,\xi_1)+(v,\xi_2))}$, $\quad J_1(u) = J_3(u, 0)$; $\quad J_2(v) = J_3(0, v)$.

Proof:

1. ξ_1, ξ_2 independent $\quad\Rightarrow\quad \mathbf{E}\, e^{i(u,\xi_1+\xi_2)} = \mathbf{E}\, e^{i(u,\xi_1)}\mathbf{E}\, e^{i(u,\xi_2)}$
2. necessitiy: follows from 1.
   sufficiency: follows form the unique relation between the characteristic function and the distribution (see Theorem 1.14).

Definition 1.17 *Let $\xi = (\xi^1, \ldots, \xi^s)^\tau$ be a s-dimensional random vector, $s \in \mathbb{N}$. Then the quantitiy*

$$m_{j_1 j_2 \cdots j_s} = \mathbf{E}\,(\xi^1)^{j_1}(\xi^2)^{j_2} \cdot \ldots \cdot (\xi^s)^{j_s}, \quad j_1, \ldots, j_s \in \mathbb{N} \cup \{0\}$$

is called **moment of order** $q = j_1 + \cdots + j_s$ *of ξ if the expectation does exist.*

Remark 1.7 The existence of $\mathbf{E}\,|\xi^k|^p < \infty$ for $k = 1, \ldots, s$, implies the existence of all moments of order $q \leq p$.

Proof: see basic course, material between Theorem 6.7 (Polya) and Definition 6.2

Theorem 1.17 *If* $\mathbf{E}\,|\xi^k|^p < \infty$ *for some* $p \geq 1$ *then for nonnegative integer numbers* j_k *it holds*

$$m_{j_1\ldots j_s} = (-i)^q \frac{\partial^q J(u)}{\partial u_1^{j_1}\ldots\partial u_s^{j_s}}\bigg|_{u=0} \quad \textit{where} \quad q = j_1 + \ldots + j_s \quad \textit{for} \quad q \leq p.$$

Conversely, if for an even $p \in \mathbb{N}$ *the characteristic function* $J(u_1,\ldots,u_s)$ *is p-times differentiable at* $u = 0$ *then there exist all moments of order* $q \leq p$ *and it holds the above relation.*

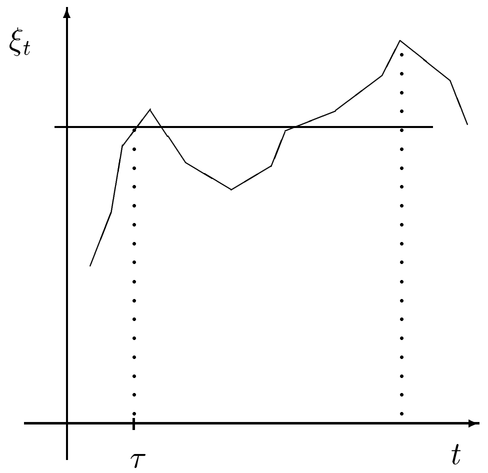
Proof:

part 1: follows from $J(u) = \mathbf{E}\, e^{i(u,\xi)}$ and the interchangeability of differentation and expectation (integation w.r.t.$\mathbf{P}$)

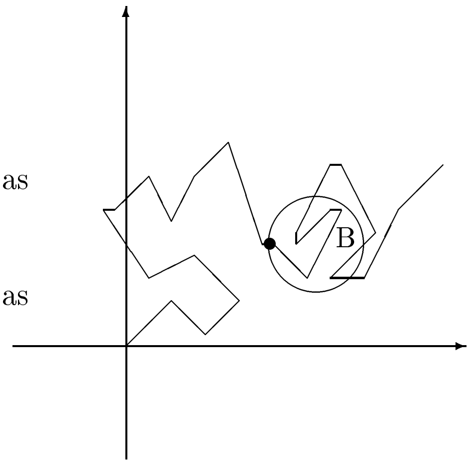
part 2: see basic course, Theorem 6.2, property 6

Stopping times

Example 1.3 stock price
$(\xi_t, t \in [0,T])$
a reasonable question:
when does the price reach a certain level for the first time
no reasonable question:
when does the price reach its maximum value in $[0,T]$
(can not be answered before T)



Example 1.4 Brownian motion
$(\xi_t, t \in \mathbb{R}_+)$
a reasonable question:
at which time a certain set B has been crossed for the first time
no reasonable question:
at which time a certain set B has been crossed for the last time



Definition 1.18 *Let* $\{\Omega, \mathcal{A}, \mathbf{P}\}$ *be a probability space and* $\{\mathcal{F}_t,\ t \in T\}$ *be a* **family of σ-algebras** *with* $\mathcal{F}_t \subset \mathcal{A} \quad \forall\, t \in T, \quad T \subseteq \mathbb{R}$.
$\{\mathcal{F}_t,\ \ t \in T\}$ *is called a* **filtration**, *if* $\mathcal{F}_{t_1} \subset \mathcal{F}_{t_2}$ *for* $t_1 < t_2$.

Remark 1.8 *interpretation:*

- $\mathcal{F}_t$ contains all those random events of $\mathcal{A}$ whose occurence or non-occurence can be decided up to time t
- filtration $\leftrightarrow$ information structure: $\mathcal{F}_t$ represents the information available at time t,
- filtration is non-decreasing: more and more information is known as time goes forward, past information is not forgotten

Definition 1.19 *A random variable $\tau(\omega) : \Omega_\tau \subset \Omega \to T$ is called a* **stopping time** *with respect to the filtration $\{\mathcal{F}_t,\ t \in T\}$ if*

$$\{\omega : \tau \le t\} \in \mathcal{F}_t \quad \forall\, t \in T.$$

Remark 1.9

1. $\{\omega : \tau \le t\} \in \mathcal{F}_t$,
   the occurence of the event "stopping time $\tau \le t$" can be decided using events of $\mathcal{F}_t$
2. Ω_τ : domain of definition of the stopping time τ with values in T
   $\Omega_\tau = \{\omega :\ \tau \in T\}$, i.e. τ is in the "observation period" T
   - if there is a maximum value $t_{\max}$ in T $\Rightarrow$ $\Omega_\tau = \{\tau \le t_{\max}\} \in \mathcal{F}_{t_{\max}}$
   - if there is no maximum value in T, let $t_k \uparrow \sup\{t :\ t \in T\}$
     $$\Rightarrow \quad \Omega_\tau = \bigcup_{k=1}^{\infty} \{\tau \le t_k\} \in \mathcal{A}$$
3. $\{\tau \le t\} \in \mathcal{F}_t \iff \{\tau > t\} \in \mathcal{F}_t$
   T countable: $\{\tau \le t\} \in \mathcal{F}_t \iff \{\tau = t\} \in \mathcal{F}_t \ \ \forall\, t \in T$

$\mathcal{F}_\tau := \{B \in \mathcal{A} :\ B \cap \{\tau \le t\} \in \mathcal{F}_t \ \ \forall\, t \in T\} \quad \Rightarrow \quad \mathcal{F}_\tau$ is a σ-algebra
$\mathcal{F}_\tau$ is the σ-algebra which is generated by the stopping time τ.
The stopping time τ is $\mathcal{F}_\tau$-measurable ($B = \Omega$).

Example 1.5 $\tau = t_0$, $t_0 \in T$ is a stopping time

$$\Rightarrow \quad \{\tau \le t\} = \begin{cases} \emptyset & \text{for} \quad t < t_0 \\ \Omega & \text{for} \quad t \ge t_0 \end{cases} \qquad \Rightarrow \quad B \cap \{\tau \le t\} = \begin{cases} \emptyset & t < t_0 \\ B & t \ge t_0 \end{cases}$$

$\Rightarrow \quad \mathcal{F}_\tau = \mathcal{F}_{t_0}$ see

1. $B \in \mathcal{F}_\tau$, i.e. $\forall\, t \ \ B \cap \{\tau \le t\} \in \mathcal{F}_t \overset{t=t_0}{\Longrightarrow} B \cap \{\tau \le t_0\} = B \in \mathcal{F}_{t_0}$
   i.e. $\mathcal{F}_\tau \subset \mathcal{F}_{t_0}$
2. $B \in \mathcal{F}_{t_0} :\ t < t_0 \ \Rightarrow \ B \cap \{\tau \le t\} = \emptyset \in \mathcal{F}_t$;
   $t \ge t_0 \ \Rightarrow \ B \cap \{\tau \le t\} = B \in \mathcal{F}_{t_0} \subset \mathcal{F}_t$, i.e. $B \in \mathcal{F}_\tau \Rightarrow \mathcal{F}_\tau \subset \mathcal{F}_{t_0}$

Theorem 1.18 (**Properties of stopping times**)
Let τ be a stopping time with respect to the filtration $\{\mathcal{F}_t,\ t \in T\}$.

1. *Let K be a Borel set of $\mathbb{R}$ with* $\sup\limits_{x \in K} x \le t$ *then* $\{\tau \in K\}$ *is* $\mathcal{F}_t$*-measurable.*
2. *Let* $\theta(t) : T \to T$ *be a Borel function with* $\theta(t) \ge t \quad \forall\, t \in T$ *then* $\theta(\tau)$ *is a stopping time.*
3. *Let* τ_1, τ_2 *be stopping times then* $\min(\tau_1, \tau_2), \quad \max(\tau_1, \tau_2)$ *are stopping times especially: if* τ *is a stopping time then it is* $\min(\tau, t_0)$, $t_0 \in T$.
4. *Let* τ_1, τ_2 *be stopping times then* $\tau_1 \le \tau_2$ *implies* $\mathcal{F}_{\tau_1} \subset \mathcal{F}_{\tau_2}$.

Proof:

1. $\{\tau \in K\} \subset \{\tau \le t\} : \ \{t_1 \le \tau \le t_2\}, \{t_1 < \tau < t_2\}, \ \{t_1 < \tau \le t_2\}, \ t_1, t_2 \le t$
$\mathcal{F}_t$-measuable;

$$K = \bigcup_{l=1}^{\infty} \bigcap_{k=1}^{p_l} A_{lk} : p_l \in N; \quad A_{lk} \text{ open (resp. closed intervals)}$$

$$\Rightarrow \quad \sup_{x \in A_{lk}} x \le t; \quad \{\tau \in K\} = \bigcup_{l=1}^{\infty} \bigcap_{k=1}^{p_l} \{\tau \in A_{lk}\} \quad \mathcal{F}_t\text{-measurable}$$

2. $K = \{s : \ \theta(s) \le t\}$ is a Borel set since θ is a Borel function;

$$\sup_{x \in K} x \le t, \text{ since } t \le \theta(t) \qquad (\text{see } x \in K \quad \Rightarrow \quad x \le \theta(x) \le t)$$

1. $\Longrightarrow \quad \{\tau \in K\} \ \ \mathcal{F}_t$-measurable; $\quad \{\tau \in K\} = \{\omega : \ \theta(\tau) \le t\} \in \mathcal{F}_t$

3. $$\{\max(\tau_1, \tau_2) \le t\} = \underbrace{\{\tau_1 \le t\}}_{\in \mathcal{F}_t} \cap \underbrace{\{\tau_2 \le t\}}_{\in \mathcal{F}_t} \in \mathcal{F}_t$$

$$\{\min(\tau_1, \tau_2) \le t\} = \{\tau_1 \le t\} \cup \{\tau_2 \le t\} \in \mathcal{F}_t$$

4. let $A \in \mathcal{F}_{\tau_1}$; $\quad$ because of $\ \{\tau_1 \le t\} \supset \{\tau_2 \le t\}$ it follows

$$A \cap \{\tau_2 \le t\} = [A \cap \{\tau_1 \le t\}] \cap \{\tau_2 \le t\} = B \cap \{\tau_2 \le t\} \in \mathcal{F}_t \quad \Rightarrow \quad A \in \mathcal{F}_{\tau_2}$$

$$(\text{see. } B = A \cap \{\tau_1 \le t\} \in \mathcal{F}_t; \quad \{\tau_2 \le t\} \in \mathcal{F}_t)$$

Let T be a finite or countable set, $\{\mathcal{F}_t, t \in T\}$ be a filtration,
τ stopping time with respect to $\{\mathcal{F}_t, t \in T\}$, $\Omega_\tau = \Omega$;
$\{\xi_t, t \in T\}$ family of r.v.'s such that ξ_t is $\mathcal{F}_t$-measurable $\quad \forall\, t \in T$,

$$\xi_\tau := \xi_t \quad \text{for} \quad \tau = t \qquad \text{i.e. } \xi_\tau \quad \text{is defined} \quad \forall\, \omega \in \Omega$$

$\Longrightarrow$ the r.v. ξ_τ is $\mathcal{F}_\tau$ - measurable.

see: $c_k, \quad k = 1, 2, \ldots$ are the possible values of τ

$$\begin{aligned} \{\omega : \ \xi_\tau < x\} \cap \{\omega : \ \tau \le t\} &= \bigcup_{c_k \le t} (\{\omega : \ \xi_\tau < x\} \cap \{\omega : \ \tau = c_k\}) \\ &= \bigcup_{c_k \le t} (\underbrace{\{\omega : \ \xi_{c_k} < x\}}_{\in \mathcal{F}_{c_k}} \cap \underbrace{\{\omega : \ \tau = c_k\}}_{\in \mathcal{F}_{c_k}}) \ \in \ \mathcal{F}_t \end{aligned}$$

$$\Rightarrow \quad \{\omega : \ \xi_\tau < x\} \ \in \ \mathcal{F}_\tau \qquad\qquad \Rightarrow \quad \text{assertion}$$

1.2 Independence

Definition 1.20 *Let $\{\Omega, \mathcal{A}, P\}$ be a probability space.*
The random events $A, B \in \mathcal{A}$ are called **independent** *if*

$$\mathbf{P}(A \cap B) \;=\; \mathbf{P}(A)\mathbf{P}(B).$$

Theorem 1.19 Properties of independent events
Let $A,\ B_i,\quad i = 1, 2, \ldots \in \mathcal{A}$

1. $\Omega,\ A$ *are independent*
2. $N,\ A \quad (\mathbf{P}(N) = 0)$ *are independent*
3. *Let $A,\ B_i$ be independent for $i = 1, 2,\quad B_1 \supset B_2$*
   $\Rightarrow\quad A,\ B_1 \setminus B_2$ *are independent (especially: $A,\ \overline{B_2}$ are independent)*
4. *Let $A,\ B_i$ be independent for $i = 1, 2, \ldots, n$;*
   $B_1, \ldots, B_n$ pairwise disjoint
   $\Rightarrow\quad A,\ \bigcup\limits_{i=1}^{n} B_i$ *independent*
   (without "'paarwise disjoint"' the statement is not true)
5. *A is independent of A if and only if $\mathbf{P}(A) = 0$ or $\mathbf{P}(A) = 1$*

Proof:

1. $\mathbf{P}(\Omega \cap A) = \mathbf{P}(A) = \mathbf{P}(\Omega)\mathbf{P}(A)$
2. $\mathbf{P}(N \cap A) = 0 = \mathbf{P}(N)\mathbf{P}(A)$
3. $$\begin{aligned}\mathbf{P}(A \cap (B_1 \setminus B_2)) &= \mathbf{P}((A \cap B_1) \setminus (A \cap B_2)) \\ &= \mathbf{P}(A \cap B_1) - \mathbf{P}(A \cap B_2) = \mathbf{P}(A)(\mathbf{P}(B_1) - \mathbf{P}(B_2)) \\ &= \mathbf{P}(A)\mathbf{P}(B_1 \setminus B_2)\end{aligned}$$
   especially: let $B_1 = \Omega$, A, Ω and A, B_2 are independent, $\Omega \supset B_2$
   $\Rightarrow\quad A, \Omega \setminus B_2 = \overline{B_2}$ are independent
4. $$\begin{aligned}\mathbf{P}(A \cap \bigcup_{i=1}^{n} B_i) &= \mathbf{P}(\bigcup_{i=1}^{n}(A \cap B_i)) = \sum_{i=1}^{n} \mathbf{P}(A \cap B_i) = \sum_{i=1}^{n} \mathbf{P}(A)\mathbf{P}(B_i) \\ &= \mathbf{P}(A)\mathbf{P}(\bigcup_{i=1}^{n} B_i) \qquad (\text{see } (A \cap B_k) \cap (A \cap B_h) = \emptyset \text{ for } k \neq h)\end{aligned}$$
5. A, A are independent $\quad\Rightarrow\quad \mathbf{P}(A \cap A) = \mathbf{P}(A)\mathbf{P}(A) \quad\Rightarrow\quad \mathbf{P}(A) = \mathbf{P}(A)^2$
   $\Rightarrow\quad \mathbf{P}(A) = 0$ or $\mathbf{P}(A) = 1 \qquad\qquad \Rightarrow\quad$ assertion

Let I be an index set and $\{\mathcal{M}_i, i \in I\}$ be classes of events

Definition 1.21 (independence of n events)
For $n = 2, 3, \ldots$ the random events $A_1, \ldots, A_n \in \mathcal{A}$ are called **(jointly) independent** *if for $k = 1, \ldots, n$ and $1 \leq i_1 < i_2 \ldots < i_k \leq n$ it holds*

$$\mathbf{P}(A_{i_1} \cap \ldots \cap A_{i_k}) = \mathbf{P}(A_{i_1}) \cdot \ldots \cdot \mathbf{P}(A_{i_k}).$$

Definition 1.22 (independence of classes of events)
The classes of events $\{\mathcal{M}_i,\ i \in I\}$ are called **independent** *if for any $n = 2, 3, \ldots$ and arbitrary pairwise different indices $i_1, i_2, \ldots, i_n \in I$ arbitrary events $A_{i_k} \in \mathcal{M}_{i_k},\ k = 1, 2, \ldots, n,$ are independent.*

Remark 1.10

1. *Note, that the* pairwise *independence of events* $A_1, \ldots, A_n$ *does not imply their jointly independence.*

2. *In the above definition the set* I *may have an infinite number of elements.*

notation: $\sigma\{\mathcal{M}\}-$ minimal σ-algebra containing the events of $\mathcal{M}$

Definition 1.23 (independence of random variables)
Random variables $\{\xi_i,\ i \in I\}$ *are called* **(jointly) independent** *if*

- *the classes of events* $\{\mathcal{M}_i,\ i \in I\}$ *are independent where*

$$\mathcal{M}_i = \{\{\omega : \xi_i \le a\},\, a \in \mathbb{R}\}.$$

$$\Updownarrow \quad \textit{(equivalence)}$$

- *for any* $n = 2, 3, \ldots,$ *and pairwise different* $i_1, \ldots, i_n \in I$, *the joint distribution function of* $\xi_{i_1}, \ldots, \xi_{i_n}$ *is equal to the product of the distribution functions of* ξ_{i_k}*:*

$$\mathbf{P}(\xi_{i_1} \le a_1, \xi_{i_2} \le a_2, \ldots, \xi_{i_n} \le a_n) \;=\; \prod_{k=1}^{n} \mathbf{P}(\xi_{i_k} \le a_k), \quad a_1, \ldots, a_n \in \mathbb{R}.$$

Definition 1.24 (independence of classes of r.v.'s)
Let J *be an index set. Classes of r.v.'s* $\{\xi_i^\mu,\ i \in I_\mu\},\ \mu \in J$ *are called* **(mutually) independent**, *if the classes of events* $\{\mathcal{M}_\mu, \mu \in J\}$ *with*

$$\mathcal{M}_\mu = \{\{\omega : \xi_{i_1}^\mu \le a_1, \ldots, \xi_{i_n}^\mu \le a_n\} :\; n = 1, 2, \ldots;\; i_k \in I_\mu;\; a_k \in \mathbb{R}\},$$

are independent.

Definition 1.25 *The* σ*-algebra of events* $\sigma\{\xi_i,\ i \in I\}$, *generated by the events of the form*

$$\{\omega :\ \xi_{i_1} \le a_1, \ldots, \xi_{i_n} \le a_n\}, \quad n = 1, 2, \ldots; \quad i_k \in I; \quad a_k \in \mathbb{R}$$

is called the σ**-algebra generated by the class of r.v.'s** $\{\xi_i, i \in I\}$.

Remark 1.11 $\sigma\{\xi_i,\ i \in I\}$ is the minimal σ-algebra with respect to which all the ξ_i are random variables, i.e. all the functions $\xi_i(\omega)$ are measurable.

Special case: a single r.v. ξ (I contains one element only)
$\Rightarrow \quad \sigma\{\xi\} = \sigma\{\ \{\omega :\ \xi \le a\},\ \ a \in \mathbb{R}\}$

Theorem 1.20 *If classes of r.v.'s* $\{\xi_i^\mu,\ i \in I_\mu\},\ \mu \in J$ *are independent, then the collection of* σ*-algebras* $\sigma\{\xi_i^\mu,\ i \in I_\mu\},\ \mu \in J$ *and their closures* $\tilde{\sigma}\{\xi_i^\mu,\ i \in I_\mu\},\ \mu \in J$ *are independent.*

Proof: not given

Remark 1.12 *concept of closure: see measure theory, Theorem 1.6 shows that the closures $\tilde{\sigma}\{\ldots\}$ are σ-algebras*

Theorem 1.21 *Let $g_\mu(t_1,\ldots,t_s)$, $\mu \in J$, be a set of Borel functions of s real variables ($g_\mu : \mathbb{R}^s \to \mathbb{R}$). If the classes of r.v.'s $\{\xi_1^\mu,\ldots,\xi_s^\mu\}$, $\mu \in J$, are independent, then the r.v.'s $\eta_\mu := g_\mu(\xi_1^\mu,\ldots,\xi_s^\mu)$, $\mu \in J$, are also independent.*

Proof: Definition 1.24 $\Rightarrow$ the classes of events

$$\mathcal{M}_\mu \;=\; \{\ \{\omega:\ \xi_1^\mu \le a_1,\ldots,\xi_s^\mu \le a_s\},\ a_1,\ldots,a_s \in \mathbb{R}\} \quad \mu \in J$$

are independent. Since the g_μ are Borel functions it follows

$$\begin{aligned}\mathcal{N}_\mu &= \{\ \{\omega:\ \eta_\mu = g_\mu(\xi_1^\mu,\ldots,\xi_s^\mu) \le a\},\ a \in \mathbb{R}\}\\ &= \{\ \{\omega:\ (\xi_1^\mu,\ldots,\xi_s^\mu) \in g_\mu^{-1}(-\infty,a]\},\ a \in \mathbb{R}\} \ \subset\ \sigma\{\mathcal{M}_\mu\},\end{aligned}$$

Theorem 1.20 $\Rightarrow$ $\{\mathcal{N}_\mu, \mu \in J\}$ are independent
Definition 1.23 $\Rightarrow$ assertion

Definition 1.26 (independence of random elements)
The random elements $\{\xi_i,\ i \in I\}$ with $\xi_i(\omega) : \{\Omega,\mathcal{A}\} \to \{\mathcal{S}_i,\mathcal{B}_i\}$, $i \in I$, are called **(jointly) independent** *if for any $n = 2,3,\ldots$, any pairwise different $i_1,\ldots,i_n \in I$, and arbitrary $B_{i_k} \in \mathcal{B}_{i_k} \in I$ it holds*

$$\mathbf{P}\left(\bigcap_{k=1}^{n}\{\xi_{i_k} \in B_{i_k}\}\right) = \prod_{k=1}^{n}\mathbf{P}\left(\xi_{i_k} \in B_{i_k}\right).$$

It is possible to generalize the above definitions and statements concerning classes of r.v.'s to r.e.'s.
Let $\xi_k(\omega)$ be r.e.'s with values in $\{\mathcal{S}_k,\mathcal{B}_k\}$, $k = 1,2,\ldots,n$
Finite sequences of r.e.'s can be considered as a r.e. with values in $\left\{\prod_{k=1}^{n}\mathcal{S}_k, \mathcal{B}^{(n)}\right\}$,
where $\mathcal{B}^{(n)}$ denotes the product of σ-algebras $\mathcal{B}_1,\ldots,\mathcal{B}_n$
consider cylinder sets $C = B_1 \times B_2 \times \ldots \times B_n$, $B_i \in \mathcal{B}_i$

$$\Rightarrow \quad \{\omega:\ (\xi_1(\omega),\ldots,\xi_n(\omega)) \in C\} \;=\; \bigcap_{i=1}^{n}\{\omega:\ \xi_i(\omega) \in B_i\}$$

i.e. the preimage of C is $\mathcal{A}$-measurable
$\Rightarrow$ the preimage of any set of $\mathcal{B}^{(n)}$ is $\mathcal{A}$-measurable (since $\mathcal{B}^{(n)}$ is the minimal σ-algebra containing all sets C)

$$m_{12\ldots n}(B) := \mathbf{P}((\xi_1,\ldots,\xi_n) \in B), \quad B \in \mathcal{B}^{(n)}$$

denotes the measure on $\left\{\prod_{k=1}^{n}\mathcal{S}_k, \mathcal{B}^{(n)}\right\}$, induced by the sequence $(\xi_1,\ldots,\xi_n)$
Let $\xi_1,\ldots,\xi_n$ be independent r.e.'s. Definition 1.26 $\Rightarrow$

$$m_{12\ldots n}(B_1 \times B_2 \times \ldots \times B_n) = m_1(B_1)m_2(B_2)\cdot\ldots\cdot m_n(B_n)$$

where $m_k(B_k) = \mathbf{P}(\xi_k \in B_k)$.
Hahn's Theorem on the extension of measures (see measure theory Theorem 1.5) $\Rightarrow$ uniqueness of the continuation of $m_{12\ldots n}$ onto the σ - Algebra $\mathcal{B}^{(n)}$

$\Rightarrow \quad m_{12\ldots n}$ is the product measure on $\left\{\prod\limits_{k=1}^{n} \mathcal{S}_k, \mathcal{B}^{(n)}\right\}$

Conversely, if $m_{12\ldots n}$ coincides with the product of the measures $m_1, \ldots, m_n$ then the r.e.'s $\xi_1, \ldots, \xi_n$ are independent. $\Rightarrow$

Theorem 1.22 *The r.e.'s $\xi_1, \ldots, \xi_n$ are independent if and only if the measure $m_{1\ldots n}$ induced by the sequence $(\xi_1, \ldots, \xi_n)$ on the σ-algebra $\mathcal{B}^{(n)}$ is the product of the measures m_k, $k = 1, 2, \ldots, n$, induced by the elements ξ_k on $\mathcal{B}_k$.*

Theorem 1.23 *Let $g(x_1, x_2)$ be a $\mathcal{B}^{(2)}$-measurable finite real function and ξ_1, ξ_2 be independent r.e.'s with values in $\{\mathcal{S}_i, \mathcal{B}_i\}$, $i = 1, 2$, and let, moreover, $\mathbf{E}\,|g(\xi_1, \xi_2)| < \infty$. Then $\phi(x_1) = \mathbf{E}\, g(x_1, \xi_2)$ is a $\mathcal{B}_1$-measurable function and*

$$\mathbf{E}\, g(\xi_1, \xi_2) = \mathbf{E}\, \phi(\xi_1) \qquad or \qquad \mathbf{E}\, g(\xi_1, \xi_2) = \int\limits_{\mathcal{S}_1} m_1(dx_1) \int\limits_{\mathcal{S}_2} g(x_1, x_2) m_2(dx_2).$$

Proof:
$g(x_1, x_2)$ is a $\mathcal{B}^{(2)}$-measurable function, i.e. $g^{-1}(G) \in \mathcal{B}^{(2)}$ $\forall$ open subsets $G \subset \mathbb{R}$

$$\begin{aligned} \mathbf{E}\, g(\xi_1, \xi_2) &= \int\limits_{\mathcal{S}_1 \times \mathcal{S}_2} g(x_1, x_2) m_{12}(d(x_1, x_2)) \\ &\overset{\text{independence}}{=} \int\limits_{\mathcal{S}_1 \times \mathcal{S}_2} g(x_1, x_2) m_1(dx_1) m_2(dx_2) \\ &\overset{\text{Fubini}}{=} \int\limits_{\mathcal{S}_1} m_1(dx_1) \int\limits_{\mathcal{S}_2} g(x_1, x_2) m_2(dx_2) \end{aligned}$$

$\int\limits_{\mathcal{S}_2} g(x_1, x_2) m_2(dx_2) = \mathbf{E}\, g(x_1, \xi_2) = \phi(x_1)$ is $\mathcal{B}_1$ - measurable (Fubini's Theorem).

Corollary 1.1 *If ξ_1 and ξ_2 are independent r.v.'s with finite expectations, then*

$$\mathbf{E}\, \xi_1 \xi_2 = \mathbf{E}\, \xi_1 \mathbf{E}\, \xi_2 \qquad \textit{(see } g(x_1, x_2) = x_1\, x_2\textit{).}$$

<u>0-1-laws</u>

Theorem 1.24 *Let $(A_n)_{n \geq 1}$ be a sequence of events.*
If $\sum\limits_{n=1}^{\infty} \mathbf{P}(A_n) < \infty$, then the event $\overline{\lim\limits_n}\, A_n$ is of probability 0.

Proof: it holds $\overline{\lim\limits_n}\, A_n = \{\omega \in \Omega : \ \omega \in A_n$ for infinite many $n\} = \bigcap\limits_{m=1}^{\infty} \bigcup\limits_{n=m}^{\infty} A_n$.

$$\mathbf{P}(\overline{\lim_n}\, A_n) = \mathbf{P}(\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_n) = \lim_{m \to \infty} \mathbf{P}(\bigcup_{n=m}^{\infty} A_n) \leq \lim_{m \to \infty} \sum_{n=m}^{\infty} \mathbf{P}(A_n) = 0$$

(see Theroem 1.1, 6. and subadditivity of $\mathbf{P}$)

The following stronger result is valid in case of independent events.

Theorem 1.25 Borel-Cantelli's Lemma
Let $(A_n)_{n\geq 1}$ be a sequence of independent events, then it holds:

$$\text{(i) } \mathbf{P}(\overline{\lim_n}\, A_n) = 0, \quad \text{for} \quad \sum_{k=1}^{\infty} \mathbf{P}(A_k) < \infty$$
$$\text{(ii) } \mathbf{P}(\overline{\lim_n}\, A_n) = 1, \quad \text{for} \quad \sum_{k=1}^{\infty} \mathbf{P}(A_k) = \infty.$$

Proof: Theorem 1.24 $\Rightarrow$ (i), for (ii) see basic course Theorem 5.7.

Let $(\mathcal{D}_n)_{n\geq 1}$ be a sequence of independent σ-algebras.
let $A = \overline{\lim_n}\, A_n, \quad A_n \in \mathcal{D}_n \quad \stackrel{\text{Borel-Cantelli}}{\Longrightarrow} \quad \mathbf{P}(A) = 0$ or $\mathbf{P}(A) = 1$

generalization:

let $(\mathcal{C}_k)_{k\geq 1} := (\sigma\{\mathcal{D}_k, \mathcal{D}_{k+1}, \dots\})_{k\geq 1}$ be a monoton decreasing sequence of σ-algebras

$$\mathcal{C} = \bigcap_{k=1}^{\infty} \mathcal{C}_k =: \overline{\lim_n}\, \mathcal{D}_n \text{ is also a } \sigma\text{-algebra (tail or terminal } \sigma\text{-algebra)}$$

$\overline{\lim_n}\, \mathcal{D}_n$ remains unchanged if a finite number of σ-algebras $\mathcal{D}_k$ is changed.

Theorem 1.26 Kolmogorov's general 0-1-law
If $\mathcal{D}_n,\ n = 1, 2, \dots$ are mutually independent σ-algebras, then any event in $\overline{\lim_n}\, \mathcal{D}_n$ has probability either 0 or 1.

Proof: let $A \in \overline{\lim_n}\, \mathcal{D}_n = \bigcap_{k=1}^{\infty} \mathcal{C}_k$ where $\mathcal{C}_k = \sigma\{\mathcal{D}_k, \mathcal{D}_{k+1}, \dots\} \Rightarrow \quad A \in \mathcal{C}_k \ \ \forall\, k$

$\Rightarrow \quad A,\ \sigma\{\mathcal{D}_1, \dots, \mathcal{D}_{k-1}\}$ independent $\forall k$
$\Rightarrow \quad A,\ \sigma\{\mathcal{D}_1, \mathcal{D}_2, \dots\} = \mathcal{C}_1$ independent
since $A \in \sigma\{\mathcal{D}_1, \mathcal{D}_2, \dots\} \quad \Rightarrow \quad A,\ A$ independent
Theorem 1.19,6.$\Rightarrow$ either $\mathbf{P}(A) = 0$ or $\mathbf{P}(A) = 1 \Rightarrow$ assertion

Theorem 1.26 implies

Theorem 1.27 *Let $(\xi_n)_{n\geq 1}$ be a sequence of independent random elements in a fixed metric space $\{\mathcal{S}, \mathcal{B}\}$. Further, let $\mathcal{D}_n = \sigma\{\xi_n\}$ be the σ-algebra generated by ξ_n and $\mathcal{C}_k = \sigma\{\mathcal{D}_k, \mathcal{D}_{k+1}, \dots\}$. Then it holds*

1. *the limit of the sequence $(\xi_n)_{n\geq 1}$ exists either with probability 0 or 1,*

2. *if $\mathcal{S}$ is seperable and complete, then the a.s. limit of the sequence $(\xi_n)_{n\geq 1}$ is constant* (mod P) *provided the limit exists,*

3. *if $z = f(x_1, x_2, \dots)$ is a real function of infinitely many arguments $x_n \in \mathcal{S},\ n = 1, 2, \dots$ and $f(\xi_1, \xi_2, \dots, \xi_n, \dots)$ is $\mathcal{C}_n$-measurable $\forall\ n$, then the function f is constant with probability 1.*

Proof: not given

Remark 1.13 *Examples for functions $f(\xi_1, \xi_2, \ldots)$ in 3. are functions which do not depend on a finite number of variables:*

$$\overline{lim}_n \xi_n,\ \underline{lim}_n \xi_n,\ I_{\left\{\sum\limits_{n=1}^{\infty} \xi_n < \infty\right\}} = I_{\left\{\sum\limits_{n=k}^{\infty} \xi_n < \infty\right\}}$$

$$\textit{not:} \quad I_{\left\{\sum\limits_{n=1}^{\infty} \xi_n < c\right\}},\ I_{\{\xi_1 = \xi_2 = \ldots = 0\}}$$

1.3 Conditional probability and conditional expectation

1.3.1 Conditional expectation

elementary case: conditional probability of an event A given B with $\mathbf{P}(B) \neq 0$)

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}$$

for fixed B is $\mathbf{P}(.|B)$ a measure on the $\sigma-$algebra $\mathcal{A}$ (as the "unconditional" probability $\mathbf{P}(.)$)

let $\xi(\omega)$ be a random variable then $\mathbf{E}\{\xi|B\} := \int\limits_{\Omega} \xi(\omega)\mathbf{P}(d\omega|B)$ is called conditional expectation of ξ given B

$\Rightarrow \quad \mathbf{P}(B)\mathbf{E}\{\xi|B\} = \int\limits_{B} \xi d\mathbf{P}$ $\qquad (+)$

see ξ discrete: $\quad A_k = \{\omega : \; \xi = x_k\} \quad \Rightarrow$

$$\begin{aligned} \mathbf{E}\{\xi|B\} &= \int\limits_{\Omega} \xi\, \mathbf{P}(d\omega|B) = \sum_k x_k \mathbf{P}(A_k|B) \\ &= \frac{1}{\mathbf{P}(B)} \sum_k x_k \mathbf{P}(A_k \cap B) = \frac{1}{\mathbf{P}(B)} \int\limits_{B} \xi\, \mathbf{P}(d\omega) \end{aligned}$$

especially for $\xi = I_A(\omega) \quad \Rightarrow$

$$\mathbf{E}\{I_A|B\} = \frac{1}{\mathbf{P}(B)} \int\limits_{B} I_A\, d\mathbf{P} = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)} = \mathbf{P}(A|B)$$

i.e. conditional probability is a special case of the conditional expectation

How can the above definitions be extended to the case $\mathbf{P}(B) = 0$?

Let $\mathcal{M} = \{B_i \in \mathcal{A} : \; \bigcup\limits_{i=1}^{\infty} B_i = \Omega, \;\; B_i \cap B_j = \emptyset, \;\; i \neq j\}$ be a countable class of disjoint events

$\Rightarrow \quad \mathbf{E}\{\xi|\mathcal{M}\}(\omega) := \mathbf{E}\{\xi|B_i\}$ for $\omega \in B_i$ is a random variable

$\mathbf{E}\{\xi|\mathcal{M}\}$ is called **conditional expectation of the r.v. ξ given $\mathcal{M}$**

- $\mathbf{E}\{\xi|\mathcal{M}\}$ is defined with probability 1 (a.e. on Ω), since:
  - for $B_i \in \mathcal{M}$, $\mathbf{P}(B_i) \neq 0$ it holds $\mathbf{E}\{\xi|\mathcal{M}\} = \mathbf{E}\{\xi|B_i\} = \text{const}$
  - $\mathbf{E}\{\xi|\mathcal{M}\}$ is not defined only for B_i with $\mathbf{P}(B_i) = 0$
- $\int\limits_{B_i} \xi d\mathbf{P} = \int\limits_{B_i} \mathbf{E}\{\xi|\mathcal{M}\} d\mathbf{P} \qquad \forall\, B_i \in \mathcal{M}$ $\qquad (\circ)$

  see $(+)$ and $\int\limits_{B_i} \mathbf{E}\{\xi|\mathcal{M}\} d\mathbf{P} = \mathbf{P}(B_i)\mathbf{E}\{\xi|B_i\}$ ($\mathbf{E}\{\xi|\mathcal{M}\}$ is constant on B_i)

  $\Rightarrow$ it also holds: $\int\limits_{B} \xi d\mathbf{P} = \int\limits_{B} \mathbf{E}\{\xi|\mathcal{M}\} d\mathbf{P} \qquad \forall\, B \in \sigma\{\mathcal{M}\}, \quad \mathbf{P}(B) \neq 0$ $(*)$

see: $\forall\ B \in \sigma\{\mathcal{M}\}\ \exists$ a representation $B = \bigcup\limits_{k=1}^{\infty} B_{j_k},\ B_{j_k} \in \mathcal{M}$

$$\int\limits_B \mathbf{E}\{\xi|\mathcal{M}\}d\mathbf{P} \;=\; \sum_{k=1}^{\infty} \underbrace{\int\limits_{B_{j_k}} \mathbf{E}\{\xi|\mathcal{M}\}d\mathbf{P}}_{=\quad(o)}$$

$$\int\limits_B \xi d\mathbf{P} \;=\; \sum_{k=1}^{\infty} \overbrace{\int\limits_{B_{j_k}} \xi d\mathbf{P}}$$

- the r.v. $\eta = \mathbf{E}\{\xi|\mathcal{M}\}$ is measurable with respect to $\sigma\{\mathcal{M}\}$

  see $\eta(\omega) = \mathbf{E}\{\xi|\mathcal{M}\} = \mathbf{E}\{\xi|B_k\}$ for $\omega \in B_k, \quad B_k \in \mathcal{M}$

  $$\Rightarrow \quad \{\omega :\ \eta \le a\} = \bigcup_{E\{\xi|B_k\}\le a} B_k \;\in\; \sigma\{\mathcal{M}\} \quad \Rightarrow \quad \text{measurability}$$

  (η is a r.v. with only countable many values $\mathbf{E}\{\xi|B_k\}$)

- the r.v. $\eta = \mathbf{E}\{\xi|\mathcal{M}\}$ is uniquely defined $(\mathrm{mod}\, P)$

  see: let η_1, η_2 be measurable w.r.t. some σ-algebra $\mathcal{F}$ then

  $\int\limits_B \eta_1 d\mathbf{P} = \int\limits_B \eta_2 d\mathbf{P} \quad \forall\ B \in \mathcal{F} \quad \Rightarrow \quad \eta_1 = \eta_2\ (\mathrm{mod}\, P)$

  since: if $\mathbf{P}(\omega :\ \eta_1 - \eta_2 > 0) > 0 \quad \Rightarrow \quad \int\limits_{\{\omega:\ \eta_1-\eta_2>0\}} (\eta_1 - \eta_2)d\mathbf{P} > 0$

  $\Rightarrow$ contradicition

  analogously $\quad \mathbf{P}(\omega :\ \eta_1 - \eta_2 < 0) > 0$

Definition 1.27 *Let $\mathcal{B}$ be a arbitrary σ-algebra of events contained in $\mathcal{A}$ and ξ an arbitrary random variable for which the expectation $\mathbf{E}\,\xi$ exists. The* **conditional expectation of ξ given the σ-algebra $\mathcal{B}$** *is the random variable $\mathbf{E}\{\xi|\mathcal{B}\}$ which is*

- *measurable with respect to $\mathcal{B}$*
- *satisfying the equality* $\int\limits_B \mathbf{E}\{\xi|\mathcal{B}\}d\mathbf{P} \;=\; \int\limits_B \xi d\mathbf{P} \quad \forall \quad B \in \mathcal{B}.$

Remark 1.14

1. *In general, ξ is not measurable with respect to $\mathcal{B}$ (but only w.r.t. $\mathcal{A}$)*

2. *Existence and uniqueness follows directly from the Radon-Nikodym theorem: $\int\limits_B \xi d\mathbf{P}$ is a σ-finite, countable-additive set function on $\mathcal{B}$ which is absolutely continuous w.r.t. $\mathbf{P}$*

   $\Rightarrow \quad \exists\ $ *$\mathcal{B}$-measurable function $\eta(\omega)$ such that* $\int\limits_B \xi d\mathbf{P} = \int\limits_B \eta\, d\mathbf{P}$ *and*

   $\eta(\omega)$ *is unique* $(mod\,\mathbf{P}) \qquad \Rightarrow \quad \eta = \mathbf{E}\{\xi|\mathcal{B}\}$

3. *Let $\tilde{\mathcal{B}}$ the completion of $\mathcal{B}$ w.r.t. $\mathbf{P}$, then it holds*

   $$\mathbf{E}\{\xi|\mathcal{B}\} \;=\; \mathbf{E}\{\xi|\tilde{\mathcal{B}}\} \quad (mod\,\mathbf{P}).$$

Definition 1.28 *For a fixed event $A \in \mathcal{A}$ the* **conditional probability** $\mathbf{P}\{A|\mathcal{B}\}$ *is a $\mathcal{B}$-measurable random variable satisfying*

$$\int_B \mathbf{P}\{A|\mathcal{B}\}\, d\mathbf{P} = \mathbf{P}(A \cap B) \quad \forall \;\; B \in \mathcal{B}.$$

see Definition 1.27 with $\xi = I_A(\omega)$: $\quad \mathbf{E}\{I_A|\mathcal{B}\} = \mathbf{P}\{A|\mathcal{B}\}$

Example 1.6 random experiment: throwing of two dice

$\Omega = \{\omega_1, \dots, \omega_{36}\},\ \omega_l = (i,j),$
$1 \le i,j \le 6,\ (l = 6(i-1)+j)$

$\mathcal{A} = \mathcal{P}(\Omega), \quad \mathrm{card}(\mathcal{A}) = 2^{36}$

$\mathcal{M} = \{B_2, \dots, B_{12}\}$
with $B_k = \{\omega = (i,j) :\ i+j = k\}$
(sum of the two numbers $= k$)

a_1/a_2	1	2	3	4	5	6
1	B_2	B_3	B_4	B_5	B_6	B_7
2	B_3	B_4	B_5	B_6	B_7	B_8
3	B_4	B_5	B_6	B_7	B_8	B_9
4	B_5	B_6	B_7	B_8	B_9	B_{10}
5	B_6	B_7	B_8	B_9	B_{10}	B_{11}
6	B_7	B_8	B_9	B_{10}	B_{11}	B_{12}

$\mathcal{B} = \sigma\{\mathcal{M}\} = \mathcal{P}(\mathcal{M}) \subset \mathcal{A}$

ξ : number of the first die; $\quad \mathbf{E}\,\xi = 3.5$

$\mathbf{E}\{\xi|\mathcal{B}\}$ is a r.v. with values $\mathbf{E}\{\xi|B_k\} = \frac{1}{\mathbf{P}(B_k)} \int\limits_{B_k} \xi d\mathbf{P} = \frac{k}{2} \qquad$ on $\quad B_k$

$$
\begin{array}{lrcl}
\text{see} \quad \forall B \in \mathcal{B}: & \int_B \mathbf{E}\{\xi|\mathcal{B}\} d\mathbf{P} & = & \int_B \xi d\mathbf{P} \\
B = B_k: & \mathbf{E}\{\xi|B_k\}\underbrace{\mathbf{P}(B_k)} & = & \int_{B_k} \xi d\mathbf{P} \\
2 \le k \le 7: & \mathbf{E}\{\xi|B_k\}\frac{k-1}{36} & = & \sum\limits_{i=1}^{6} i\,\underbrace{\mathbf{P}(A_i \cap B_k)}_{\frac{1}{36} \text{ or } 0} = \sum\limits_{i=1}^{k-1} i\frac{1}{36} = \frac{k(k-1)}{2}\frac{1}{36} \\
 & A_i & = & \{\omega : \xi(\omega) = i\} = \{(i,j) : 1 \le j \le 6\} \\
 & A_i \cap B_k & = & \{(i,j) : 1 \le i,j \le 6,\ i+j = k\} \\
 & \Rightarrow \mathbf{E}\{\xi|B_k\} & = & \frac{k}{2} \\
8 \le k \le 12: & \text{analogously} & &
\end{array}
$$

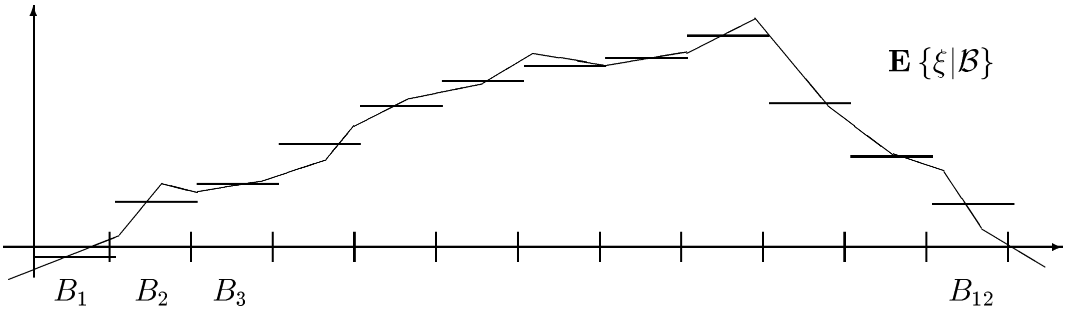
$\mathbf{E}\{\xi|\mathcal{B}\}$ is $\mathcal{B}$-measurable (it is constant on $B_k,\ k = 2, \dots, 12$)
ξ is not $\mathcal{B}$-measurable (it is not constant on $B_k,\ k = 2, \dots, 12$)

Example 1.7 random experiment: observation of midday temperature

ξ midday temperature;
$\Omega = \{\omega_1, \dots, \omega_{365}\}$ observations of one year
$\mathcal{A} = \mathcal{P}(\Omega)$,

$$
\mathcal{M} = \{B_1, \dots, B_{12}\} \quad \text{with} \quad
\begin{array}{ll}
B_1 = \{\omega_1, \dots, \omega_{31}\} & \text{january} \\
B_2 = \{\omega_{32}, \dots, \omega_{59}\} & \text{february} \\
\vdots & \\
B_{12} = \{\omega_{335}, \dots, \omega_{365}\} & \text{dezember}
\end{array}
\qquad \mathcal{B} = \mathcal{P}(\mathcal{M}) :
$$

$$\Rightarrow \quad \mathbf{E}\{\xi|\mathcal{B}\} \overset{\omega \in B_k}{=} \frac{1}{\mathbf{P}(B_k)} \int\limits_{B_k} \xi d\mathbf{P} \qquad \hat{=} \quad \text{monthly average}$$



Example 1.8 US presidential elections
(in memoriam to the 2000-elections between of G.W. Bush and A. Gore)

$$
\begin{array}{rl}
n & \text{persons are entitled to vote } (n \approx 200 \text{ millions}) \\
m & \text{candidates, assume } m = 2 \\
r & \text{states (groups of persons, r=50)} \\
\Omega = & \{\omega_1, \dots, \omega_n\}, \quad \mathcal{A} = \mathcal{P}(\Omega), \quad \mathbf{P}(\omega_i) = \frac{1}{n},\ i = 1 \dots, n \\
A = & \{\omega\colon \text{ person } \omega \text{ votes in favour of candidate 1 }\} \\
\overline{A} = & \{\omega\colon \text{ person } \omega \text{ votes in favour of candidate 2 }\} \\
\mathbf{P}(A) = & \frac{1}{n}\,\mathrm{card}\,A = p, \quad \mathbf{P}(\overline{A}) = 1 - p \\
\xi = & I_A, \quad \mathbf{E}\,\xi = \mathbf{P}(A) \\
& \text{if } p > 0.5 \text{ then 1 is not necessarily the winner !} \\
B_k = & \{\omega\colon \text{ person } \omega \text{ lives in state } k\},\ \mathbf{P}(B_k) = \frac{1}{n}\,\mathrm{card}\,B_k,\ k = 1 \dots, r \\
\mathcal{M} = & \{B_1, \dots, B_r\}, \quad \mathcal{B} = \mathcal{P}(\mathcal{M})
\end{array}
$$

$\mathbf{E}\{\xi|\mathcal{B}\} = \mathbf{P}\{A|\mathcal{B}\}$ is a r.v. with values

$$\mathbf{E}\{\xi|B_k\} = \frac{1}{\mathbf{P}(B_k)} \int\limits_{B_k} \xi\, d\mathbf{P} = \frac{1}{\mathbf{P}(B_k)} \mathbf{P}(A \cap B_k) = \mathbf{P}(A|B_k) =: p_k$$

if $p_k > 0.5$ then 1 gets c_k electors, $c_k \in \mathbb{N}, c_k \sim \mathbf{P}(B_k)$
assume $\mathbf{P}(B_k) = c_k \left(\sum\limits_{k=1}^{r} c_k\right)^{-1}$
percentage of electors voting in favour of 1 is

$$\sum_{k=1}^{r} I_{\{\mathbf{E}\{\xi|B_k\} > 0.5\}} \mathbf{P}(B_k) = \mathbf{E}\, I_{\{\mathbf{E}\{\xi|B_k\} > 0.5\}}$$

in general it holds

$$\mathbf{P}(A) = \mathbf{E}\,\xi = \mathbf{E}\,\mathbf{E}\{\xi|\mathcal{B}\} \neq \mathbf{E}\, I_{\{\mathbf{E}\{\xi|B_k\} > 0.5\}}$$

$\Rightarrow$ if 1 has the majority of votes then 1 gets not necessarily the majority of electors

Theorem 1.28 (properties of conditional expectation)
(see basic course, Theorem 4.2)
Let ξ be a r.v. on $\{\Omega, \mathcal{A}, P\}$ with $\mathbf{E}\,|\xi| < \infty$ and $\mathcal{B}$ a σ-algebra contained in $\mathcal{A}$ Then it holds with probability 1:

1. $\xi \geq 0 \quad \Rightarrow \quad \mathbf{E}\{\xi|\mathcal{B}\} \geq 0; \qquad \xi = 0 \quad \Rightarrow \quad \mathbf{E}\{\xi|\mathcal{B}\} = 0;$

2. ξ *is* $\mathcal{B}$*-measurable* $\quad \Rightarrow \quad \mathbf{E}\{\xi|\mathcal{B}\} = \xi;$
   especially: $B \in \mathcal{B} \quad \Rightarrow \quad \mathbf{P}\{B|\mathcal{B}\} = I_B;$

3. $\mathbf{E}\,\mathbf{E}\{\xi|\mathcal{B}\} = \mathbf{E}\,\xi;$

4. $\mathbf{E}\,|\xi_1|, \mathbf{E}\,|\xi_2| < \infty,\ a, b \in \mathbb{R} \ \Rightarrow \quad \mathbf{E}\{a\xi_1 + b\xi_2|\mathcal{B}\} = a\mathbf{E}\{\xi_1|\mathcal{B}\} + b\mathbf{E}\{\xi_2|\mathcal{B}\};$
   especially: $C_1 \cap C_2 = \emptyset \quad \Rightarrow \quad \mathbf{P}\{C_1 \cup C_2|\mathcal{B}\} = \mathbf{P}\{C_1|\mathcal{B}\} + \mathbf{P}\{C_2|\mathcal{B}\};$

5. $(\xi_n)_{n\geq 1}, \quad \xi_n \geq 0$ *monotonically increasing sequence of r.v.'s*
   $$\Rightarrow \quad \lim_{n\to\infty} \mathbf{E}\{\xi_n|\mathcal{B}\} = \mathbf{E}\{\lim_{n\to\infty} \xi_n|\mathcal{B}\};$$
   especially:
   (a): $(B_n)_{n\geq 1}$ *monotonically increasing sequence of events*
   $$\Rightarrow \quad \lim_{n\to\infty} \mathbf{P}\{B_n|\mathcal{B}\} = \mathbf{P}\{\bigcup_{n=1}^{\infty} B_n|\mathcal{B}\};$$
   (b): $(C_n)_{n\geq 1}$ *sequence of events* $\quad C_i \cap C_j = \emptyset, \quad i \neq j$
   $$\Rightarrow \quad \sum_{n=1}^{\infty} \mathbf{P}\{C_n|\mathcal{B}\} = \mathbf{P}\{\bigcup_{n=1}^{\infty} C_n|\mathcal{B}\};$$

6. $\xi, \ \mathcal{B}$ *independent* $\quad \Rightarrow \quad \mathbf{E}\{\xi|\mathcal{B}\} = \mathbf{E}\,\xi;$
   especially: $A, \ \mathcal{B}$ *independent* $\quad \Rightarrow \quad \mathbf{P}\{A|\mathcal{B}\} = \mathbf{P}(A);$

7. η *is* $\mathcal{B}$ *- measurable* $\quad \Rightarrow \quad \mathbf{E}\{\xi\eta|\mathcal{B}\} = \eta\mathbf{E}\{\xi|\mathcal{B}\};$

8. $\mathcal{B}_1 \subset \mathcal{B}_2 \quad \Rightarrow \quad \mathbf{E}\{\mathbf{E}\{\xi|\mathcal{B}_1\}|\mathcal{B}_2\} = \mathbf{E}\{\xi|\mathcal{B}_1\};$
   $\mathcal{B}_2 \subset \mathcal{B}_1 \quad \Rightarrow \quad \mathbf{E}\{\mathbf{E}\{\xi|\mathcal{B}_1\}|\mathcal{B}_2\} = \mathbf{E}\{\xi|\mathcal{B}_2\};$

9. $\xi(\omega): \ \{\Omega, \mathcal{A}\} \to \{\mathcal{S}, \mathcal{C}\}$*-measurable,* $\eta(\omega): \ \{\Omega, \mathcal{A}\} \to \{\mathcal{Z}, \mathcal{D}\}$*-measurable;*
   $g(x, z)$ *function on* $\mathcal{S} \times \mathcal{Z}$*; measurable w.r.t.* $\sigma\{\mathcal{C} \times \mathcal{D}\}$; $\quad \mathbf{E}\,|g(\xi, \eta)| < \infty;$
   $\Rightarrow$*if* η *is* $\mathcal{B}$*-measurable* $(\mathcal{B} \subset \mathcal{A})$*, then* $\quad \mathbf{E}\{g(\xi, \eta)|\mathcal{B}\} = \mathbf{E}\{g(\xi, z)|\mathcal{B}\}\big|_{z=\eta}.$

1.3.2 Conditional expectation with respect to random elements

starting point:
η discrete r.v. with values $y_1, y_2, \ldots, \ B_n = \{\omega: \ \eta = y_n\}; \quad \mathbf{P}(B_n) > 0$

$\mathbf{P}_n(A) := \mathbf{P}(A|B_n) = \frac{\mathbf{P}(A \cap B_n)}{\mathbf{P}(B_n)}$ conditional probability of $A \in \mathcal{A}$ given $\eta = y_n$

$\mathbf{E}\{\xi|\eta = y_n\} = \int\limits_{\Omega} \xi d\mathbf{P}_n = \frac{1}{\mathbf{P}(B_n)} \int\limits_{B_n} \xi d\mathbf{P}$ (see subsection 1.3.1)

$\Rightarrow \quad \mathbf{E}\{\xi|\eta\}$ is a r.v. depending on the outcome of the random experiment described by η: values $\mathbf{E}\{\xi|\eta = y_n\}$ for $\eta = y_n$

Definition 1.29 *Let* ξ *be a random variable with* $\mathbf{E}\,|\xi| < \infty$, $\eta: \{\Omega, \mathcal{A}\} \to \{\mathcal{S}, \mathcal{C}\}$ *be a random element and* $\mathcal{F}_\eta := \{A: \ A = \eta^{-1}(C), \ C \in \mathcal{C}\} \quad (\mathcal{F}_\eta \subset \mathcal{A})$*.*
The random variable $\mathbf{E}\{\xi|\mathcal{F}_\eta\}$ *is called the* **conditional expectation of the random variable $\boldsymbol{\xi}$ given the random element $\boldsymbol{\eta}$** *and is denoted by* $\mathbf{E}\{\xi|\eta\}$*.*

$$\Rightarrow \qquad \int\limits_{\eta^{-1}(C)} \mathbf{E}\{\xi|\eta\} d\mathbf{P} = \int\limits_{\eta^{-1}(C)} \xi d\mathbf{P} \qquad \forall \ C \in \mathcal{C} \qquad \text{(see Definition 1.27)}$$

conditional probability given $\boldsymbol{\eta}$: $\qquad \mathbf{P}\{A|\eta\} = \mathbf{P}\{A|\mathcal{F}_\eta\}$

Theorem 1.29 *The conditional expectation given the random element η is a $\mathcal{C}$-measurable function, i.e. there exists a $\mathcal{C}$-measurable function $s(y)$ such that $\mathbf{E}\{\xi|\eta\} = s(\eta)$.*

Proof:

$$
\begin{tikzcd}
\{\mathcal{S},\mathcal{C},P_\eta\} \arrow[drr, "\eta^{-1}"] \arrow[dd, bend left=80, dotted, "s(\eta)"] & & \\
\{\Omega,\mathcal{A},P\} \arrow[u, "\eta"] \arrow[d, "\xi(\omega)"] & \supset & \{\Omega,\mathcal{F}_\eta,P\} \arrow[dll, "\mathbf{E}\{\xi\,|\,\eta\}"] \\
\{\mathrm{R},\mathcal{B}\} & &
\end{tikzcd}
$$

Borel sets

without loss of generality let $\xi \geq 0$

$$\Rightarrow \quad q(C) := \int\limits_{\eta^{-1}(C)} \xi d\mathbf{P} = \int\limits_{\eta^{-1}(C)} \mathbf{E}\{\xi|\eta\} d\mathbf{P}$$

is a σ-finite measure on $\mathcal{C}$ (moreover, it is finite)

$$\text{see} \quad q(\mathcal{S}) = \int\limits_{\eta^{-1}(\mathcal{S})} \xi d\mathbf{P} = \int\limits_{\Omega} \xi d\mathbf{P} = \mathbf{E}\,\xi < \infty$$

it holds $q(C) = 0$ for $\mathbf{P}(\eta^{-1}(C)) = 0$,

i.e. q is absolutely continuous with respect to $\mathbf{P}_\eta$ where $\mathbf{P}_\eta(C) := \mathbf{P}(\eta^{-1}(C))$

$\overset{\text{Radon-Nikodym}}{\Longrightarrow}$ $\exists$ a $\mathcal{C}$-measurable nonnegative function $s(y)$ such that

$$q(C) = \int\limits_{C} s(y) d\mathbf{P}_\eta = \int\limits_{\eta^{-1}(C)} s(\eta(\omega)) d\mathbf{P}$$

(see rule of change of variables for integrals, basic course, Theorem 3.25)

$$\Longrightarrow \qquad \int\limits_{\eta^{-1}(C)} \mathbf{E}\{\xi|\eta\} d\mathbf{P} = \int\limits_{\eta^{-1}(C)} s(\eta) d\mathbf{P} \qquad \forall\ C \in \mathcal{C}$$

$\overset{\text{uniqueness}}{\Longrightarrow}$ $\mathbf{E}\{\xi|\eta\} = s(\eta)$

$s(\eta(\omega))$ is $\mathcal{F}_\eta$ - measurable since s it is $\mathcal{C}$-measurable $\qquad \Rightarrow$ assertion

Corollary 1.2 $\mathbf{E}\{\xi|\eta\} = s(\eta)$ *can be considered as a function $s = s(y)$ of the variable y on the space $\{\mathcal{S},\mathcal{C},\mathbf{P}_\eta\}$.*

see $s(\eta):\quad \{\Omega,\mathcal{A}\} \xrightarrow{\eta} \{\mathcal{S},\mathcal{C}\} \xrightarrow{s} \{R,\mathcal{B}\}$ and

$$\int\limits_{\eta^{-1}(C)} \xi d\mathbf{P} = \int\limits_{\eta^{-1}(C)} \mathbf{E}\{\xi|\eta\} d\mathbf{P} = \int\limits_{\eta^{-1}(C)} s(\eta) d\mathbf{P} = \int\limits_{C} s(y) \mathbf{P}_\eta(dy) \qquad \forall\ C \in \mathcal{C}$$

Theorem 1.30 properties of $\mathbf{E}\{\xi|\eta\}$

Let ξ be a r.v. with $\mathbf{E}|\xi| < \infty$ and η be a random element in $\{\mathcal{S},\mathcal{C}\} \Rightarrow$

1. *Let ξ, η be independent then $\mathbf{E}\{\xi|\eta\} = \mathbf{E}\,\xi$;*

2. *Let ξ be $\mathcal{F}_\eta$-measurable then $\mathbf{E}\{\xi|\eta\} = \xi$;*

3. Let $\eta_i(\omega)$ be measurable mappings $\{\Omega, \mathcal{A}\} \to \{\mathcal{S}_i, \mathcal{C}_i\}, \quad i = 1,2$ then

$$\mathbf{E}\{\mathbf{E}\{\xi|(\eta_1,\eta_2)\}|\eta_1\} = \mathbf{E}\{\xi|\eta_1\},$$

where $(\eta_1,\eta_2): \quad \{\Omega, \mathcal{A}\} \quad \longrightarrow \quad \{\mathcal{S}_1 \times \mathcal{S}_2, \sigma\{\mathcal{C}_1 \times \mathcal{C}_2\}\}$ *(product space)*

Proof:

1. if ξ, η independent, i.e. $\xi,\ \sigma\{\eta\} = \mathcal{F}_\eta$ independent
$$\Rightarrow \qquad \mathbf{E}\{\xi|\eta\} = \mathbf{E}\{\xi|\mathcal{F}_\eta\} = \mathbf{E}\,\xi \qquad \text{(see Theorem 1.28 6.)}$$
2. follows from Theorem 1.28 2.
3. $\mathbf{E}\{\mathbf{E}\{\xi|(\eta_1,\eta_2)\}|\eta_1\} = \mathbf{E}\{\mathbf{E}\{\xi|\mathcal{F}_{(\eta_1,\eta_2)}\}|\mathcal{F}_{\eta_1}\}$
$$\overset{\text{Theorem 1.28 8.}}{=} \mathbf{E}\{\xi|\mathcal{F}_{\eta_1}\}, \text{ since } \mathcal{F}_{\eta_1} \subset \mathcal{F}_{(\eta_1,\eta_2)}$$
see $\mathcal{F}_{(\eta_1,\eta_2)} = \sigma\{A_1 \cap A_2 :\ A_1 = \eta_1^{-1}(C_1), A_2 = \eta_2^{-1}(C_2);\ C_1 \in \mathcal{C}_1,\ C_2 \in \mathcal{C}_2\}$
$$\mathcal{F}_{\eta_1} = \sigma\{A_1 :\ A_1 = \eta_1^{-1}(C_1),\ C_1 \in \mathcal{C}_1\} \quad \Rightarrow \quad A \in \mathcal{F}_{\eta_1} \ \Rightarrow \ A \in \mathcal{F}_{(\eta_1,\eta_2)}$$

1.3.3 Regular probabilites

Let $\{\Omega, \mathcal{A}, P\}$ be a probability space and $\mathcal{B} \subset \mathcal{A}$ a σ-algebra;
$h(\omega, A) := \mathbf{P}\{A|\mathcal{B}\}$ is a function of $\omega \in \Omega$ and $A \in \mathcal{A}$ defined for each fixed A only up to an event of probability 0;

Definition 1.30 *If for $h(\omega, A) = \mathbf{P}\{A|\mathcal{B}\}$ there exists a function $p(\omega, A)$ such that*

- *for fixed ω the function $p(\omega, A)$ is a probability measure on $\mathcal{A}$*
- *$h(\omega, A) = p(\omega, A)$ almost surely for an arbitrary fixed $A \in \mathcal{A}$*

then the **family of conditional probablites** $\mathbf{P}\{A|\mathcal{B}\}$ *is called* **regular**.
In this case $\mathbf{P}\{A|\mathcal{B}\}$ is identified with $p(\omega, A)$.

Theorem 1.31 *If $\mathbf{P}\{A|\mathcal{B}\}$ is a regular conditional probability and $\xi(\omega)$ a random variable then it holds*

$$\mathbf{E}\{\xi|\mathcal{B}\} = \int_\Omega \xi(\omega)\mathbf{P}\{d\omega|\mathcal{B}\} \quad (mod\ \mathbf{P}) \qquad (*)$$

(conditional expectation can be expressed by means of integrals w.r.t. conditional probability measure).

Proof: Let $A \in \mathcal{A}, \quad \xi = I_A$:

$$\mathbf{E}\{I_A|\mathcal{B}\} = \mathbf{P}\{A|\mathcal{B}\}, \qquad \int_\Omega I_A(\omega)\mathbf{P}(d\omega|\mathcal{B}) = \int_A \mathbf{P}(d\omega|\mathcal{B}) = \mathbf{P}\{A|\mathcal{B}\}$$

$\Rightarrow \quad (*)$ is valid for the indicator of random events I_A,

$\Rightarrow \quad$ linearity of $\mathbf{E}$ and integral implies $(*)$ is valid for discrete r.v.
let $\xi \geq 0$, Theorem 1.4 $\Rightarrow \ \exists$ monotonically increasing sequence $(\xi_n)_{n\geq 1}$ of discrete r.v.'s with countable many values such that $\lim\limits_{n\to\infty} \xi_n = \xi$ uniformly $\forall\, \omega \in \Omega$

Theorem 1.28, 5, and Lebesgue's Theorem on monotone convergence imply

$$\mathbf{E}\{\xi|\mathcal{B}\} = \lim_{n\to\infty} \mathbf{E}\{\xi_n|\mathcal{B}\} = \lim_{n\to\infty} \int_\Omega \xi_n(\omega)\mathbf{P}(d\omega|\mathcal{B}) = \int_\Omega \xi\ \mathbf{P}(d\omega|\mathcal{B}),$$

i.e. $(*)$ is valid for all $\xi \geq 0$

$\xi = \xi^+ - \xi^-$, linearity $\Rightarrow$ $(*)$ is valid $\forall$ r.v. ξ $\Rightarrow$ assertion

Remark 1.15 It is possiblie to emphasize that $\mathbf{P}\{A|\mathcal{B}\}$ is a function of elementary events ω by:

$$\mathbf{P}\{A|\mathcal{B}\} = \mathbf{P}_{\mathcal{B}}(\omega, A) \qquad \text{or} \qquad = \mathbf{P}(\omega, A), \quad \text{if} \quad \mathcal{B} \quad \text{is fixed.}$$

analogously: $\mathbf{P}\{A|\xi\} = \mathbf{P}_{\xi}(\omega, A)$

Definition 1.31 *Let $\{\mathcal{S}, \mathcal{C}\}$ be a measurable space, $\eta : \{\Omega, \mathcal{A}, \mathbf{P}\} \rightarrow \{\mathcal{S}, \mathcal{C}\}$ be a random element and $\mathcal{B}$ be a σ-algebra with $\mathcal{B} \subset \mathcal{A}$.*
The function $Q(\omega, C)$ defined on $\Omega \times \mathcal{C}$ is called the **regular conditional distribution** *of the r.e. η given a σ-algebra $\mathcal{B}$, if*

1. *for a fixed $C \in \mathcal{C}$ $\quad Q(\omega, C)$ is $\mathcal{B}$ - measurable*

2. *for a fixed ω $\quad Q(\omega, C)$ is a probability measure on $\mathcal{C}$ with probability 1*

3. *for each $C \in \mathcal{C}$ it holds $\quad Q(\omega, C) = \mathbf{P}\{\eta \in C\}|\mathcal{B}\}$ $(mod\, \mathbf{P})$*

Remark 1.16 Condition 3 is equivalent to

$$\int_B Q(\omega, C) d\mathbf{P} = \mathbf{P}(\{\eta \in C\} \cap B) \quad \forall \;\; B \in \mathcal{B}.$$

see: $\displaystyle\int_B Q(\omega, C) d\mathbf{P} = \int_B \mathbf{P}(\{\eta \in C\}|\mathcal{B}) d\mathbf{P} = \int_B I_{\{\eta \in C\}} d\mathbf{P} = \mathbf{P}(\{\eta \in C\} \cap B)$

Theorem 1.32 *Let*

$\mathcal{S}$:	*be a complete, separable, metric space;*	*Then there exists a regular*
$\mathcal{C}$:	*the σ-algebra of Borel sets of $\mathcal{S}$;*	
η :	$\{\Omega, \mathcal{A}, P\} \rightarrow \{\mathcal{S}, \mathcal{C}\}$ *a random element*	
$\mathcal{B} \subset \mathcal{A}$:	*an arbitrary σ - algebra.*	

conditional distribution of the random element η given the σ-algebra $\mathcal{B}$.

Proof: not given

1.3.4 Conditional densites

Definition 1.32 *Let $\{\mathcal{S}, \mathcal{C}, m\}$ be a measure space and $\eta(\omega)$ be a random element in $\{\mathcal{S}, \mathcal{C}\}$ i.e. $\eta : \{\Omega, \mathcal{A}\} \rightarrow \{\mathcal{S}, \mathcal{C}\}$;*
η possesses a **distribution density** *$\rho(x)$ w.r.t. the measure m if it holds*

$$\mathbf{P}(\eta \in C) = \int_C \rho(x) dm = \int_C \rho(x) m(dx), \quad \forall \; C \in \mathcal{C}.$$

Corollary 1.3 *Let $f: \{\mathcal{S}, \mathcal{C}\} \to \{R, \mathcal{B}\}$ be a measurable mapping, then*

$$\mathbf{E}\, f(\eta) = \int\limits_{\Omega} f(\eta(\omega) d\mathbf{P} = \int\limits_{\mathcal{S}} f(x)\rho(x) dm$$

Theorem 1.33 *The random element η possesses a distribution density if and only if the measure $\mathbf{P}_\eta$ defined by $\mathbf{P}_\eta(C) = \mathbf{P}(\eta^{-1}(C))$ is absolutely continuous w.r.t. the measure m.*

Remark 1.17 *μ is absolutely continuous w.r.t. m if*
$m(C) = 0$ implies $\mu(C) = 0$, $\quad C \in \mathcal{C}$
or if $\forall\, \varepsilon > 0 \quad \exists\, \delta > 0$ such that $m(C) \leq \delta \quad \Rightarrow \quad |\mu(C)| \leq \varepsilon$

Proof: apply Radon-Nikodym Theorem for $\mu(C) = \mathbf{P}_\eta(C)$ and $m(C)$

1. let $\mu = \mathbf{P}_\eta$ be absolutely continuous *w.r.t.* m : $\quad \mu \geq 0, \quad \mu(\mathcal{S}) = 1$ (μ is finite)

   $$\Rightarrow \quad \exists \;\; \rho(x) \geq 0 \;\; \text{with} \;\; \mathbf{P}_\eta(C) = \int\limits_C \rho(x) dm$$

2. let η possess a distribution density, i.e. $\mathbf{P}(\eta \in C) = \int\limits_C \rho(x) dm$:

   $\Rightarrow \; m(C) = 0$ implies $\mathbf{P}_\eta(C) = 0 \;\Rightarrow\; \mathbf{P}_\eta$ is absolutely continuous w.r.t. m

$\Rightarrow$ assertion

consider the random elements $\eta_i, \; i = 1, 2$

$$\eta_i : \quad \{\Omega, \mathcal{A}, P\} \overset{\text{measurable}}{\longrightarrow} \{\mathcal{S}_i, \mathcal{C}_i, m_i\},$$

$$\Rightarrow \quad (\eta_1, \eta_2) : \quad \{\Omega, \mathcal{A}, P\} \overset{\text{measurable}}{\longrightarrow} \{\mathcal{S}_1 \times \mathcal{S}_2, \sigma\{\mathcal{C}_1 \times \mathcal{C}_2\}, m_1 \times m_2\}$$

assume (η_1, η_2) possesses a distribution density $\rho(x_1, x_2)$ w.r.t. $m_1 \times m_2$, i.e.

$$\mathbf{P}(\eta_1 \in C_1, \eta_2 \in C_2) = \int\limits_{C_1} \int\limits_{C_2} \rho(x_1, x_2) m_1(dx_1) m_2(dx_2),$$

ρ is called the **joint distribution density function**

Theorem 1.34 *The existence of the joint distribution density function $\rho(x_1, x_2)$ of (η_1, η_2) implies the existence of the distribution densities of the r.e.'s η_1 and η_2. It holds*

$$\rho_{\eta_1}(x_1) = \int\limits_{\mathcal{S}_2} \rho(x_1, x_2) m_2(dx_2), \qquad \rho_{\eta_2}(x_2) = \int\limits_{\mathcal{S}_1} \rho(x_1, x_2) m_1(dx_1).$$

Proof: $\mathbf{P}(\eta_1 \in C_1) = \mathbf{P}(\eta_1 \in C_1, \eta_2 \in \mathcal{S}_2) = \int\limits_{C_1} \int\limits_{\mathcal{S}_2} \rho(x_1, x_2) m_2(dx_2) m_1(dx_1)$

$$= \int\limits_{C_1} \rho_{\eta_1}(x_1) m_1(dx_1) \qquad \Rightarrow \quad \text{assertion}$$

Theorem 1.35 *Let $f: \{\mathcal{S}_2, \mathcal{C}_2\} \to \{\mathbb{R}, \mathcal{B}\}$ ($\mathcal{B}$: Borel sets) be a measurable function. If (η_1, η_2) possesses a joint distribution density $\rho(x_1, x_2)$, then it holds*

$$\mathbf{E}\{f(\eta_2) | \eta_1\} = \int\limits_{\mathcal{S}_2} f(x_2) \tfrac{\rho(\eta_1, x_2)}{\rho_{\eta_1}(\eta_1)} m_2(dx_2)$$

Here, $\rho(x_2|x_1) := \frac{\rho(x_1, x_2)}{\rho_{\eta_1}(x_1)}$ is called the **conditional distribution density** *of the random element η_2 given $\eta_1 = x_1$.*

Proof: $\forall\ C_1 \in \mathcal{C}_1$ it holds:

$$
\begin{aligned}
\int\limits_{\eta_1^{-1}(C_1)} \mathbf{E}\{f(\eta_2)|\eta_1\}d\mathbf{P} &= \int\limits_{\eta_1^{-1}(C_1)} f(\eta_2)d\mathbf{P} \quad = \quad \mathbf{E}\left(f(\eta_2)I_{C_1}(\eta_1)\right) \\
&\qquad (\text{see } \eta_1^{-1}(C_1) = \{\omega :\ \eta_1(\omega) \in C_1\}) \\
&= \int\limits_{\mathcal{S}_1}\int\limits_{\mathcal{S}_2} f(x_2)I_{C_1}(x_1)\rho(x_1,x_2)\, m_1(dx_1)m_2(dx_2) \\
&= \int\limits_{C_1} \underbrace{\left(\int\limits_{\mathcal{S}_2} f(x_2)\frac{\rho(x_1,x_2)}{\rho_{\eta_1}(x_1)}\, m_2(dx_2)\right)}_{\bar{f}(x_1)} \rho_{\eta_1}(x_1)m_1(dx_1) \\
&= \int\limits_{C_1} \bar{f}(x_1)\rho_{\eta_1}(x_1)m_1(dx_1) = \int\limits_{\eta_1^{-1}(C_1)} \bar{f}(\eta_1)d\mathbf{P} \\
\Rightarrow \quad \mathbf{E}\{f(\eta_2)|\eta_1\} &= \bar{f}(\eta_1) = \int\limits_{\mathcal{S}_2} f(x_2)\frac{\rho(\eta_1,x_2)}{\rho_{\eta_1}(\eta_1)}\, m_2(\mathrm{d}x_2) \qquad \Rightarrow \quad \text{assertion}
\end{aligned}
$$

Remark 1.18 $\mathbf{E}\{f(\eta_2)|\eta_1\}$ is a r.v. with values

$$
\mathbf{E}\{f(\eta_2)|\eta_1 = x_1\} = \int\limits_{\mathcal{S}_2} f(x_2)\frac{\rho(x_1,x_2)}{\rho_{\eta_1}(x_1)} m_2(dx_2).
$$

(see beginning of subsection 1.3.1)

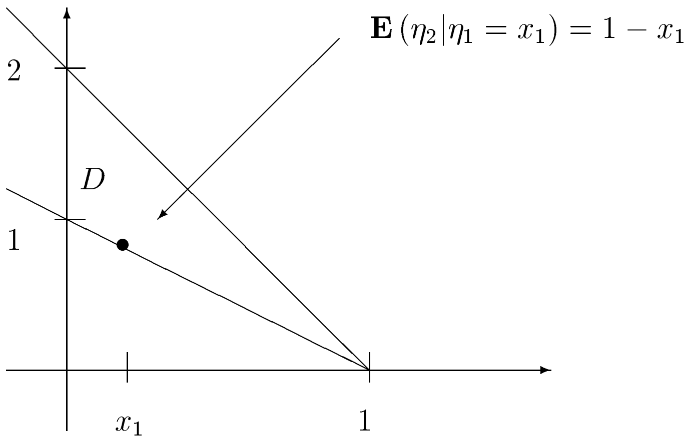
Example 1.9 Let (η_1,η_2) be a random vector which is uniformly distributed on the triangle D with vertices (0,0), (1,0), (0,2))

$\mathcal{S}_1 = [0,1], \quad \mathcal{S}_2 = [0,2]; \quad m_1, m_2$ – Lebesgue measure

$\Rightarrow \quad x := (x_1,x_2); \quad x \in D: \quad \rho(x) = 1, \qquad x \notin D: \quad \rho(x) = 0;$

$f(y) = y: \qquad x_1 \in [0,1] \quad \Rightarrow$

$$
\begin{aligned}
\mathbf{E}\{\eta_2|\eta_1 = x_1\} &= \int\limits_0^2 x_2 \frac{\rho(x_1,x_2)}{\rho_{\eta_1}(x_1)} dx_2 \\
&= \int\limits_0^{2-2x_1} x_2\ \frac{1}{\int\limits_0^{2-2x_1} 1dx_2} dx_2 = \int\limits_0^{2-2x_1} x_2\ \frac{1}{2(1-x_1)} dx_2 \\
&= \frac{1}{2(1-x_1)} \frac{x_2^2}{2}\Bigg|_0^{2-2x_1} = 1 - x_1
\end{aligned}
$$


$\mathbf{E}(\eta_2|\eta_1 = x_1) = 1 - x_1$
2
D
1
x_1
1

2 Random functions and random mappings

Let $\{\Omega, \mathcal{A}, \mathbf{P}\}$ be a given probability space and let the realization of a random experiment be described by a function $\xi(x), \quad x \in \mathcal{X}$
$\Rightarrow$ we say: on $\{\Omega, \mathcal{A}, \mathbf{P}\}$ is given a random function
i.e. random function is a mapping $\omega \to \xi(x) = \xi(x, \omega), \quad \omega \in \Omega$
require: $\xi(x, \omega)$ is a random variable (or element) for any fixed $x \in \mathcal{X}$

Definition 2.1 *Let $\mathcal{X}$ be some set and $\{\mathcal{S}, \mathcal{C}\}$ be a measurable space. A* **random mapping** $\xi(x, \omega)$ *of the set $\mathcal{X}$ into a measurable space $\{\mathcal{S}, \mathcal{C}\}$ is called the mapping $\mathcal{X} \times \Omega \to \mathcal{S}$, which is for arbitrary fixed x a measurable mapping of $\{\Omega, \mathcal{A}\}$ into $\{\mathcal{S}, \mathcal{C}\}$, i.e. $\{\omega : \ \xi(x, \omega) \in C\} \in \mathcal{A} \quad \forall \ C \in \mathcal{C}$.*

"random mapping" $\longleftrightarrow$ "random function with values in $\mathcal{S}$"
$x \in \mathcal{X}$: argument of the random function

- $\mathcal{X} = (-\infty, \infty)$ or $\mathcal{X} = (T_1, T_2) \subset \mathbb{R} \quad \Rightarrow \quad x$ is interpreted as time (instead of x we use t)
  random function is called **random process**
- $\mathcal{X} = \{0, 1, 2, \ldots\}$ or $\mathcal{X} = \{\ldots, -n, \ldots, -1, 0, 1, \ldots, n, \ldots\}$
  random function is called **discrete time random process** or **time series**
- $\mathcal{X} = G \subset \mathbb{R}^m$: random function is called **random field**

Definition 2.2 *Let $\xi(x), \ x \in \mathcal{X}$ be a random function with values in $\{\mathcal{S}, \mathcal{C}\}$, i.e. $\xi(x, .) : \{\Omega, \mathcal{A}, \mathbf{P}\} \overset{measurable}{\longrightarrow} \{\mathcal{S}, \mathcal{C}\}$. Further, let $n \in \mathbb{N}$ and $x_1, \ldots, x_n \in \mathcal{X}$. Consider the random elements $(\xi(x_1), \ldots, \xi(x_n)) : \{\Omega, \mathcal{A}, \mathbf{P}\} \to \{\mathcal{S}^n, \mathcal{C}^n\}$ which induce measures on $\mathcal{C}^n$ defined by*

$$\mathbf{P}_{x_1,\ldots,x_n}(C^n) := \mathbf{P}(\omega : \ (\xi(x_1, \omega), \ldots, \xi(x_n, \omega)) \in C^n); \qquad C^n \in \mathcal{C}^n.$$

The above meaures are called **finite dimensional distributions (f.d.d.'s)** *or* **marginal distributions** *of the random function $\xi(x)$.*

Theorem 2.1 *The family of finite dimensional distributions of a random function $\xi(x)$ possesses the following* **conditions of compatibility** *or* **consistence***:*

1. $\mathbf{P}_{x_1 \ldots x_n \, x_{n+1} \ldots x_{n+m}}(C^n \times \mathcal{S}^m) = \mathbf{P}_{x_1 \ldots x_n}(C^n) \quad \forall \ C^n \in \mathcal{C}^n$
2. *let* π : *permutation of* $(1, 2, \ldots, n)$,
   $\pi(x_1, \ldots, x_n) := (x_{\pi(1)}, \ldots, x_{\pi(n)})$;
   Π : *corresponding permutation of sets in* $\mathcal{S}^n$, *i.e.*
   for any any subset C^n of points $(z_1, \ldots, z_n) \in \mathcal{S}^n$ with $C^n \in \mathcal{C}^n$ it holds $\Pi(C^n) = \{(z_{\pi(1)}, \ldots z_{\pi(n)}) \in \mathcal{S}^n : (z_1, \ldots, z_n) \in C^n)\}$

   then $\mathbf{P}_{\pi(x_1,\ldots,x_n)}(\Pi C^n) = \mathbf{P}_{x_1,\ldots,x_n}(C^n) \qquad \forall \ C^n \in \mathcal{C}^n; \quad \forall \ n \in N$

Proof: obvious

From the practical point of view two random functions are different only if the corresponding f.d.d's are different.

Definition 2.3 *Two random functions* $\xi(x)$ *and* $\xi'(x)$, $x \in \mathcal{X}$ *with values in* $\mathcal{S}$ *defined on possibly different proabability spaces, i.e.*

$$\xi(x): \ \{\Omega, \mathcal{A}, \mathbf{P}\} \to \{\mathcal{S}, \mathcal{C}\} \qquad \xi'(x): \ \{\Omega', \mathcal{A}', \mathbf{P}'\} \to \{\mathcal{S}, \mathcal{C}\}$$

are called **stochastically equivalent in the wide sense**, *if* $\forall\ n \in \mathbb{N}, \forall\ x_k \in \mathcal{X},\ k = 1, \ldots, n$ *and* $\forall\ C^n \in \mathcal{C}^n$ *their f.d.d's coincide:*

$$\mathbf{P}(\omega: \ (\xi(x_1), \ldots, \xi(x_n)) \in C^n) = \mathbf{P}'(\omega': \ (\xi'(x_1), \ldots, \xi'(x_n)) \in C^n).$$

Definition 2.4 *Two random functions* $\xi_1(x)$ *and* $\xi_2(x)$, $x \in \mathcal{X}$ *defined on the same probability space* $\{\Omega, \mathcal{A}, \mathbf{P}\}$ *are called* **stochastically equivalent**, *if*

$$\mathbf{P}(\omega: \ \xi_1(x, \omega) \neq \xi_2(x, \omega)) = 0 \qquad \forall \ \ x \in \mathcal{X}$$

and are said to be **modifications** *or* **versions** *of each other.*
They are called **indistinguishable** *if*

$$\mathbf{P}(\omega: \ \xi_1(x, \omega) = \xi_2(x, \omega), \quad \forall \ \ x \in \mathcal{X}) = 1.$$

Theorem 2.2 *If* $\xi_1(x)$ *and* $\xi_2(x)$ *are stochastically equivalent, then they are stochastically equivalent in the wide sense.*

Proof: For any set $N \subset \Omega$ with

$$N = \{\omega : (\xi_1(x_1), \ldots, \xi_1(x_n)) \neq (\xi_2(x_1), \ldots, \xi_2(x_n))\} = \bigcup_{i=1}^{n} \{\omega : \xi_1(x_i) \neq \xi_2(x_i)\}$$

it holds $\mathbf{P}(N) = 0$ (n is finite) $\Rightarrow$

$$\mathbf{P}(\omega: \ (\xi_1(x_1), \ldots, \xi_1(x_n)) \in C^n) = \mathbf{P}(\omega: \ (\xi_2(x_1), \ldots, \xi_2(x_n)) \in C^n)$$

Example 2.1 $\Omega = [0,1]$; $\mathcal{A}$: Borel sets of $[0,1]$; $\mathbf{P}$: Lebesgue measure;
$\mathcal{S} = \mathrm{R}$; $\mathcal{C}$: Borel sets of R; $\mathcal{X} = [0,1]$;

$$\xi_1(x, \omega) = 0 \quad \forall \ \ (x, \omega) \in \mathcal{X} \times \Omega = [0,1]^2$$

$$\xi_2(x, \omega) = \begin{cases} 1 & (x, \omega) \in D = \{(x, \omega): \ x = \omega\} \quad (D \text{ diagonal}) \\ 0 & (x, \omega) \in [0,1]^2 \setminus D \end{cases}$$

paths of ξ_1 are continuous while the paths of ξ_2 are not continuous
$\forall \ \ x \in \mathcal{X}: \quad \mathbf{P}(\omega: \ \xi_1(x, \omega) \neq \xi_2(x, \omega)) = \mathbf{P}(\omega: \ \omega = x) = 0$
(Lebesgue measure of a single point is 0)
$\Rightarrow$ ξ_1, ξ_2 are stochastically equivalent
above theorem $\Rightarrow$ ξ_1, ξ_2 are stochastically equivalent in the wide sense
but:

- paths do not coincide
- consider: $A_i = \{\omega: \ \xi_i(x, \omega) \leq \frac{1}{2}, \ \ \forall\ x \in \mathcal{X}\}, \quad i = 1, 2$

  $$\mathbf{P}(A_1) = \mathbf{P}(\Omega) = 1 \text{ but } \mathbf{P}(A_2) = \mathbf{P}(\emptyset) = 0$$

  in this special case it holds $A_1, A_2 \in \mathcal{A}$,
  in the general case $A = \{\omega: \ \xi(x, \omega) \leq c, \ \ \forall\ x \in \mathcal{X}\}$ needs not to be contained in $\mathcal{A}$, since

  $$A = \bigcap_{x \in \mathcal{X}} \{\omega: \ \ \xi(x, \omega) \leq c\} \quad (\text{ no countable intersection!})$$

Definition 2.5 *Let $\{\mathcal{S}, \mathcal{C}\}$ be a measurable space. A family of finite dimensional distributions*

$$\{\mathbf{P}_{x_1 \dots x_n}(C^n), \quad n = 1, 2, \dots; \quad x_k \in \mathcal{X}; \quad C^n \in \mathcal{C}^n\}$$

satisfying the conditions of compatibility is called a **random function in the wide sense** *with values in $\{\mathcal{S}, \mathcal{C}\}$.*

usually random functions are only given in the wide sense, since

1. in many practical problems random functions are characterized by their f.d.d.'s while the underlying probability space is not given
2. it is simpler to define the f.d.d's than the corresponding probability space and the mapping $\xi(x, \omega)$;
3. for solving many practical problems it is sufficient to know only the f.d.d's

Construction of a random function by its f.d.d.'s

Given the f.d.d's of two random functions it is possible to decide whether they are stochastically equivalent in the wide sense or not.
It is possible to change $\{\Omega, \mathcal{A}, \mathbf{P}\}$ and $\xi(x, \omega)$ (arbitrarily) as long the f.d.d's remain unchanged.
Problem : Does there exist a random function with a family of f.d.d's which coincides with a given family

$$\{\mathbf{P}_{x_1 \dots x_n}(C^n); \quad n \in \mathbb{N}; \quad x_k \in \mathcal{X}; \quad C^n \in \mathcal{C}^n\}? \tag{1}$$

(the answer will be: yes)

Definition 2.6 *If there exists a probability space $\{\Omega, \mathcal{A}, \mathbf{P}\}$ and a random function $\xi(x, \omega)$ with values in $\mathcal{S}, \mathcal{C}$ such that the f.d.d.'s of ξ coincide with the family (1), i.e.*

$$\mathbf{P}(\omega : \ (\xi(x_1, \omega), \dots, \xi(x_n, \omega)) \in C^n) = \mathbf{P}_{x_1 \dots x_n}(C^n), \quad \forall C^n \in \mathcal{C}^n$$

then $\{\Omega, \mathcal{A}, \mathbf{P}\}$ and $\xi(x, \omega)$ are called **representation of the family of the distributions** *(1).*

Theorem 2.3 Kolmogorov
Let $\mathcal{S}$ be a complete, metric and separable space. The family of distributions

$$\{\mathbf{P}_{x_1 \dots x_n}(C^n), \quad n = 1, 2, \dots, \quad x_k \in \mathcal{X}, \quad C^n \in \mathcal{C}^n\},$$

satisfying the conditions of compatibility (see Theorem 2.1) possesses a representation.

Proof: not given

3 Gaussian random functions

Definition 3.1 *A real-valued random function $\xi(x),\ x \in \mathcal{X}$, i.e.*
$\xi(x):\ \{\Omega, \mathcal{A}, P\} \overset{measurable}{\longrightarrow} \{\mathbb{R}, \mathcal{B}\}, \forall\ x \in \mathcal{X}$, *is called* **Gaussian**, *if for any $n \in \mathbb{N}$ and any $x_k \in \mathcal{X},\ k = 1, 2, \ldots, n$, the random vector $(\xi(x_1), \ldots, \xi(x_n))^\tau$ has a joint normal distribution, i.e. for its characteristic function it holds*

$$J(x_1, \ldots, x_n; u_1, \ldots, u_n) = \exp\left(iu^\tau a - \frac{1}{2}u^\tau B u\right),$$

$$\begin{aligned} \textit{where}\quad u &= (u_1, \ldots, u_n)^\tau \\ a &= (a(x_1), \ldots, a(x_n))^\tau \quad \textit{with}\ \ a(x) = \mathbf{E}\,\xi(x), \\ B &= (b(x_i, x_j))_{1\le i,j\le n} \qquad \textit{with}\ \ b(x,y) = \mathbf{E}\,(\xi(x) - a(x))(\xi(y) - a(y)). \end{aligned}$$

The function $a(x)$ is called **mean value function** *and $b(x, y)$* **correlation function**.

Theorem 3.1 *The correlation function $b(x, y)$ possesses the following properties*

1. *$b(x, y) = b(y, x)$,*
2. *$b(x, y)$ is a positive semi-definite function:*
   $\sum\limits_{i,j=1}^{n} b(x_i, x_j)u_i u_j \ge 0 \quad \forall\ \ n \in \mathbb{N}, \quad \forall\ u_i \in \mathbb{R}, \quad \forall\ x_i \in \mathcal{X};$

i.e. the matrix $B = (b(x_i, x_j))_{1\le i,j\le n}$ is symmetric and positive semi-definite.
Proof:
1. obvious
2. $0 \le \mathbf{E}\,\Big(\sum\limits_{i=1}^{n}(\xi(x_i) - a(x_i))u_i\Big)^2 = \sum\limits_{i,j=1}^{n} \mathbf{E}\,(\xi(x_i) - a(x_i))(\xi(x_j) - a(x_j))u_i u_j$

   $= \sum\limits_{i,j=1}^{n} b(x_i, x_j)u_i u_j \qquad \Rightarrow$ assertion

Theorem 3.2 *Let $\xi(x)$ be a Gaussian random function with a (strictly) positve definite correlation function b, then for the distribution of $(\xi(x_1), \ldots, \xi(x_n))^\tau$ it holds for $C^n \in \mathcal{B}^n$*

$$\mathbf{P}((\xi(x_1), \ldots, \xi(x_n)) \in C^n) = \int_{C^n}\!\!\cdots\!\int f(y)dy,$$

with the distribution density

$$f(y) = \frac{1}{\sqrt{(2\pi)^n \det B}} \exp\left(-\frac{1}{2}(y-a)^\tau B^{-1}(y-a))\right)$$

with a and B from Defintion 3.1.

Proof: b is strictly positive definite $\Rightarrow$ B is symmetric and positive definite $\Rightarrow$ $\exists$ diagonal matrix $\Lambda =$ diag$(\lambda_1, \ldots, \lambda_n)$ with $\lambda_i > 0$ (eigenvalues) and an orthogonal matrix Q such that $B = Q\Lambda Q^\tau = \underbrace{Q\Lambda^{\frac{1}{2}}}_{L}\underbrace{\Lambda^{\frac{1}{2}}Q^\tau}_{L^\tau}$

$(Q^{-1} = Q^\tau;\quad \det Q = 1;\quad$ eigenvectors of B are the columns of $Q) \Rightarrow$

- $B^{-1} = (LL^\tau)^{-1} = (L^\tau)^{-1}L^{-1} = L^{-\tau}L^{-1}$
- $\det B = \det Q \cdot \det \Lambda \cdot \det Q = \det \Lambda = \lambda_1 \cdot \ldots \cdot \lambda_n$
- $\det L = \underbrace{\det Q}_{=1} \det \Lambda^{\frac{1}{2}} = \sqrt{\lambda_1} \cdot \ldots \cdot \sqrt{\lambda_n} = \sqrt{\det \Lambda} = \sqrt{\det B}$ $\quad (*)$

evaluation of the distribution density:
J is absolutely integrable $\Rightarrow$ it exists a distribution density (see basic course, Theorem 6.4) which is given by

$$
\begin{aligned}
f(y) &= \frac{1}{(2\pi)^n} \int\limits_{\mathbb{R}^n} \exp\left(-iu^\tau y\right) J(x_1, \ldots, x_n; u_1, \ldots, u_n) du \\
&= \frac{1}{(2\pi)^n} \int\limits_{\mathbb{R}^n} \exp\left(-iu^\tau (y-a) - \frac{1}{2} u^\tau B u\right) du \\
& \qquad\qquad u = L^{-\tau} v,\ du = \det L^{-\tau} dv \underset{(*)}{=} \frac{1}{\sqrt{\det B}} dv, \\
& \qquad\qquad u^\tau B u = v^\tau L^{-1} \underbrace{B}_{LL^\tau} L^{-\tau} v = v^\tau v \\
&= \frac{1}{(2\pi)^n \sqrt{\det B}} \int\limits_{\mathbb{R}^n} \exp\left(-iv^\tau \underbrace{L^{-1}(y-a)}_{c} - \frac{1}{2} v^\tau v\right) dv \qquad (+)
\end{aligned}
$$

Set $c := L^{-1}(y-a)$, then for the exponent it follows

$$
\begin{aligned}
-iv^\tau c - \frac{1}{2} v^\tau v &= -\frac{1}{2}(v^\tau v + \underbrace{iv^\tau c + ic^\tau v}_{=2iv^\tau c} \underbrace{-c^\tau c + c^\tau c}_{=0}) \\
&= -\frac{1}{2}((v+ic)^\tau (v+ic) + c^\tau c) \qquad (++) \\
\text{it holds } \ c^\tau c &= (y-a)^\tau L^{-\tau} L^{-1} (y-a) = (y-a)^\tau B^{-1} (y-a) \\
& \quad \text{is independent of } v \\
\text{set } \quad w = v + ic \ &\Rightarrow \ (v+ic)^\tau (v+ic) = w^\tau w, \quad dw = dv
\end{aligned}
$$

$$
\begin{aligned}
(+), (++) \Rightarrow \ f(y) &= \frac{1}{(2\pi)^n \sqrt{\det B}} \int\limits_{\mathbb{R}^n} \exp\left(-\frac{1}{2}((v+ic)^\tau (v+ic) + c^\tau c)\right) dv \\
&= \frac{1}{(2\pi)^n \sqrt{\det B}} \exp\left(-\frac{1}{2}(y-a)^\tau B^{-1}(y-a)\right) \underbrace{\int\limits_{\mathbb{R}^n} \exp\left(-\frac{1}{2} w^\tau w\right) dw}_{\prod\limits_{k=1}^n \int\limits_{-\infty}^{\infty} e^{-\frac{w_k^2}{2}} dw_k = \sqrt{2\pi}^n} \\
&= \frac{1}{\sqrt{(2\pi)^n \det B}} \exp\left(-\frac{1}{2}(y-a)^\tau B^{-1}(y-a)\right) \quad \Rightarrow \text{ assertion}
\end{aligned}
$$

Remark 3.1 Without the assumption of a strictly positive definite correlation function b the matrix B can be singular, i.e. the density $f(y)$ is not well defined. This case corresponds to a degenerated random vector. Let rank $B = m < n$ then the distribution of $(\xi(x_1, \ldots, \xi(x_n))^\tau$ is concentrated on a m-dimensional submanifold of $\mathbb{R}^n$. A well defined density exists only for projections on this submanifold.
In case of $n = 1$ the matrix B corresponds to the variance of $\xi(x_1)$. If $B = 0$ then $\mathbf{P}(\xi(x_1) = a(x_1)) = 1$.

Theorem 3.3 *Let*

$\mathcal{X}$: *be an arbitrary set,*
$a(x)$ *a real function on* $\mathcal{X}$,
$b(x, y)$: *a positive semi-definite, symmetric, real function on* $\mathcal{X}^2$.

Then there exists a Gaussian random function $\xi(x)$ *for which* $a(x)$ *is its mean value and* $b(x, y)$ *its correlation function.*

Proof: consider the family of distributions from Definition 3.1;
it satisfies the conditions of compatibility;
Theorem of Kolmogoroff (Theorem 2.3) $\Rightarrow$ assertion

Definition 3.2 *The* **random vector function** $\xi(x) = (\xi_1(x), \ldots, \xi_m(x))^\tau \in \mathbb{R}^m$, $m \in \mathbb{N}$, $x \in \mathcal{X}$ *is called* **Gaussian**, *if* $\forall\ n \in \mathbb{N}$ *and* $x_1, \ldots, x_n \in \mathcal{X}$ *the joint distributions of all components of the sequences*

$$(\xi(x_1), \ldots, \xi(x_n)) = (\xi_1(x_1), \ldots, \xi_m(x_1), \ldots, \xi_1(x_n), \ldots, \xi_m(x_n))$$

are normal distributions, i.e. the characteristic function of $(\xi(x_1), \ldots, \xi(x_n))$ *has the form*

$$J(x_1, \ldots, x_n; u_1, \ldots, u_n) = \exp\left(i \sum_{k=1}^{n} u_k^\tau a(x_k) - \frac{1}{2} \sum_{j,k=1}^{n} u_j^\tau b(x_j, x_k) u_k \right)$$

$$\begin{aligned} \textit{where}\quad u_k &= (u_{k1}, \ldots, u_{km})^\tau \in \mathbb{R}^m, \\ a(x) &= \mathbf{E}\,\xi(x) = (a_1(x), \ldots, a_m(x))^\tau \\ b(x, y) &= \mathbf{E}\left((\xi(x) - a(x))(\xi(y) - a(y))^\tau\right) = (b_{pq}(x, y))_{1 \le p,q \le m} \end{aligned}$$

and $\quad u_k^\tau\, a(x) = \sum_{p=1}^{m} u_{kp} a_p(x); \quad u_j^\tau\, b(x, y) u_k = \sum_{p,q=1}^{m} b_{pq}(x, y) u_{jp} u_{kq}.$

Corollary 3.1 $a(x)$ *can be an arbitrary vector function on* $\mathcal{X}$ *with values in* $\mathbb{R}^m$ *and* $b(x, y)$ *is a symmetric and positive semi-definite function on* $\mathcal{X}^2$ *with values in* $\mathbb{R}^{m \times m}$, *i.e.*

$$\sum_{j,k=1}^{n} u_j^\tau\, b(x_j, x_k) u_k \ \ge\ 0, \quad x_. \in \mathcal{X}, \quad u_. \in \mathbb{R}^m, \quad \forall\ \ n \in \mathbb{N}.$$

Proof: evaluate $\mathbf{E}\left(\sum_{p=1}^{m} \sum_{k=1}^{n} (\xi_p(x_k) - a_p(x_k)) u_{kp} \right)^2 \ \ge\ 0$

Theorem 3.4 (counterpart to Theorem 3.3)
Let $\mathcal{X}$: *be an arbitrary set,*
$a(x)$: *a vector function on* $\mathcal{X}$*,*
$b(x,y)$: *a positive semi-definite, symmetric matrix function on* $\mathcal{X}^2$*.*
Then there exists a Gaussian random function $\xi(x)$ *for which* $a(x)$ *is its mean value and* $b(x,y)$ *its correlation function, i.e.*

$$\mathbf{E}\,\xi_p(x) = a_p(x)$$

$$\mathbf{E}\,(\xi_p(x) - a_p(x))(\xi_q(y) - a_q(y)) = b_{pq}(x,y), \quad p,q = 1,\ldots,m.$$

Computation of moments for scalar Gaussian random functions

let $\quad a(x) = 0, \quad u = (u_1,\ldots,u_n)^\tau, \quad B = (b(x_j,x_k))_{1\le j,k\le n}$

$$m_{j_1,\ldots,j_n}(x_1,\ldots,x_n) := \mathbf{E}\left([\xi(x_1)]^{j_1}\cdot\ldots\cdot[\xi(x_n)]^{j_n}\right)$$

$$\Rightarrow \quad \sum_{k=1}^{n} j_k \ := \ \text{order of the moment}$$

Theorem 3.5 *For the moments of a scalar Gaussian random function it holds*

$$m_{j_1,\ldots,j_n}(x_1,\ldots,x_n) = \begin{cases} 0 & \textit{order is odd} \\ \frac{\partial^{2s}}{\partial {u_1}^{j_1}\ldots\partial {u_n}^{j_n}}\left(\frac{1}{2^s\,s!}(Bu,u)^s\right)\Big|_{u=0} & \textit{order is even } (=2s). \end{cases}$$

Proof: $x = (x_1,\ldots,x_n)^\tau, \quad u = (u_1,\ldots,u_n)^\tau, \quad u^\tau Bu = (Bu,u) = \sum_{k,l=1}^{n} b_{kl}u_ku_l;$

$$\overset{\text{Theorem 3.2}}{\Longrightarrow} \quad J(x;u) = \exp\left(-\frac{1}{2}(Bu,u)\right) = \sum_{p=0}^{\infty}\frac{1}{p!}\left(-\frac{1}{2}\right)^p(Bu,u)^p$$

$$\overset{\text{Theorem 1.17}}{\Longrightarrow} \quad m_{j_1\,\ldots\,j_n}(x) = (-i)^q\frac{\partial^q J(x;u)}{\partial {u_1}^{j_1}\ldots\partial {u_n}^{j_n}}\Big|_{u=0} \qquad (q = \sum_{k=1}^{n} j_k)$$

$$\sum_{k=1}^{n} j_k = 2s{+}1: \ m_{j_1\,\ldots\,j_n}(x) = 0 \quad (J(.;u) \text{ is even, } (B(-u),-u)) = (Bu,u))$$

$$\sum_{k=1}^{n} j_k = 2s: \ m_{j_1\,\ldots\,j_n}(x) = (-i)^{2s}\frac{\partial^{2s}}{\partial {u_1}^{j_1}\ldots\partial {u_n}^{j_n}}\left(\frac{1}{s!}\left(-\frac{1}{2}\right)^s(Bu,u)^s\right)\Big|_{u=0}$$

$$= \frac{\partial^{2s}}{\partial {u_1}^{j_1}\ldots\partial {u_n}^{j_n}}\left(\frac{1}{2^s s!}(Bu,u)^s\right)\Big|_{u=0}$$

$$\Rightarrow \quad \text{assertion}$$

special cases:

$2s = 2$:

$$m_{11}(x_1,x_2) = \frac{1}{2}\frac{\partial^2}{\partial u_1\partial u_2}\left(\sum_{j,k=1}^{2} b(x_j,x_k)u_ju_k\right)\Big|_{u=0} = b(x_1,x_2) = \mathbf{E}\,\xi(x_1)\xi(x_2)$$

$$m_2(x) = \frac{1}{2}\frac{\partial^2}{\partial {u_1}^2}(b(x,x)u_1^2)\Big|_{u=0} = b(x,x) = m_{11}(x,x) = \mathbf{E}\,\xi^2(x)$$

$2s = 4$:

$$\begin{aligned} m_{1111}(x_1,x_2,x_3,x_4) &= \frac{\partial^4}{\partial u_1 \ldots \partial u_4}\left(\frac{1}{2^2\cdot 2!}(Bu,u)^2\right)\Bigg|_{u=0} \\ \text{where}\quad (Bu,u)^2 &= \left(\sum_{i,j=1}^{4} b_{ij}u_iu_j\right)^2 = \sum_{i,j,k,l=1}^{4} b_{ij}b_{kl}u_iu_ju_ku_l \Rightarrow \\ m_{1111}(x_1,x_2,x_3,x_4) &= b(x_1,x_2)b(x_3,x_4) + b(x_1,x_3)b(x_2,x_4) + b(x_1,x_4)b(x_2,x_3) \end{aligned}$$

it results: $m_{j_1\,\ldots\,j_n}(x_1,\ldots,x_n) = \sum\prod b(x_p,x_q) \qquad \text{for} \qquad \sum_{k=1}^{n} j_k = 2s$

description:

- write down the sequence: $\underbrace{x_1,\ldots,x_1}_{j_1-\text{times}},\underbrace{x_2,\ldots,x_2}_{j_2-\text{times}},\ldots,\underbrace{x_n,\ldots,x_n}_{j_n-\text{times}}$
- subdivide the sequence into pairs
  (pairs counted only once if they are permutations of one another)
- $\prod$ is taken over all pairs of the partition (s factors)
- $\sum$ is taken over all partitions ($(2s-1)!!$ terms)

example $m_{31}(x_1,x_2)$: $x_1\ x_1\ x_1\ x_2$

$\Rightarrow\ (x_1,x_1),\ (x_1,x_2), \quad (x_1,x_1),\ (x_1,x_2), \quad (x_1,x_2),\ (x_1,x_1)$

$\Rightarrow\ m_{31}(x_1,x_2) = 3b(x_1,x_1)b(x_1,x_2)$

$m_{2s}(x) = (2s-1)!!\,(b(x,x))^s = (2s-1)!!\,\sigma^{2s}(x) \quad \text{for} \quad \sigma^2(x) = b(x,x)$

($(2s-1)!!$ equal terms in the sum)

Complex valued Gaussian random functions

Let $\zeta(x) \in \mathbb{C}$ (complex numbers); $\zeta(x) = \xi(x) + i\eta(x)$; $\xi(x),\ \eta(x) \in \mathbb{R},\ x \in \mathcal{X}$

Definition 3.3 *A random function $\zeta(x),\ x \in \mathcal{X}$ with $\zeta(x) = \xi(x) + i\eta(x)$ is called a* **complex Gaussian random function** *if the real vector function $(\xi(x),\eta(x)),\ x \in \mathcal{X}$ is Gaussian and*

$$\mathbf{E}\big(\zeta(x)-a(x))(\zeta(y)-a(y)\big) = 0 \quad \forall\ x,y \in \mathcal{X} \quad \textit{where} \quad a(x) = \mathbf{E}\,\zeta(x).$$

The function $b(x,y) := \mathbf{E}\big(\zeta(x)-a(x))(\overline{\zeta(y)-a(y)}\big),\ x,y \in \mathcal{X}$ is called **correlation function** *of ζ.*

without loss of generality let $a(x) = 0$ and let

$$\begin{pmatrix} b_{11}(x,y) & b_{12}(x,y) \\ b_{21}(x,y) & b_{22}(x,y) \end{pmatrix} := \begin{pmatrix} \mathbf{E}\,\xi(x)\xi(y) & \mathbf{E}\,\xi(x)\eta(y) \\ \mathbf{E}\,\eta(x)\xi(y) & \mathbf{E}\,\eta(x)\eta(y) \end{pmatrix} = \mathbf{E}\begin{pmatrix} \xi(x) \\ \eta(x) \end{pmatrix}(\xi(y),\eta(y))$$

$$\begin{aligned} \mathbf{E}\,\zeta(x)\zeta(y) = 0 \iff\ & \mathbf{E}\big(\xi(x)\xi(y) - \eta(x)\eta(y)\big) = 0 \quad \text{(real part)} \\ & \mathbf{E}\big(\xi(x)\eta(y) + \eta(x)\xi(y)\big) = 0 \quad \text{(imaginary part)} \\ \iff\ & b_{11}(x,y) = \mathbf{E}\,\xi(x)\xi(y) = \mathbf{E}\,\eta(x)\eta(y) = b_{22}(x,y) \\ & b_{12}(x,y) = \mathbf{E}\,\xi(x)\eta(y) = -\mathbf{E}\,\eta(x)\xi(y) = -b_{21}(x,y) \end{aligned} \tag{2}$$

For the correlation function then it follows:

$$\begin{aligned} b(x,y) = \mathbf{E}\,\zeta(x)\overline{\zeta(y)} &= \mathbf{E}\left((\xi(x)+i\eta(x))(\xi(y)-i\eta(y))\right) \\ &= \mathbf{E}\left(\xi(x)\xi(y)+\eta(x)\eta(y)\right) + i\,\mathbf{E}\left(\xi(y)\eta(x)-\xi(x)\eta(y)\right) \\ &\underset{(2)}{=} 2(b_{11}(x,y) - i\,b_{12}(x,y)) \end{aligned}$$

(2) $\Rightarrow \quad b_{12}(x,x) = \mathbf{E}\,\xi(x)\eta(x) = 0 \quad \Rightarrow \quad \xi(x)$ and $\eta(x)$ are independent (since $\xi(x),\ \eta(x)$ are Gaussian)

Theorem 3.6 *Let $\rho(x),\ \phi(x)$ be real random functions on $\mathcal{X}$. If $\zeta(x) = \rho(x)e^{i\phi(x)}$ is a complex Gaussian random function then*

1. *$\rho(x),\ \phi(x)$ are independent $\forall x \in \mathcal{X}$*
2. *$\rho(x)$ is Rayleigh-distributed, its density is*

$$f_\rho(r) = \frac{r}{\sigma^2(x)} \exp\left(\frac{-r^2}{2\sigma^2(x)}\right), \quad u > 0,\ \sigma^2(x) = \mathbf{D}^2\,\xi(x) = b_{11}(x,x)$$

$\phi(x)$ is uniformly distributed on $(-\pi,\pi)$, i.e. its density is

$$g_\phi(v) = \frac{1}{2\pi}, \quad v \in (-\pi,\pi).$$

Proof: $\zeta(x) = \xi(x) + i\eta(x) \quad \Rightarrow \quad \xi(x) = \rho(x)\cos\phi(x), \quad \eta(x) = \rho(x)\sin\phi(x);$

$\zeta(x)$ is Gaussian $\overset{\text{Definition 3.3}}{\Longrightarrow}$ $(\xi(x),\eta(x))$ is a Gaussian vector function

$\Rightarrow$ density of $(\xi(x),\eta(x))$ is

$$\begin{aligned} f_{\xi\eta}(u,v) &= \frac{1}{2\pi}\frac{1}{\sqrt{b_{11}\,b_{22}}} \exp\left(-\frac{1}{2}\left(\frac{u^2}{b_{11}} + \frac{v^2}{b_{22}}\right)\right) \quad \text{(since } b_{12}(x,x) = 0\text{)} \\ &= \frac{1}{2\pi b_{11}} \exp\left(-\frac{u^2+v^2}{2b_{11}}\right) \quad \text{(since } b_{11} = b_{22}\text{)} \end{aligned}$$

transformation of r.v.'s: density of $(\rho(x),\phi(x))$

$$\begin{aligned} f_{\rho\phi}(r,s) &= f_{\xi\eta}(u(r,s),v(r,s)) \left|\frac{\partial(u,v)}{\partial(r,s)}\right| \qquad (u = r\cos s,\ v = r\sin s) \\ &= \frac{1}{2\pi b_{11}} \exp\left(-\frac{r^2}{2b_{11}}\right) \underbrace{\begin{vmatrix} \cos s & -r\sin s \\ \sin s & r\cos s \end{vmatrix}}_{=r} = \frac{1}{2\pi}\frac{r}{b_{11}} \exp\left(-\frac{r^2}{2b_{11}}\right) \end{aligned}$$

$\Rightarrow$ marginal densities of ρ and ϕ:

$$\begin{aligned} f_\rho(r) &= \int_{-\pi}^{\pi} f_{\rho\phi}(r,s)ds = \frac{r}{\sigma^2(x)} \exp\left(-\frac{r^2}{2\sigma^2(x)}\right), \ r > 0,\ b_{11}(x,x) = \sigma^2(x) \\ f_\phi(s) &= \int_0^\infty f_{\rho\phi}(r,s)dr = \frac{1}{2\pi}, \quad s \in (-\pi,\pi) \end{aligned}$$

$\Rightarrow f_{\rho\phi}(r,s) = f_\rho(r)f_\phi(s) \quad \Rightarrow$ assertion

Theorem 3.7 properties of b_{11}, b_{12}, $b = 2(b_{11} - ib_{12})$:

1. $b_{11}(x,y)$ *is a positive semi-definite function*
2. $b_{12}(x,y) = -b_{12}(y,x)$ *(anti-symmetry)*
3. $b(x,y)$ *is a positive semi-definite function, i.e.*
$$\sum_{k,l=1}^{n} b(x_k, x_l) z_k \bar{z}_l \ \geq \ 0 \qquad \forall\, n \in \mathrm{N}, \quad \forall\, x_{\cdot} \in \mathcal{X}, \quad \forall\, z_{\cdot} \in \mathbb{C}\,.$$

Proof:

1. $b_{11}(x,y) = \mathbf{E}\,\xi(x)\xi(y) \quad \overset{\text{Theorem 3.1 2.}}{\Longrightarrow} \quad$ assertion
2. follows from (2)
3. $0 \leq \mathbf{E}\left|\sum_{k=1}^{n} \zeta(x_k) z_k\right|^2 = \mathbf{E}\sum_{k,l=1}^{n} \zeta(x_k)\overline{\zeta(x_l)} z_k \bar{z}_l = \sum_{k,l=1}^{n} b(x_k, x_l) z_k \bar{z}_l$
$\Rightarrow$ assertion

Theorem 3.8 *Let $b(x,y)$, $x,y \in \mathcal{X}$ be an arbitrary positive semi-definite function. Then there exists a complex Gauusian random function $\zeta(x)$ such that*
$$\mathbf{E}\,\zeta(x) = 0 \qquad \textit{and} \qquad \mathbf{E}\,\zeta(x)\overline{\zeta(y)} = b(x,y).$$
Proof: plan of the proof:

(i): consider the real 2×2 matrix function
$$B(x,y) = \begin{pmatrix} b_{11}(x,y) & b_{12}(x,y) \\ b_{21}(x,y) & b_{22}(x,y) \end{pmatrix} := \frac{1}{2}\begin{pmatrix} \mathrm{Re}\, b(x,y) & -\mathrm{Im}\, b(x,y) \\ \mathrm{Im}\, b(x,y) & \mathrm{Re}\, b(x,y) \end{pmatrix}$$

(ii): to prove: $B(x,y)$ is

(a) symmetric: $B(x,y) = B(y,x)^{\tau}$

(b) positive semi-definite

(iii): Theorem 3.4 $\Rightarrow$ $\exists$ real Gaussian random vector function $(\xi(x), \eta(x))$ such that $B(x,y)$ is its correlation function $(\mathbf{E}\,\xi(x) = \mathbf{E}\,\eta(x) = 0)$

(iv): consider $\zeta(x) = \xi(x) + i\eta(x)$ and prove $\mathbf{E}\,\zeta(x)\zeta(y) = 0$
Definition 3.3 $\Rightarrow$ $\zeta(x)$ is a complex Gaussian random function
remains to prove: $\mathbf{E}\,\zeta(x)\overline{\zeta(y)} = b(x,y)$
$\Rightarrow$ assertion

(ii): $b(x,y)$ is positive semi-definite
$\Rightarrow$ $\forall\, n \in \mathrm{N}$, $\forall\, x_a \in \mathcal{X}$, $\forall\, z_a \in \mathrm{C}$, $a = 1, \ldots, n$ gilt:

(a)
$$0 \leq \sum_{p,q=1}^{n} b(x_p, x_q) z_p \bar{z}_q = \overline{\sum_{p,q=1}^{n} b(x_p, x_q) z_p \overline{z_q}} = \sum_{p,q=1}^{n} \overline{b(x_p, x_q)}\, \overline{z_p} z_q$$
$$= \sum_{p,q=1}^{n} \overline{b(x_q, x_p)} z_p \overline{z_q}$$

$\Rightarrow \quad b(x,y) = \overline{b(y,x)}$ (coefficients of identical polynomials)

$\Rightarrow \quad b_{11}(x,y) = \frac{1}{2}\mathrm{Re}\, b(x,y) = \frac{1}{2}\mathrm{Re}\, b(y,x) = b_{11}(y,x);$
$b_{22}(x,y) = b_{11}(x,y)$

$\Rightarrow \quad b_{12}(x,y) = -\frac{1}{2}\mathrm{Im}\, b(x,y) = \frac{1}{2}\mathrm{Im}\, b(y,x) = b_{21}(y,x)$

$\Rightarrow \quad B(x,y) \ = \ B(y,x)^{\tau}$

$$\text{(b)}\quad 0 \le \sum_{p,q=1}^{n} b(x_p,x_q) z_p \bar{z}_q$$
$$= \sum_{p,q=1}^{n} 2\,(\underbrace{b_{11}(x_p,x_q)}_{\frac{1}{2}\mathrm{Re}\, b(x_p,x_q)} - i\,\underbrace{b_{12}(x_p,x_q)}_{-\frac{1}{2}\mathrm{Im}\, b(x_p,x_q)}) \cdot \underbrace{(u_1^{(p)} + i u_2^{(p)})}_{z_p} \underbrace{(u_1^{(q)} - i u_2^{(q)})}_{\overline{z_q}}$$
$$= 2 \sum_{p,q=1}^{n} \Big(b_{11}(x_p,x_q) u_1^{(p)} u_1^{(q)} + \underbrace{b_{11}(x_p,x_q)}_{=b_{22}(x_p,x_q)} u_2^{(p)} u_2^{(q)}$$
$$\underbrace{-b_{12}(x_p,x_q)}_{=b_{21}(x_p,x_q)} u_1^{(p)} u_2^{(q)} + b_{12}(x_p,x_q) u_1^{(q)} u_2^{(p)} \Big) \quad + \quad i \cdot 0$$

(since it is real)

$$= 2 \sum_{p,=1}^{n} \sum_{k,l=1}^{2} b_{kl}(x_p,x_q) u_l^{(p)} u_k^{(q)} = 2 \sum_{p,q=1}^{n} \big(B(x_p,x_q) z_p, z_q \big)$$

(with the vector $z_p = (u_1^{(p)}, u_2^{(p)})^\tau$)$\Rightarrow$ $B(x,y)$ is positive semi-definite

$$\text{(iv):}\quad \begin{aligned} \mathbf{E}\,\zeta(x)\zeta(y) &= \mathbf{E}\,(\xi(x)+i\eta(x))(\xi(y)+i\eta(y)) \\ &= \underbrace{(b_{11}(x,y) + i^2 b_{22}(x,y))}_{=\,0} + i\,\underbrace{(b_{12}(x,y) + b_{21}(x,y))}_{=\,0} = 0 \\ \mathbf{E}\,\zeta(x)\overline{\zeta(y)} &= \mathbf{E}\,(\xi(x)+i\eta(x))(\xi(y)-i\eta(y)) \\ &= b_{11}(x,y) + b_{22}(x,y) + i(-b_{12}(x,y) + b_{21}(x,y)\,) \\ &= \mathrm{Re}\, b(x,y) + i\,\mathrm{Im}\, b(x,y) = b(x,y) \end{aligned}$$

$\Rightarrow$ assertion

Why do Gaussian random functions play an important role in practical problems ? The answer gives a multivariate generalization of the central limit theorem (theorem on normal correlation).
Under very general conditions the sum of a large number of small (in magnitude) random functions is approximately a Gaussian random function independently of the probabilistic nature of the summands.

Consider a double sequence of random functions:
$\xi_{nk}(x),\quad x \in \mathcal{X},\quad k = 1,2,\ldots,m_n,\quad n = 1,2,\ldots;\quad \lim_{n\to\infty} m_n = \infty$
Set $\eta_n(x) := \sum_{k=1}^{m_n} \xi_{nk}(x)$ and define the "truncated" variables
$\xi^{\varepsilon}_{nk}(x) := I_\varepsilon(\xi_{nk}(x))\,\xi_{nk}(x)$ where for $\varepsilon > 0$ $I_\varepsilon(z) = I_{(-\varepsilon,\varepsilon)}(x) = \begin{cases} 1 & x \in (-\varepsilon,\varepsilon) \\ 0 & \text{else} \end{cases}$
and their moments
$a^{\varepsilon}_{nk}(x) := \mathbf{E}\,\xi^{\varepsilon}_{nk}(x);\qquad b^{\varepsilon}_{nk}(x,y) := \mathbf{E}\,(\xi^{\varepsilon}_{nk}(x) - a^{\varepsilon}_{nk}(x))(\xi^{\varepsilon}_{nk}(y) - a^{\varepsilon}_{nk}(y))$

Theorem 3.9 *Let the random functions $\xi_{n1}(x),\ldots,\xi_{nm_n}(x)$ be mutually independent for each n and satisfy the conditions*

1. *for any $\varepsilon > 0$ it holds:* $\displaystyle\lim_{n\to\infty} \sum_{k=1}^{m_n} \mathbf{P}(|\xi_{nk}(x)| > \varepsilon) = 0$
2. *for some $\varepsilon = \varepsilon_0 = \varepsilon_0(x) > 0$ it holds*
$$\lim_{n\to\infty} \sum_{k=1}^{m_n} a^{\varepsilon_0}_{nk}(x) = a(x), \qquad \lim_{n\to\infty} \sum_{k=1}^{m_n} b^{\varepsilon_0}_{nk}(x,y) = b(x,y),$$

Then for $n \to \infty$ the finite dimensional distributions of the random function $\eta_n(x)$ converge weakly (in the sense of weak convergence of measures) to the corresponding distributions of a Gaussian random function with mean value $a(x)$ and correlation function $b(x,y)$.

Proof: not given

Special case:
let $(\rho_k(x))_{k\geq 1}$ be a sequence of mutually independent random functions with values in $[-1,1]$ possessing identical f.d.d.'s with

$a(x) = \mathbf{E}\,\rho(x) = 0$ and $b(x,y) = \mathbf{E}\,\rho(x)\rho(y)$
$m_n = n$, $\xi_{nk} := \frac{1}{\sqrt{n}}\rho_k \in [-\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}]$ and $\eta_n := \sum\limits_{k=1}^{n} \xi_{nk} = \frac{1}{\sqrt{n}} \sum\limits_{k=1}^{n} \rho_k \in [-\sqrt{n}, \sqrt{n}]$
$\Rightarrow$ for every n it holds $\mathbf{E}\,\eta_n(x) = 0$
For $n > \frac{1}{\varepsilon^2}$ it holds $\sum\limits_{k=1}^{n} \underbrace{\mathbf{P}(|\xi_{nk}(x)| > \varepsilon)}_{=0} = 0 \quad \Rightarrow 1.$
For $\varepsilon_0 > \frac{1}{\sqrt{n}}$ it holds $\xi_{nk}^{\varepsilon_0} \equiv \xi_{nk} \Rightarrow$

$$b_{nk}^{\varepsilon_0}(x,y) = b_{nk}(x,y) = \mathbf{E}\,\xi_{nk}(x)\xi_{nk}(y) = \frac{1}{n}\mathbf{E}\,\rho_k(x)\rho_k(y) = \frac{1}{n}\,b(x,y) \quad \Rightarrow \; 2.$$

The correlation function of η_k is

$$\mathbf{E}\,\eta_n(x)\eta_n(y) = \frac{1}{n}\sum_{k,l=1}^{n} \mathbf{E}\,\rho_k(x)\rho_l(y) = \frac{1}{n}\sum_{k=1}^{n} b(x,y) = b(x,y), \quad \forall n \in \mathbb{N}$$

$\Rightarrow$ limit distribution of $(\eta_n)_{n\geq 1}$ is Gaussain with mean 0 and correlation function $b(x,y)$

4 Processes with independent increments

Let $T \subset \mathbb{R}$ be a finite or infinite interval closed on the left, $a = \min T > -\infty$.

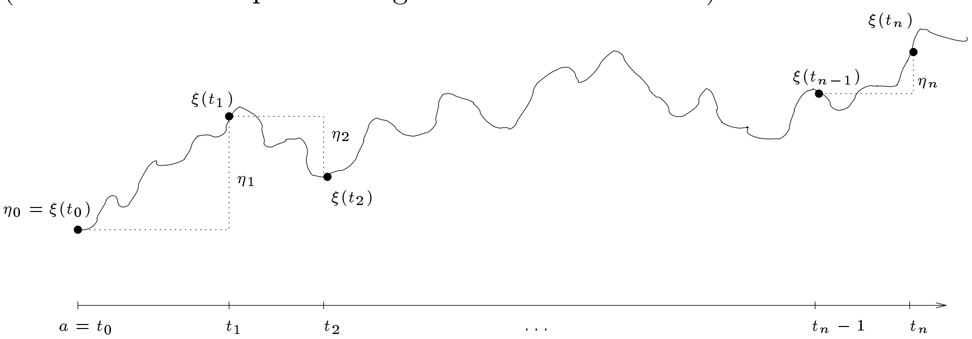
Definition 4.1 *A random process $\xi(t)$, $t \in T$ with values in $\mathbb{R}^m$ is called a* **process with independent increments** *if for any $n \in \mathbb{N}$, $t_k \in T$ with $t_1 < t_2 < \ldots < t_n$, the random vectors $\xi(a)$, $\xi(t_1) - \xi(a)$, $\ldots$, $\xi(t_n) - \xi(t_{n-1})$ are mutually independent.*
The vector $\xi(a)$ is called **initial state** *or* **initial value** *of the process and its distribution the* **initial distribution**.

A process with independent increments is defined in the wide sense if

$$\begin{array}{ll} P_0(B) & \text{the initial distribution and} \\ P(t,h,B) & \text{the distribution of the vector } \xi(t+h) - \xi(t) \\ & \qquad\qquad\text{for } h > 0;\ t \geq a;\ B \in \mathcal{B}^m \text{ (Borel sets of } \mathrm{R}^m) \end{array}$$

are given.

to prove: the finite dimensional distributions are uniquely defined
(Definition 2.5 $\Rightarrow$ process is given in the wide sense)



define:

$$\begin{aligned} \eta_k &:= \xi(t_k) - \xi(t_k - 1),\ k = 1, \ldots, n, \text{ for } n \in \mathbb{N} \quad \text{(increments)} \\ \eta_0 &:= \xi(t_0) \qquad\qquad\qquad\qquad\qquad\qquad \text{(initial value)} \\ Q_{t_0 t_1 \ldots t_n}(B^{(n+1)}) &:= \mathbf{P}(\omega : (\xi(t_0), \xi(t_1), \ldots, \xi(t_n)) \in B^{(n+1)}), \\ & \qquad\qquad B^{(n+1)} \in (\mathcal{B}^m)^{(n+1)}, \quad \text{(f.d.d.'s of } \xi) \\ R_{t_0 t_1 \ldots t_n}(\tilde{B}^{(n+1)}) &:= \mathbf{P}(\omega : (\eta_0, \eta_1, \ldots, \eta_n) \in \tilde{B}^{(n+1)}), \quad \tilde{B}^{(n+1)} \in (\mathcal{B}^m)^{(n+1)} \\ & \qquad\qquad \text{(joint distribution of increments)} \end{aligned}$$

Q and R are measures on $\{(\mathbb{R}^m)^{(n+1)}, (\mathcal{B}^m)^{(n+1)}\}$
determine $Q_{t_0 t_1 \ldots t_n}(B^{(n+1)})$ to given $P_0(B)$, $P(t,h,B)$:

restrict to sets $B^{(n+1)} = B_0 \times B_1 \times \ldots B_n$ and $\tilde{B}^{(n+1)} = \tilde{B}_0 \times \tilde{B}_1 \times \ldots \tilde{B}_n$ with $B_k, \tilde{B}_k \in \mathcal{B}^m$ (rectangles with Borel sides),
independence of increments implies for their joint distribution

$$\begin{aligned} R_{t_0 t_1 \ldots t_n}(\tilde{B}_0 \times \tilde{B}_1 \times \ldots \times \tilde{B}_n) &= \mathbf{P}(\omega : \eta_0 \in \tilde{B}_0, \ldots \eta_n \in \tilde{B}_n) \\ &= \int\limits_{\tilde{B}_0} \int\limits_{\tilde{B}_1} \ldots \int\limits_{\tilde{B}_n} P_0(dy_0) P(t_0, t_1 - t_0, dy_1) \cdot \ldots \cdot P(t_{n-1}, t_n - t_{n-1}, dy_n) \end{aligned}$$

represent f.d.d.'s Q in terms of R:

$$\begin{aligned}
&Q_{t_0t_1\ldots t_n}(B_0\times B_1\times\ldots\times B_n)=\mathbf{P}(\omega:\xi(t_0)\in B_0,\xi(t_1)\in B_1,\ldots,\xi(t_n)\in B_n)\\
&= \mathbf{P}(\eta_0\in B_0,\eta_0+\eta_1\in B_1,\ldots,\eta_0+\ldots+\eta_n\in B_n)\\
&= \mathbf{P}(\eta_0\in B_0,\eta_0\in B_1-\eta_1,\ldots,\eta_n\in B_n-(\eta_0+\ldots+\eta_{n-1}))\\
&\qquad\qquad\text{where } B-z:=\{y:y=x-z,\ x\in B\}\\
&= \int\limits_{B_0} P_0(dy_0)\int\limits_{B_1-y_0} P(t_0,t_1-t_0,dy_1)\ldots\\
&\qquad \ldots\int\limits_{B_n-(y_0+\ldots y_{n-1})} P(t_{n-1},t_n-t_{n-1},dy_n) \qquad (*)
\end{aligned}$$

Question:	Let $P_0(B),\ P(t,h,B)$ be given. Does there exist an random process $\xi(t,\omega)$ with independent increments, possessing these distributions ?
Answer:	Yes, if the f.d.d.'s satisfy the conditions of compatibility.

To fulfill these conditions $P_0(B)$ can be chosen arbitrarily and $P(t,h,B)$ is required to possess the following

Property (+): $\forall\ n,\quad \forall\ a\le t=t_0<t_1<\ldots<t_n=t+h$ *it holds:* *$P(t,h,B)$ is the distribution of sums of independent random vectors: $\xi_1+\ldots+\xi_n$ where $\xi_k:=\xi(t_k)-\xi(t_{k-1})$ is distributed according to $P(t_{k-1},t_k-t_{k-1},B)$.*

Remark 4.1 This property of $P(t,h,B)$ (for fixed t,h) possess the so-called infinitely divisble distributions e.g. normal and Poisson (but not uniform, exponential, lognormal and many other) distribution
see for example $m=1;\ \ n=2,\ t=t_0<t_1<t_2=t+h$:
to prove: $(+)\Rightarrow$ conditions of compatibility, e.g.

$$Q_{t_0t_1t_2}(B_0\times\mathbb{R}\times B_2)=Q_{t_0t_2}(B_0\times B_2)$$

$(*)\Rightarrow$

$$Q_{t_0t_1t_2}(B_0\times\mathbb{R}\times B_2)=\int\limits_{B_0}P_0(dy_0)\underbrace{\int\limits_{\underbrace{\mathbb{R}-y_0}_{=\mathbb{R}}}P(t_0,t_1-t_0,dy_1)\int\limits_{B_2-(y_0+y_1)}P(t_1,t_2-t_1,dy_2)}_{=H}$$

From (+) it follows

$$\begin{aligned}
\mathbf{P}(\omega:\eta_1+\eta_2\in C) &= \int\limits_C P(t_0,t_2-t_0,dz)\\
&= \int\limits_{\mathbb{R}}\int\limits_{C-y_1} P(t_0,t_1-t_0,dy_1)\cdot P(t_1,t_2-t_1,dy_2)
\end{aligned}$$

set $C=B_2-y_0$ the following identity for H can derived:

$$\int\limits_{B_2-y_0}P(t_0,t_2-t_0,dz)=\int\limits_{\mathbb{R}}P(t_0,t_1-t_0,dy_1)\int\limits_{B_2-(y_0+y_1)}P(t_1,t_2-t_1,dy_2)=H$$

$$\Rightarrow \quad Q_{t_0t_1t_2}(B_0 \times \mathbb{R} \times B_2) = \int\limits_{B_0} P_0(dy_0) \int\limits_{B_2-y_0} P(t_0, t_2 - t_0, dz)$$
$$\underset{(+)}{=} Q_{t_0t_2}(B_0 \times B_2)$$

Theorem 2.3 of Kolmogoroff implies the existence of a random process ξ with f.d.d.'s given in $(*)$
structure of these f.d.d.'s $\Rightarrow$ ξ possesses independent increments

It is convenient to study processes with independent increments using characteristic functions

Definition 4.2 $\qquad J(t,h,u) := \int\limits_{\mathbb{R}^m} e^{i(u,x)} P(t,h,dx)$
is called **characteristic function** *of the process with independent increments.*
$J(t,h,u)$ *is the characteristic function of the increment* $\xi(t+h) - \xi(t)$.

Theorem 4.1 *A process with independent increments is given in the wide sense if and only if*
$P_0(B)$ *the initial distribution and*
$J(t,h,u)$ *the characteristic function of* $\xi(t+h) - \xi(t)$
are given where $J(t,h,u)$ *satisfies the condition*

$$J(t, h_1 + h_2, u) = J(t, h_1, u) \cdot J(t + h_1, h_2, u). \qquad (\circ)$$

Remark 4.2 *The condition* $(\circ)$ *to* $J(t,h,u)$ *corresponds to property (+) of* $P(t,h,B)$.

Proof: Given $J(t,h,u)$ and using the independence of the increments the characteristic function of the joint distribution of

$$\big(\xi(t_1) - \xi(t_0), \xi(t_2) - \xi(t_1), \ldots, \xi(t_n) - \xi(t_{n-1})\big)$$

is given by

$$J(t_1, \ldots, t_n; u_1, \ldots, u_n) = \prod_{k=1}^{n} J(t_{k-1}, t_k - t_{k-1}, u_k)$$

Characteristic function defines uniquely the distribution $\overset{(*)}{\Rightarrow}$ f.d.d.'s of ξ are given.

condition $(\circ)$: consider $t_0 = t, \quad t_1 = t + h_1, \quad t_2 = t + h_1 + h_2;$
$\eta_1 = \xi(t_1) - \xi(t_0), \quad \eta_2 = \xi(t_2) - \xi(t_1);$

r.v.	distribution		characterristic function
η_1	$P(t_0, h_1, B)$	$\Leftrightarrow$	$J(t_0, h_1, u)$
η_2	$P(t_1, h_2, B)$	$\Leftrightarrow$	$J(t_1, h_2, u)$
$\eta_1 + \eta_2$	$P(t_0, h_1 + h_2, B)$	$\Leftrightarrow$	$J(t_0, h_1 + h_2, u)$

$$\Rightarrow \quad J(t_0, h_1 + h_2, u) = \mathbf{E} \exp(i(u, \eta_1 + \eta_2))$$
$$\overset{\eta_1,\eta_2 \text{ independent}}{=} \mathbf{E} \exp(i(u, \eta_1))\mathbf{E} \exp(i(u, \eta_2))$$
$$= J(t_0, h_1, u) J(t_1, h_2, u)$$

$\Rightarrow$ condition $(\circ)$

Definition 4.3 *A random process* $\xi(t)$, $t \in T$, *is called* **stochastically continuous** *in* t_0, *if for any* $\varepsilon > 0$ *it holds*

$$\lim_{t \to t_0} \mathbf{P}(\,|\xi(t,\omega) - \xi(t_0,\omega)| \geq \varepsilon) = 0.$$

Definition 4.4 *A process with independent increments is called* **homogeneous**, *if the differences* $\xi(t+h) - \xi(t)$ *are distributed independently of* t, *i.e. it holds*

$$P(t,h,B) = P(h,B) \qquad \text{and consequently} \qquad J(t,h,B) = J(h,B) \qquad .$$

$\Rightarrow$ A homogeneous process with independent increments is stochastically continuous, if

$$\begin{aligned} \lim_{h \to 0} P(h, \overline{U_\varepsilon}) &= 0 \\ \text{where} \qquad \overline{U_\varepsilon} &= \mathbb{R}^m \setminus U_\varepsilon \quad \text{and} \quad U_\varepsilon = \{x \in \mathbb{R}^{\mathrm{m}} : |\mathrm{x}| < \varepsilon\}, \quad \varepsilon > 0. \end{aligned}$$

see $P(h, \overline{U_\varepsilon}) = \mathbf{P}(|\xi(t_0+h) - \xi(t_0)| \geq \varepsilon)$

Definition 4.5 *A stochastically continuous and homogeneous process with independent increments is called* **Levy process**.

Theorem 4.2

1. *Let* $\xi(t)$, $t \in T$ *be a Levy process. Then for* $h \to 0$
   (a) *the increment* $\xi(t+h) - \xi(t)$ *converges to zero in probability,*
   (b) *the distribution of* $\xi(t+h) - \xi(t)$ *converges weakly to zero (the measure concentrated in zero).*
2. $\xi(t)$ *is stochastically continuous if and only if* $\lim\limits_{h \to 0} J(h,u) = 1$ *uniformly in any bounded region* $|u| \leq K$.

Proof:

1. (a) $\xi(t)$ stochastically continuous $\overset{\text{Definition 4.3}}{\Rightarrow}$

   $$\lim_{h \to 0} \mathbf{P}(\,|(\xi(t+h) - \xi(t)) - 0| \geq \varepsilon) = 0$$

   $\Rightarrow$ $\xi(t+h) - \xi(t)$ converges in probability to 0

   (b) convergence in probability implies weak convergene (see basic course Theorem 5.8.),
   i.e. for any continuous and bounded function f it holds
   $\int\limits_{\mathbb{R}^m} f(x)P(h,dx) \overset{h \to 0}{\longrightarrow} f(0)$

2. $(\Rightarrow)$: $J(h,u) = \int\limits_{R^m} e^{i(u,x)} P(h,dx) \overset{1(b).}{\longrightarrow} 1, \quad (e^{i(u,x)}$ bounded for $|u| \leq K)$

   $(\Leftarrow)$: let $\lim\limits_{h \to 0} J(h,u) = 1$, limit function $J^*(u) \equiv 1$ is continuous at $u = 0$
   $\overset{\text{Theorem 1.15}}{\Longrightarrow}$ $J^*(u)$ is a characteristic function of a distribution $P^*(B)$ in $\mathbb{R}^m$ and $\lim\limits_{h \to 0} P(h,B) = P^*(B)$

   here $P^*(B) = \begin{cases} 1 & \text{for} \quad 0 \in B \\ 0 & \text{for} \quad 0 \notin B \end{cases}$ (Dirac measure)

   $\Rightarrow \lim\limits_{h \to 0} P(h, \overline{U_\varepsilon}) = 0 \Rightarrow \xi(t)$ is stochastically continuous

Theorem 4.3 *The characteristic function of a Levy process possesses the following properties:*

1. $J(h_1 + h_2, u) = J(h_1, u)J(h_2, u)$, *especially* $J(nh, u) = [J(h, u)]^n \quad \forall\ n \in \mathbb{N}$
2. $J(h, u) \neq 0$
3. $J(h, u) = \exp(h\, g(u))$ *with a uniquely defined continuous function* $g(u)$;
   ($\Rightarrow$ *property completely characterizes the dependence of* $J(h, u)$ *on* h)

Proof:

1. follows from Theorem 4.1 applied to homogeneous processes
2. Theorem 4.2 $\Rightarrow \lim\limits_{h\to 0} J(h, u) = 1 \quad \forall\ |u| \leq K$
   $\Rightarrow \quad \exists\, h_0 > 0 : \ |J(h, u)| \geq \frac{1}{2} \quad \forall \ h \in [0, h_0]$
   let h be arbitrary $\Rightarrow \quad h = h_0(n + \theta)$ where $0 \leq \theta < 1$
   1. $\Rightarrow \quad J(h, u) = J(h_0 n, u)J(h_0\theta, u) = [J(h_0, u)]^n J(h_0\theta, u)$
   $\Rightarrow \quad |J(h, u)| \geq (\frac{1}{2})^{n+1} \qquad \Rightarrow$ assertion
3. since $\lim\limits_{h\to 0} J(h, u) = 1$ uniformly $\forall\ |u| \leq K$ there exists a uniquely defined function $g_1(h, u) = \ln J(h, u) \quad \forall\ |u| \leq K; \quad h \in [0, s]$ where $s = s(K)$
   $g_1(h, u)$ is continuous in $[0, s] \times \{|u| \leq K\}$
   because of 1. it holds :

$$\begin{aligned} g_1(h_1 + h_2, u) &= g_1(h_1, u) + g_1(h_2, u),\ \forall\ |u| \leq K,\ h_1, h_2 > 0,\ h_1 + h_2 \leq s \\ \Rightarrow \ g_1(h, u) &= h \cdot g(u) \text{ with an uniquely defined continuous function } g \\ \Rightarrow \ J(h, u) &= \exp(h \cdot g(u)) \quad \text{holds for any fixed } u \text{ and } \textit{all } h \end{aligned}$$

if $J(h, u) = \exp(hg(u))$ holds only for $0 < h \leq h_0$ then for any $h > 0$ it follows
$J(h, u) \underset{1.}{=} [J(\frac{h}{n}, u)]^n = [\exp(\frac{h}{n} g(u))]^n = \exp(hg(u)) \quad \text{for} \ \ n > \frac{h}{h_0} \ (\frac{h}{n} < h_0)$
$\Rightarrow$ assertion

$g(u)$ can be an arbitrary continuous function provided that $\exp(hg(u))$ is the characteristic function of a certain distribution for each h

Theorem 4.4 Levy-Chinchin-representation
Let $J(h, u),\ h > 0,\ u \in \mathbb{R}^m$ *be a family of characteristic functions such that the limit*

$$g(u) = \lim_{h\to 0} \tfrac{1}{h}(J(h, u) - 1) \quad \textit{for} \quad |u| \leq K \quad (K > 0 \ \textit{ arbitrarily}),$$

uniformly exists. Then there exists a finite measure Π *on* $\{\mathbb{R}^m, \mathcal{B}^m\}$ *with* $\Pi(\{0\}) = 0$, *a non-negative definite matrix* $B \in \mathbb{R}^{m\times m}$ *and a vector* $a \in \mathbb{R}^m$, *such that*

$$g(u) = i(a, u) - \frac{1}{2}(B\, u, u) + \int\limits_{\mathbb{R}^m} \left[e^{i(u,z)} - 1 - \frac{i(u, z)}{1 + |z|^2} \right] \frac{1 + |z|^2}{|z|^2}\ \Pi(dz).$$

Proof: not given

If ξ is a Levy process then $J(h,u) = \exp(hg(u))$, hence $g(u) = \lim\limits_{h\to 0} \frac{1}{h}(J(h,u)-1)$, convergence is uniform in every bounded domain $|u| \le K,\ 0 < K < \infty$

see $\quad \frac{1}{h}(J(h,u)-1) = g(u) + \frac{h}{2!}g^2(u)\exp(\theta h g(u)),\ \theta \in [0,1]$, (Taylor expansion)

$g(u)$ is continuous for $|u| \le K \quad \Rightarrow \quad |g(u)| \le C = C(K)$

$\Rightarrow \quad \left|\frac{1}{h}(J(h,u)-1) - g(u)\right| \le \frac{h}{2}C^2 e^{hC}$ independent of u

$\Rightarrow$ uniform convergence

Theorem 4.5 *If $\xi(t),\ t > 0$ is a Levy process with values in R^m, then the characteristic function $J(h,u)$ of the increment $\xi(t+h) - \xi(t)$ is of the form*

$$J(h,u) = \exp(h\,g(u)),$$

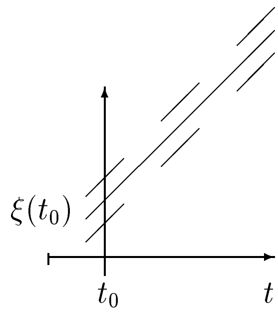
where $g(u)$ is given in Theorem 4.4.

Some special cases:

(1) $B = 0,\quad \Pi \equiv 0 \quad \Rightarrow \quad g(u) = i(a,u); \quad J(h,u) = \exp(ih(a,u))$

$J(h,u)$ is the characteristic function of a degenerated distribution concentrated at the point $ha \in \mathbb{R}^m \quad \Rightarrow \xi(t+h) - \xi(t) = ah$ a.s. $\Rightarrow$
$\xi(t) = \xi(0) + at$ a.s. moves uniformly with constant velocity a

(only the initial value $\xi(0)$ is random,
increments are non-random)



(2) $\Pi \equiv 0 \quad \Rightarrow \quad J(h,u) = \exp(ih(a,u) - \frac{1}{2}h(B\,u,u))$

Theorem 3.2 $\Longrightarrow$ increments $\xi(t+h) - \xi(t)$ are normally distributed
with mean value ha and covariance matrix hB
let $\xi(0) = 0$, then $\xi(t)$ is a Gaussian process
it can be shown that $\xi(t)$ is stochastically equivalent to a process with continuous paths which is called Brownian motion
(see motion of a small particle in a gas or liquid)
for $m = 1$ the one-dimensional process $\xi(t)$ is the so-called Wiener process.

Definition 4.6 *A (scalar) stochastic process $W(t),\ t \in [0,\infty)$ is called* **Wiener process**, *if*

1. $\mathbf{P}(W(0) = 0) = 1$,
2. *for any* $\quad 0 \le t_0 < t_1 < \ldots < t_n \quad$ *the increments* $W(t_1) - W(t_0), \ldots, W(t_n) - W(t_{n-1})$ *are mutually independent,*

3. *for all t and $h > 0$ the increment $W(t+h) - W(t)$ is normally distributed with :*
$$\mathbf{E}\left(W(t+h) - W(t)\right) = 0,$$
$$\mathbf{E}\left(W(t+h) - W(t)\right)^2 = \sigma^2 h,$$
$\sigma^2 > 0$ is the variance of the Wiener process,

4. *W possesses continuous paths.*

For $\sigma^2 = 1$ the process is called **standard Wiener process**.

correlation function: $b(t_1, t_2) = \mathbf{E}\, W(t_1) W(t_2) = \sigma^2 \min(t_1, t_2)$
proof: let $t_1 < t_2$:

$$\begin{aligned}
&\mathbf{E}\, W(t_1)W(t_2)\\
&= \mathbf{E}\left(W(t_1) - W(0)\right)\left(W(t_2) - W(t_1) + W(t_1) - W(0)\right) \quad (W(0) = 0 \ \text{ a.s.})\\
&= \mathbf{E}\left(W(t_1) - W(0)\right)\left(W(t_2) - W(t_1)\right) + \mathbf{E}\left(W(t_1) - W(0)\right)^2\\
&= \underbrace{\mathbf{E}\left(W(t_1) - W(0)\right)}_{0}\underbrace{\mathbf{E}\left(W(t_2) - W(t_1)\right)}_{0} + \underbrace{\mathbf{E}\left(W(t_1) - W(0)\right)^2}_{\sigma^2 t_1}\\
&\quad \text{(independence of increments)}\\
&= \sigma^2 t_1
\end{aligned}$$

(3) $a = 0,\ B = 0,\ \Pi:$ mass $q > 0$ is concentrated in $z_0 \in \mathbb{R}^m \setminus \{0\}$

$$\Rightarrow \int\limits_{\mathbb{R}^m} f(z)\Pi(dz) = q f(z_0)$$

$$\text{Theorem 4.4} \Rightarrow \ J(h,u) = \exp\left(h \underbrace{q\,\frac{(1+|z_0|^2)}{|z_0|^2}}_{=:\lambda} \left(e^{i(u,z_0)} - 1 - \frac{i(u,z_0)}{1+|z_0|^2}\right)\right)$$

$$= \underbrace{\exp(h\lambda(e^{i(u,z_0)} - 1))}_{\text{Poisson distribution}} \underbrace{\exp\left(-h\frac{q}{|z_0|^2} i(u,z_0)\right)}_{\text{degenerated distribution}}$$

it can be shown: $\xi(t) - \xi(0) \underset{t=h}{=} z_0 \left(N(t) - \frac{qt}{|z_0|^2}\right)$,

where $N(t)$ is a so-called Poisson process with the parameter $\lambda = q\frac{(1+|z_0|^2)}{|z_0|^2}$

Definition 4.7 *A (scalar) stochastic process $N(t),\ \ t \in [0,\infty)$ with homogeneous and independent increments is called* **Poisson process** *with the parameter $\lambda > 0$, if*

1. $\mathbf{P}(N(0) = 0) = 1$,

2. *for any $0 \le t_0 < t_1 < \ldots < t_n$ the increments $N(t_1) - N(t_0), \ldots, N(t_n) - N(t_{n-1})$ are mutually independent,*

3. *for all t and $h > 0$ the increment $N(t+h) - N(t)$ is Poisson distributed with the parameter λh,*

4. the paths of N are continuous on the right.

Definition 4.8 *A random variable η with values $k = 0, 1, \ldots$ possesses a Poisson distribution with the parameter $\nu > 0$, if*

$$\mathbf{P}(\eta = k) = \frac{1}{k!}\nu^k e^{-\nu} \quad \textit{for} \quad k = 0, 1, 2, \ldots .$$

it holds:

$$\begin{aligned}
\mathbf{E}\,\eta = \mathbf{D}^2\,\eta &= \nu \\
J(u) &= \exp\big(\nu(e^{iu} - 1)\big) \qquad \text{(characteristic function)} \\
\Rightarrow \quad \mathbf{P}\big(N(t+h) - N(t) = k\big) &= \frac{1}{k!}(\lambda h)^k e^{-\lambda h} \\
\mathbf{P}\big(N(t) = k\big) &= P\big(N(t) - N(0) = k\big) = \frac{1}{k!}(\lambda t)^k e^{-\lambda t} \\
\Rightarrow \ \mathbf{E}\,N(t) &= \lambda t \\
\mathbf{D}^2\,N(t) &= \mathbf{E}\,\big(N(t) - \mathbf{E}\,N(t)\big)^2 = \lambda t
\end{aligned}$$

correlation function:

$$\begin{aligned}
b(t_1, t_2) &= \mathbf{E}\,\big(N(t_1) - \lambda t_1\big)\big(N(t_2) - \lambda t_2\big) = \mathbf{E}\,N(t_1)N(t_2) - \lambda^2 t_1 t_2 \\
&= \lambda \min(t_1, t_2)
\end{aligned}$$

proof: let $t_1 < t_2$:

$$\begin{aligned}
\mathbf{E}\,N(t_1)N(t_2) &= \mathbf{E}\,N(t_1)\big(N(t_2) - N(t_1) + N(t_1)\big) \qquad (N(0) = 0 \text{ f.s.}) \\
&= \mathbf{E}\,\big(N(t_1) - N(0)\big)\big(N(t_2) - N(t_1)\big) + \mathbf{E}\,N(t_1)^2 \\
&\qquad\qquad \text{independence of increments } \Rightarrow \\
&= \underbrace{\mathbf{E}\,\big(N(t_1) - N(0)\big)}_{\lambda t_1}\underbrace{\mathbf{E}\,\big(N(t_2) - N(t_1)\big)}_{\lambda(t_2 - t_1)} + \underbrace{\mathbf{E}\,\big(N(t_1)^2\big)}_{(EN(t_1))^2 + D^2N(t_1)} \\
&= \lambda t_1 \cdot \lambda(t_2 - t_1) + (\lambda t_1)^2 + \lambda t_1 = \lambda^2 t_1 t_2 + \lambda t_1
\end{aligned}$$

(4) $B = 0$, $\quad \Pi$ such that $\int\limits_{\mathbb{R}^m} \frac{1}{|z|^2}\Pi(dz) < \infty$ (property of Π with respect to zero)

Theorem 4.4 $\Rightarrow$ it exists a $q > 0$, a (probability) measure Π_0 on $\{\mathbb{R}^m, \mathcal{B}^m\}$ and $\hat{a} \in \mathbb{R}^m$ such that

$$g(u) = i(\hat{a}, u) + q \int\limits_{\mathbb{R}^m} (e^{i(u,z)} - 1)\Pi_0(dz)$$

since

$$\begin{aligned}
g(u) &= i(a, u) - \int\limits_{\mathbb{R}^m} \frac{i(u,z)}{|z|^2}\Pi(dz) + \int\limits_{\mathbb{R}^m} (e^{i(u,z)} - 1)\frac{1 + |z|^2}{|z|^2}\Pi(dz) \\
&= i\sum_{k=1}^{m} u_k \underbrace{\Big(a_k - \int\limits_{\mathbb{R}^m} \frac{z_k}{|z|^2}\Pi(dz) \Big)}_{\hat{a}_k} + \lambda \int\limits_{\mathbb{R}^m} (e^{i(u,z)} - 1)\Pi_0(dz)
\end{aligned}$$

$$\text{with } \Pi_0(B) := \frac{1}{\lambda}\int\limits_{B} \frac{1 + |z|^2}{|z|^2}\Pi(dz) \qquad \text{and} \qquad \lambda := \int\limits_{\mathbb{R}^m} \frac{1 + |z|^2}{|z|^2}\Pi(dz)$$

$$\Rightarrow\ J(h,u) = \exp\big(hg(u)\big) = \exp\big(ih(\hat{a},u)\big)\exp\left(h\lambda\int\limits_{\mathbb{R}^m} e^{i(u,z)}\Pi_0(dz) - h\lambda\right)$$

$$= exp\big(ih(\hat{a},u)\big)\sum_{k=0}^{\infty} e^{-h\lambda}\frac{1}{k!}(h\lambda)^k\left[\int\limits_{\mathbf{R}^m} e^{i(u,z)}\Pi_0(dz)\right]^k \qquad (+)$$

$\Rightarrow$ $J(h,u)$ is the characteristic function of the sum $h\hat{a}+\eta_1+\cdots+\eta_{N(h)}$ where

1. $\eta_1,\eta_2,\ldots \in \mathbb{R}^m$ are independent random vectors with the distribution Π_0,
2. $\hat{a}$ is a constant vector and
3. $N(h)$ is a random variable which is independent of $\{\eta_k,\ k=0,1,\ldots\}$ possessing a Poisson-distribution with the parameter $h\lambda$ ($N(.)$ is a Poisson process).

Let $\xi(0)=0$ a.s. then

$$\xi(t) = \xi(t)-\xi(0) \underset{h=t}{=} \hat{a}t + \sum_{k=1}^{N(t)} \eta_k$$

Such a process is called **generalized** or **compound Poisson process** with values in $\mathbb{R}^m$.

$\boxed{\text{figure}}$

Let $\quad I_k = \begin{cases} 1 & \text{for} \quad N(h)=k \\ 0 & \text{else} \end{cases} \quad$ then:

$$\mathbf{E}\exp\big(i(u,h\hat{a}+\eta_1+\ldots+\eta_{N(h)})\big) = \mathbf{E}\sum_{k=0}^{\infty} I_k \exp\big(i\,(u,h\hat{a}+\eta_1+\ldots+\eta_k)\big)$$

$$= \sum_{k=0}^{\infty}\mathbf{E}\,I_k\mathbf{E}\,\exp\big(i\,(u,h\hat{a}+\eta_1+\cdots+\eta_k)\big)$$

$$= \sum_{k=0}^{\infty} \underbrace{\frac{1}{k!}(hq)^k e^{-hq}}_{EI_k=P(N(h)=k)} \quad e^{i(u,h\hat{a})} \quad \underbrace{\prod_{p=1}^{k}\mathbf{E}\,e^{i(u,\eta_p)}}_{\eta_1,\ldots,\eta_k \text{ are i.i.d}}$$

$$= e^{i(u,h\hat{a})}\sum_{k=0}^{\infty} e^{-hq}\frac{1}{k!}(hq)^k\left[\int\limits_{\mathbb{R}^m} e^{i(u,z)}\Pi_0(dz)\right]^k$$

$$\Rightarrow (+)$$

(5) a,Π as in (4), $B\neq 0$

Theorem 4.4 $\Rightarrow$

$$J(h,u) = \underbrace{\exp\left(h(i(a,u)+\int\ldots\Pi(dz)\right)}_{(4)\ \Rightarrow\ \text{c.f. of } h\hat{a}+\eta_1+\ldots\eta_{N(h)}}\cdot\underbrace{\exp(h(Bu,u))}_{(2)\ \Rightarrow\ \text{c.f. of } \beta}$$

where β is a Gaussain random vector
with zero mean and covariance matrix hB
independent of $\eta_1, \eta_2, \ldots$, and the Poisson process $N(h)$.

Let $\xi(0) = 0$ and $m = 1 \underset{(2),(4)}{\Longrightarrow}$

$$\xi(t) = \hat{a}t + W(t) + \sum_{k=1}^{N(t)} \eta_k$$

ξ is an example of a **jump-diffusion process** with values in $\mathbb{R}^1$

$\boxed{\text{figure}}$

$\hat{a}$ - drift
$W(t)$ - Wiener process (diffusion)
η_k - jump heights
$N(t)$ - jump times

5 Markov processes

Let $\xi(t),\ t \in T$ be a random process with values in $\{\mathcal{S}, \mathcal{C}\}$

$\mathcal{S}$: complete, separable, metric space (state space)

$\mathcal{C}$: σ-algebra of Borel sets in $\mathcal{S}$ (w.r.t. the metric in $\mathcal{S}$)

T : a finite or infinite interval of time

assume: <u>independence of the future from the past</u> or <u>absence of an aftereffect</u>:

$$\mathbf{P}\{\xi(t) \in C |\, \xi(t_1), \ldots, \xi(t_n)\} \;=\; \mathbf{P}\{\xi(t) \in C |\, \xi(t_n)\} \qquad \forall\;\; C \in \mathcal{C};\quad \forall\;\; t_1 < t_2 < \ldots < t_n < t$$

remember: $\mathbf{P}\{\underbrace{\xi(t) \in C}_{=:A} |\, \xi(s)\} = \mathbf{P}\{A |\, \xi(s)\} = \mathbf{P}\{A | \mathcal{F}_{\xi(s)}\}$

where ($\mathcal{F}_{\xi(s)}$ is the smallest σ-algebra such that $\xi(s)$ is measurable)

<u>interpretation :</u> $\xi(t)$ is the position of a particle, the stock price,…, at time t

let $\underbrace{t}_{\text{future}} > \underbrace{t_n}_{\text{present}} > \underbrace{t_{n-1} > \ldots > t_1}_{\text{past}}$;

$\xi(t)$ depends only on the present value $\xi(t_n)$ but is independent of the past values $\xi(t_i)$ with $i < n$;

Theorem 1.29: $\Rightarrow$ $\quad \mathbf{E}\{\xi|\eta\} = d(\eta)$, where

$\mathbf{E}\{\xi\|\eta\}$	is η-measurable	i.e. $\{\Omega, \mathcal{F}_\eta\}$	$\overset{\text{measurable}}{\longrightarrow}$ $\{\mathbb{R}, \mathcal{B}\}$
η	is a random element in $\{\mathcal{S}, \mathcal{C}\}$	i.e. $\{\Omega, \mathcal{A}\}$	$\overset{\text{measurable}}{\longrightarrow}$ $\{\mathcal{S}, \mathcal{C}\}$
d	is a $\mathcal{C}$-measurable function	i.e. $\{\mathcal{S}, \mathcal{C}\}$	$\overset{\text{measurable}}{\longrightarrow}$ $\{\mathbb{R}, \mathcal{B}\}$

$\Rightarrow \quad \mathbf{P}\{\xi(t) \in C |\, \xi(s)\} \;=\; P(s, \xi(s), t, C) \quad$ for $s < t$,

since $\mathbf{P}\{\xi(t) \in C |\, \xi(s)\} = \mathbf{E}\{\chi_{\xi(t)\in C} |\, \xi(s)\}$ is a $\xi(s)$-measurable function ($\mathcal{F}_{\xi(s)}$-measurable)

Definition 5.1 *A random process $\xi(t),\ t \in T$, with values in $\mathcal{S}$ is called a* **Markov process**, *if*

1. $\mathbf{P}\{\xi(t) \in C|\xi(t_1), \ldots, \xi(t_n)\} = \mathbf{P}\{\xi(t) \in C|\xi(t_n)\}$ (mod $\mathbf{P}$)
   $\forall\;\; t_1 < t_2 < \ldots < t_n < t,\;\; t_k,\; t \in T$

2. *there exists a function $P(s, y, t, C)$ with $s \le t$ such that*
   $P(s, ., t, C)$ is for fixed s, t, C; a $\mathcal{C}$-measurable function,
   $P(s, y, t, .)$ is for fixed s, y, t a probability measure on $\mathcal{C}$ and

   (a) P satisfies the **Chapman-Kolmogorov-equation**

   $$P(t_1, y, t_3, C) = \int_{\mathcal{S}} P(t_2, y_2, t_3, C) P(t_1, y, t_2, dy_2)$$

   for $t_1 < t_2 < t_3$ and

   (b) P coincides with probability 1 with the conditional probability $\mathbf{P}\{\xi(t) \in C|\xi(s)\}$, i.e. $P(s, \xi(s), t, C) = \mathbf{P}\{\xi(t) \in C|\xi(s)\}$ (mod P).

The functions $P(s, y, t, C)$ are called **transition probabilities** *of the Markov process.*

- condition 1 is called the **Markov-property** or as the **absence of after-effects**
- condition 2a (Chapman-Kolmogorov equation) is a condition of compatibility or consistence to $P(s, y, t, C)$ for the times $t_1 < t_2 < t_3$

  $\boxed{\text{figure}}$
- condition 2b means that $\mathbf{P}\{\xi(t) \in C \,|\, \xi(s)\}$ is a regular conditional distribution (see Definition 1.30), i.e.
  - $P(s, \xi(s), t, C)$ is for fixed ω a probability on $\mathcal{C}$ $(\xi(s) = \xi(s, \omega)$!)
  - $P(s, \xi(s), t, C) = \mathbf{P}\{\xi(t) \in C|\xi(s)\}$ a.s. for any $C \in \mathcal{C}$
  - for $s = t$ it holds $P(t, y, t, C) = \begin{cases} 1 & \text{for } y \in C \\ 0 & \text{else} \end{cases}$

For a Markov process also holds the **generalized Markov property**:

$$\begin{aligned}\mathbf{P}\{(\xi(t_{m+1}), \ldots, \xi(t_{m+n})) \in C^n \,|\, \xi(t_1), \ldots, \xi(t_m)\} = \\ \mathbf{P}\{(\xi(t_{m+1}), \ldots, \xi(t_{m+n})) \in C^n \,|\, \xi(t_m)\} \quad (\mathrm{mod}\ \mathbf{P})\end{aligned} \qquad (*)$$

with a Borel set C^n from $\mathcal{S}^n$ and $\forall\ t_1 < t_2 < \ldots < t_{n+m}$, $\forall\ n, m \in \mathrm{N}$
(i.e. it holds the generalization of Definition 5.1 1. !)

Further generalization of $(*)$ leads to the following

Theorem 5.1 *Let $\xi(t)$, $t \in T$, be a Markov process. Moreover, let $\mathcal{F}_t = \sigma\{\xi(s), s \leq t,\ s \in T\}$ and $\mathcal{F}_t^* = \sigma\{\xi(s),\ s > t,\ s \in T\}$. Then it holds*

$$\mathbf{P}\{C \,|\, \mathcal{F}_t\} = \mathbf{P}\{C \,|\, \xi(t)\} \quad (\mathrm{mod}\ \mathbf{P}) \qquad \forall\ \ C \in \mathcal{F}_t^*.$$

Proof: not given

Remark 5.1 The theorem shows that the dependence of an event in the future $C \in \mathcal{F}_t^*$ on the past and present $\mathcal{F}_t$ is completely determined by the dependence on the present state $\xi(t)$.

$\boxed{\text{figure}}$

Definition 5.2 *Let $T = [0, b]$ or $T = [0, \infty)$ and $\{\mathcal{S}, \mathcal{C}\}$ be a measurable space. A family consiting of*

1. *an initial distribution P_0 on $\{\mathcal{S}, \mathcal{C}\}$ and*
2. *transition probabilities $P(s, y, t, C)$ with $t > s$, t, $s \in T$, $C \in \mathcal{C}$, satisfying the conditions of Definition 5.1 2.*

is called **Markov process in the wide sense** *with values in $\{\mathcal{S}, \mathcal{C}\}$*

Theorem 5.2 (f.d.d.'s of a Markov process)
Let $\xi(t),\ t \in T$ be a Markov process in the wide sense with values in $\{\mathcal{S}, \mathcal{C}\}$ with
the initial distribution $P_0(C) \quad (P_0(C) = \mathbf{P}(\xi(0) \in C); \quad C \in \mathcal{C})$ *and*
the transition probabilities $P(s, y, t, C)$.
Then it holds $\forall \quad 0 < t_1 < \ldots < t_n, \quad C_k \in \mathcal{C}$:

$$\mathbf{P}(\xi(t_1) \in C_1, \ldots, \xi(t_n) \in C_n) =$$

$$\int\limits_{\mathcal{S}} \int\limits_{C_1} \ldots \int\limits_{C_n} \mathbf{P}(\xi(t_0) \in dy_0, \xi(t_1) \in dy_1, \ldots, \xi(t_n) \in dy_n)$$

$$= \int\limits_{\mathcal{S}} P_0(dy_0) \int\limits_{C_1} P(0, y_0, t_1, dy_1) \int\limits_{C_2} P(t_1, y_1, t_2, dy_2)$$

$$\cdot \ldots \cdot \int\limits_{C_{n-1}} P(t_{n-2}, y_{n-2}, t_{n-1}, dy_{n-1}) P(t_{n-1}, y_{n-1}, t_n, C_n).$$

<u>Special cases:</u>

1. $$\mathbf{P}(\xi(t_1) \in C_1) = \int\limits_{\mathcal{S}} \underbrace{P(0, y_0, t_1, C_1)}_{\mathbf{P}\{\xi(t_1) \in C_1 \mid \xi(0) = y_0\}} P_0(dy_0)$$

(see formula of total probability)

2. $$\mathbf{P}(\xi(t_1) \in C_1, \xi(t_2) \in C_2) = \int\limits_{\mathcal{S}} P_0(dy_0) \underbrace{\int\limits_{C_1} P(0, y_0, t_1, dy_1) P(t_1, y_1, t_2, C_2)}_{\int\limits_{C_1} P\{\xi(t_2) \in C_2 \mid \xi(t_1) = y_1\} P(0, y_0, t_1, dy_1)}$$

see $\int\limits_{C_1} P(0, y_0, t_1, dy_1) P(t_1, y_1, t_2, C_2) = \mathbf{P}\{\xi(t_2) \in C_2;\ \xi(t_1) \in C_1 \mid \xi(0) = y_0\}$

Theorem 5.3 *Let $\mathcal{S}$ be a complete, separable and metric space, then for every Markov process in the wide sense there exists a representation (in the sense of Definition 2.6).*
Proof: construction of f.d.d.'s (see Theorem 5.2);
they satisfy the conditions of compatibility
Theorem of Kolmogorov $\Rightarrow$ assertion

Definition 5.3 *A* **Markov process** *is called* **homogeneous** *if for any $t \in T$ and $h \geq 0$, $t + h \in T$, the transition probabilities $P(t, y, t + h, C)$ are independent of t, i.e.*

$$P(t, y, t + h, C) = P(y, h, C).$$

In this case the Chapman-Kolmogorov equation becomes

$$P(y, h_1 + h_2, C) = \int\limits_{\mathcal{S}} P(y_2, h_2, C) P(y, h_1, dy_2) \tag{3}$$

$\boxed{\text{figure}}$

It can be seen that

- it is enough to know $P(y,h,C)$ for all $h \le \varepsilon$, where ε is arbitrarily small, because for $\tilde{h} > \varepsilon$ $P(y,\tilde{h},C)$ can be determined by (3).
- the local behaviour of the process determines the global behaviour.

Let $\xi(t),\ t \in T$ be a process with independent increments :

$$\Rightarrow \quad \xi(t_n) = \sum_{k=1}^{n} (\xi(t_k) - \xi(t_{k-1}))$$

$\Rightarrow$ $\xi(t_n)$ is a sum of independent random elements;
it can be shown that such a process is a Markov process
i.e. it holds $\quad \mathbf{P}\{\xi(t_n) \in C \,|\, \xi(t_1), \ldots, \xi(t_{n-1})\} = \mathbf{P}\{\xi(t_n) \in C \,|\, \xi(t_{n-1})\}$;
$\Rightarrow$ Wiener and Poisson process are examples of Markov processes;

for the Wiener process the transition probabilities are given by

$$P(s,x,t,C) = \int_C \frac{1}{\sqrt{2\pi(t-s)}} \exp\left(-\frac{(y-x)^2}{2(t-s)}\right) dy = P(x, t-s, C)$$

(i.e. the Wiener process is a homogeneous Markov process)
Proof: Theorem 1.35 $\Rightarrow \mathbf{E}\{f(\eta_2)|\eta_1\} = \int_{\mathcal{S}_2} f(x_2) \frac{\varrho(\eta_1,x_2)}{\varrho_{\eta_1}(\eta_1)} m_2(dx_2)$
where $\varrho(x_2|x_1) = \frac{\varrho(x_1,x_2)}{\varrho_{\eta_1}(x_1)}$ is the conditional density of η_2 given $\eta_1 = x_1$
here $\quad \eta_1 = \xi(s), \quad \eta_2 = \xi(t); \quad f(\eta_2) = I_C(\xi(t)), \quad s < t$
$\varrho(x_1,x_2)$ density of $(\xi(s),\xi(t))$ w.r.t. $m_1 \times m_2$ (Lebesgue measure in $\mathbb{R}^2$);
$\varrho_{\eta_1}(x_1)$ density of $\eta_1 = \xi(s)$

$$\begin{aligned} \Rightarrow \quad \mathbf{E}\{f(\eta_2)|\eta_1 = x_1\} &= \mathbf{P}\{\xi(t) \in C \,|\, \xi(s) = x_1\} = P(s, x_1, t, C) \\ &= \int_{\mathcal{S}_2} I_C(x_2) \frac{\varrho(x_1,x_2)}{\varrho_{\eta_1}(x_1)} dx_2 = \int_C \frac{\varrho(x_1,x_2)}{\varrho_{\eta_1}(x_1)} dx_2 \end{aligned}$$

the Wiener process is a Gaussian process, Theorem 3.2 $\Rightarrow$ density of $(\xi(s),\xi(t))$ is given by

$$\varrho(x_1,x_2) = \frac{1}{2\pi\sqrt{\det B}} \exp\left(-\frac{1}{2} x^\tau B^{-1} x\right) \qquad \text{with } x^\tau = (x_1,x_2)^\tau$$

$$\text{and } B = \begin{pmatrix} \mathbf{E}\,\xi^2(s) & \mathbf{E}\,\xi(s)\xi(t) \\ \mathbf{E}\,\xi(t)\xi(s) & \mathbf{E}\,\xi^2(t) \end{pmatrix} = \begin{pmatrix} s & \min(s,t) \\ \min(t,s) & t \end{pmatrix} = \begin{pmatrix} s & s \\ s & t \end{pmatrix}$$

$$\det B = st - s^2 = s(t-s); \qquad B^{-1} = \frac{1}{s(t-s)} \begin{pmatrix} t & -s \\ -s & s \end{pmatrix}$$

Using the relations

$$\begin{aligned} s(t-s)\, x^\tau B^{-1} x &= x_1^2 t - 2x_1x_2 s + x_2^2 s \\ &= x_1^2 t - x_1^2 s - 2x_1x_2 s + x_2^2 s + x_1^2 s \\ &= x_1^2(t-s) + s(x_1-x_2)^2 \end{aligned}$$

it follows $\quad \varrho(x_1, x_2) = \frac{1}{2\pi\sqrt{s(t-s)}} \exp\left(-\frac{x_1^2}{2s} - \frac{(x_1-x_2)^2}{2(t-s)}\right)$

using $\varrho_{\eta_1}(x_1) = \frac{1}{\sqrt{2\pi s}} \exp\left(-\frac{x^2}{2s}\right)$ it follows

$$\varrho(x_2 \mid x_1) = \frac{\varrho(x_1, x_2)}{\varrho_{\eta_1}(x_1)} = \frac{1}{\sqrt{2\pi(t-s)}} \exp\left(-\frac{(x_1-x_2)^2}{2(t-s)}\right)$$

$\Rightarrow$ assertion

6 Weakly correlated random functions

Idea: $\xi(x)$ random function; $x \in \mathbb{R}^m$, $m \in \mathbb{N}$
$\xi(x), \xi(y)$ independent if the distance between x and y is larger than ε
$\xi(x)$ "affects" the values $\xi(y)$ only in an ε-neighbourhood of x
values of ξ show only short range dependence

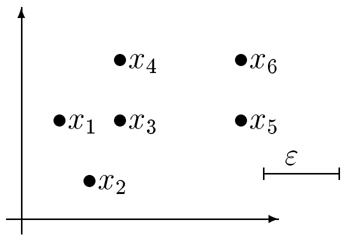
figure

introduced by Ornstein and Uhlenbeck (1930) in relation with the description of Brownian motion of a particle
$\xi(x)$-force acting on a particle caused by the impacts of other particles, x-time
($\varepsilon = 0$ corresponds to "white noise")

Definition 6.1 *Let $\varepsilon > 0$ be arbitrarily, $I = \{1, 2, \ldots, n\}$ be an index set. A set of points $Q = \{x_i : \ x_i \in \mathbb{R}^m, \ i \in I\} = \{x_1, \ldots, x_n\}$ is called* **ε-adjoining**, *if it can not be separated into two subsets with a distance larger than ε, i.e. for any two subsets I_1, I_2 of I with $I_1 \cap I_2 = \emptyset$ and $I_1 \cup I_2 = I$ it holds $J(I_1) \cap J(I_2) \neq \emptyset$, where $J(I_p) := \bigcup\limits_{i \in I_p} \{j \in I : \mid x_i - x_j \mid \leq \varepsilon\}$, $p = 1, 2$.*

A single point is always said to be ε-adjoining.

A set of points $\tilde{Q} = \{x_i : i \in \tilde{I} \subset I\}$ is called **maximally ε-adjoining** *w.r.t. the set Q, if $\tilde{Q}$ is ε-adjoining and there exists no $r \in I \setminus \tilde{I}$ such that $Q \cup \{x_r\}$ is ε-adjoining.*



Example 6.1 $Q = \{x_1, \ldots, x_6\}, \quad x_i \in \mathbb{R}^2$

$\{x_1, x_2, x_3\}$	*is ε-adjoining but it is not maximally ε-adjoining, since $\{x_1, x_2, x_3\} \cup \{x_4\}$ is ε-adjoining;*
$\{x_1, x_2, x_3, x_4\}$	*is maximally ε-adjoining;*
$\{x_1, x_2, x_5\}$	*is not ε-adjoining, since* $\underbrace{J(\{1,2\})}_{\{1,2,3\}} \cap \underbrace{J(\{5\})}_{\{5,6\}} = \emptyset$;

Any finite set of points in $\mathbb{R}^m$ can be separated uniquely into disjoint maximally ε-adjoining subsets of points.

in the above example: $Q = \{x_1, \ldots, x_6\} = Q_1 \cup Q_2$
$Q_1 = \{x_1, x_2, x_3, x_4\}$, $\quad Q_2 = \{x_5, x_6\}$ are maximally ε-adjoining

characterization of maximal ε-adjoining subsets:

- the points of the subsets form a "tree" where the length of the "branches" does not exceed ε
- the distance between the points of the subset to points outside the subset is larger than ε.

Definition 6.2 *A random function* $\xi_\varepsilon(x,\omega)$ *on* $\mathcal{D} \subset \mathbb{R}^m$ *and values in* $\mathbb{R}$ *with* $\mathbf{E}\{\xi_\varepsilon(x)\} = 0$ *is called* **weakly correlated** *with the* **correlation length** ε *if it holds*

$$\mathbf{E}\left\{\prod_{i\in I}\xi_\varepsilon(x_i)\right\} = \prod_{j=1}^{p}\mathbf{E}\left\{\prod_{i\in I_j}\xi_\varepsilon(x_i)\right\}$$

for any moments of order $k = 2, 3, \ldots,$ *where* $I = \{1, 2, \ldots, k\}$ *and*

$$Q_1 = \{x_i \in \mathcal{D} : i \in I_1\}, \ldots, Q_p = \{x_i \in \mathcal{D} : i \in I_p\} \quad \textit{with} \quad \bigcup_{j=1}^{p} I_j = I$$

is the partition of $\{x_i : i \in I\}$ *in maximally* ε*-adjoining subsets.*

interpretation: a moment w.r.t. a set of points $Q = \{x_i, i \in I\}$ is equal to the product of moments w.r.t. the disjoint maximally ε-adjoining subsets $Q_1, \ldots, Q_p$

in the above example:

$$\mathbf{E}\left\{\prod_{i=1}^{6}\xi_\varepsilon(x_i)\right\} = \mathbf{E}\left\{\xi_\varepsilon(x_1)\xi_\varepsilon(x_2)\xi_\varepsilon(x_3)\xi_\varepsilon(x_4)\right\} \cdot \mathbf{E}\left\{\xi_\varepsilon(x_5)\xi_\varepsilon(x_6)\right\}$$

for second-order moments ($k = 2$) it holds

$$\begin{aligned}
\mathbf{E}\left\{\xi_\varepsilon(x_1)\xi_\varepsilon(x_2)\right\} &= \begin{cases} K_\varepsilon(x_1, x_2) & \text{for } |x_1 - x_2| \le \varepsilon \\ \underbrace{\mathbf{E}\,\xi_\varepsilon(x_1)}_{=0}\underbrace{\mathbf{E}\,\xi_\varepsilon(x_2)}_{=0} = 0 & \text{else} \end{cases} \\
\text{e.g.} \qquad K_\varepsilon(x_1, x_2) &= \sigma^2\left(1 - \frac{|x_1 - x_2|}{\varepsilon}\right)
\end{aligned}$$

Remark 6.1 Note that correlation is usually related to second-order moments while the definition of weakly correlated function is based on a property of moments of any order.

Theorem 6.1 *Let* $\xi_\varepsilon(x,\omega)$ *be a weakly correlated with the correlation length* ε *and*

$$\sum_{k=0}^{\infty}\frac{\mathbf{E}\left\{|\ \xi_\varepsilon(x)\ |^k\right\}}{k!} < \infty \quad \forall\, x \in \mathcal{D}.$$

Further let $Q_j = \{x_i \in \mathcal{D} : i \in I_j\},\ j \in 1\ldots, p,$ *be the maximally* ε*-adjoining subsets of* $Q = \{x_i \in \mathcal{D} : i \in I\}$*.*
Then the random vectors $(\xi_\varepsilon(x_i,\omega))_{i\in I_j}, j = 1, \ldots, p,$ *are mutually independent.*
Especially $\xi_\varepsilon(x,\omega)$ *and* $\xi_\varepsilon(y,\omega)$ *are independent for* $|\ x - y\ | > \varepsilon$*.*

Proof: not given

The theorem motivates the following definition.

Definition 6.3 *A random function* $\xi_\varepsilon(x,\omega)$ *on* $\mathcal{D} \subset \mathbb{R}^m$ *is called* $\boldsymbol{\varepsilon}$**-dependent** *with the* **dependence length** ε*, if for any* $k = 2, 3\ldots$ *the random vectors* $(\xi_\varepsilon(x_i))_{i\in I_j}$*,* $j = 1, \ldots, p,$ *are mutually independent (where* $I_j, p, \ldots$ *are as above).*

Remark 6.2 If a random function is weakly correlated with correlation length ε then the assumption of the above theorem implies its ε-dependence.
But a (centered) ε-dependent function needs not to be weakly correlated because the moments $\mathbf{E}\left\{\prod_{i\in I}\xi_\varepsilon(x_i)\right\}$ may not exist.

Theorem 6.2 *For any $\varepsilon > 0$ there exists Gaussian random function which is weakly correlated with the correlation length ε.*
A centered Gaussian random function $\xi(x)$ with

$$\mathbf{E}\,\xi(x_1)\xi(x_2) = 0 \quad \textit{for} \quad |x_1 - x_2| > \varepsilon$$

is weakly correlated with a correlation length $\varepsilon' \leq \varepsilon$.

The probabilistic analysis of solutions to equations containing weakly correlated functions often leads to integral functionals of the type

$$\eta_\varepsilon(x) = \int_{\mathcal{G}} F(x,y)\xi_\varepsilon(y)dy$$

with some non-random kernel function F.

Example 6.2 Initial value problem for an ODE with a random inhomogeneous term, $\mathcal{D} = [0,\infty)$, $c > 0$

$$\begin{aligned} f'(x) &= -cf(x) + \xi_\varepsilon(x,\omega), \quad f(0) = 0 \\ \text{solution} \quad f(x,\omega) &= \int_0^x \underbrace{e^{-c(x-y)}}_{F(x,y)}\,\xi_\varepsilon(y,\omega)dy \\ &= \int_{\mathcal{G}} F(x,y)\,\xi_\varepsilon(y,\omega)dy, \qquad \mathcal{G} = \mathcal{G}_x = [0,x] \end{aligned}$$

Under suitable assumptions it can be derived

1.
$$\begin{aligned} \lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon^m}\mathbf{E}\left(\eta_\varepsilon(x_1)\eta_\varepsilon(x)\right) &= \int_{\mathcal{G}} F(x_1,y)F(x_2,y)a(y)dy \\ \text{with the intensity of } \xi_\varepsilon\text{:} \quad a(x) &:= \lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon^m}\int_{\{|y|\leq\varepsilon\}}\mathbf{E}\left(\xi_\varepsilon(x)\xi_\varepsilon(x+y)\right)dy \end{aligned}$$

2. For $\varepsilon \downarrow 0$ the f.d.d.'s of $\frac{1}{\sqrt{\varepsilon^m}}\eta_\varepsilon(x)$ converge weakly to the corresponding f.d.d.'s of a centered Gaussian random function $\zeta(x)$ with

$$\mathbf{E}\,\zeta(x_1)\zeta(x_2) = \int_{\mathcal{G}} F(x_1,y)F(x_2,y)a(y)dy$$

results can be extended to

1. weakly correlated functions with values in $\mathbb{R}^n, \mathbb{C}^n$

2. expansions of moments and distribution densities of integral functionals η_ε in powers of ε, e.g.

$$\mathbf{E}\left\{\eta_\varepsilon(x_1)\eta_\varepsilon(x_2)\right\} = \ldots\varepsilon^m + \ldots\varepsilon^{m+1} + O(\varepsilon^{m+2})$$

   leading expansion terms correspond to the Gaussian limit function while the higher order expansion terms describe deviations from the Gaussian limit function for $\varepsilon > 0$

7 Linear theory of random functions

7.1 Correlation functions

Many practical problems require only the knowledge of second-order moments and not of the complete f.d.d.'s
Consider random functions with values in a linear space with finite second-order moments
recall: Hilbert space of random variables (see Definition 1.9):

$$\begin{aligned}\mathcal{L}_2 = \mathcal{L}_2(\Omega, \mathcal{A}, \mathbf{P}) &:= \{\eta \text{ r.v. on } \{\Omega, \mathcal{A}, \mathbf{P}\} : \ \mathbf{E}\,|\eta|^2 < \infty\} \\ (\xi, \eta) &:= \mathbf{E}\,\xi\overline{\eta}\end{aligned}$$

Definition 7.1 *A random function $\xi(x)$, $x \in \mathcal{X}$, with values in $\mathbb{C}$ is called* **Hilbert random function**, *if $\xi(x) \in \mathcal{L}_2$ $\forall x \in \mathcal{X}$.*

$$\begin{array}{rll} a(x) & := \mathbf{E}\,\xi(x) & \textbf{mean value} \\ R(x,y) & := \mathbf{E}\,(\xi(x) - a(x))\overline{(\xi(y) - a(y))} & \textbf{correlation function} \\ R(x,x) & := \mathbf{E}\,|\xi(x) - a(x)|^2 = \sigma^2(x) & \textbf{variance function} \\ B(x,y) & := \mathbf{E}\,\xi(x)\overline{\xi(y)} & \textbf{covariance function} \end{array}$$

Remark 7.1 Let $x \in \mathcal{X}$ be fixed then $\xi(x) \in \mathcal{L}_2$ is a point in $\mathcal{L}_2$,
let $x \in (a,b)$ then $\xi(x)$ is the parametric equation of a curve in $\mathcal{L}_2$.

Definition 7.2 *A complex valued function $G(x_1, x_2)$, $(x_1, x_2) \in \mathcal{X}^2$ is called* **positive semi-definite kernel** *on $\mathcal{X}^2$, if*

$$\sum_{p,q=1}^{n} G(x_p, x_q) z_p \overline{z_q} \geq 0 \quad \forall\ x_p \in \mathcal{X}, \forall\ z_p \in \mathbb{C},\ p = 1, \ldots, n, \quad \forall\ n \in \mathbb{N}.$$

Theorem 7.1 *The covariance- and correlation functions are positive semi-definite kernels on $\mathcal{X}^2$.*

Proof:

$$\begin{aligned} 0 \leq \quad & \mathbf{E}\,|\sum_{p=1}^{n} \xi(x_p) z_p|^2 &= \sum_{p,q=1}^{n} B(x_p, x_q) z_p \overline{z_q} \\ 0 \leq \quad & \mathbf{E}\,|\sum_{p=1}^{n} \big(\xi(x_p) - \mathbf{E}\,\xi(x_p)\big) z_p|^2 &= \sum_{p,q=1}^{n} R(x_p, x_q) z_p \overline{z_q} \end{aligned}$$

Theorem 7.2 *Positive semi-definite kernels $G(x_1, x_2)$ possess the following properties:*

1. $G(x,x) \geq 0$
2. $G(x_1, x_2) = \overline{G(x_2, x_1)}$
3. $|G(x_1, x_2)|^2 \ \leq\ G(x_1, x_1) G(x_2, x_2)$
4. $|G(x_1, x_3) - G(x_2, x_3)|^2 \ \leq\ G(x_3, x_3)\big[G(x_1, x_1) + G(x_2, x_2) - 2 Re\, G(x_1, x_2)\big]$

Proof:

1. Definition 7.2 with $n = 1$: $\quad G(x_1, x_1) \cdot |z_1|^2 \geq 0 \quad \Rightarrow \quad$ assertion 1.

2. Definition 7.2 with $n = 2$:

$$\underbrace{G(x_1,x_1)\,|z_1|^2}_{\text{real}} + \underbrace{G(x_1,x_2)\,z_1\overline{z_2} + G(x_2,x_1) z_2\overline{z_1}}_{\Rightarrow\ \text{real}} + \underbrace{G(x_2,x_2)\,|z_2|^2}_{\text{real}} \ \geq\ 0$$

$$\Rightarrow \quad G(x_1,x_2)\,z_1\overline{z_2} + G(x_2,x_1)\,z_2\overline{z_1} = \overline{G(x_1,x_2)}\,\overline{z_1} z_2 + \overline{G(x_2,x_1)}\overline{z_2} z_1$$

$$\Rightarrow \quad \underbrace{\left(G(x_1,x_2) - \overline{G(x_2,x_1)}\right)}_{A} z_1\overline{z_2} + \underbrace{\left(G(x_2,x_1) - \overline{G(x_1,x_2)}\right)}_{B} z_2\overline{z_1} = 0$$

$$z_1 = i,\ z_2 = 1 \quad \Rightarrow \quad Ai + B(-i) = 0 \quad \Rightarrow \quad A = B$$

$$\Rightarrow \quad A(z_1\overline{z_2} + z_2\overline{z_1}) = 0 \quad \Rightarrow \quad A = B = 0 \quad \Rightarrow \quad \text{assertion 2.}$$

3. Definition 7.2 with $n = 2$:

$\sum\limits_{p,q=1}^{2} G(x_p,x_q)\, z_p\overline{z_q}$, is a positve semi-definite, Hermitian quadratic form $\Rightarrow$

$$\begin{aligned} \det\begin{pmatrix} G(x_1,x_1) & G(x_1,x_2) \\ G(x_2,x_1) & G(x_2,x_2) \end{pmatrix} &\geq 0 \\ \Rightarrow \quad G(x_1,x_1)G(x_2,x_2) &\geq G(x_1,x_2)G(x_2,x_1) \\ &\overset{2.}{=} G(x_1,x_2)\overline{G(x_1,x_2)} \\ &= |G(x_1,x_2)|^2 \Rightarrow \text{ assertion 3.} \end{aligned}$$

4. Definition 7.2 with $n = 3$: $\quad z_1 = z,\ z_2 = -z$

define: $G_{ij} := G(x_i,x_j), \quad i,j = 1,2,3$

$$\Rightarrow \quad G_{11}|z|^2 - G_{12}|z|^2 + G_{13}z\overline{z_3} - G_{21}|z|^2 + G_{22}|z|^2 - G_{23}z\overline{z_3} + G_{31}z_3\bar{z} - G_{32}z_3\bar{z} + G_{33}|z_3|^2 \ \geq\ 0$$

$$\Rightarrow \quad \underbrace{[G_{11} + G_{22} - 2Re\,G_{12}]}_{a_{11}}\,|z|^2 + \underbrace{G_{33}}_{a_{22}}\,|z_3|^2 + \underbrace{(G_{13} - G_{23})}_{a_{21}}\,z\overline{z_3} + \underbrace{(G_{31} - G_{32})}_{a_{12}}\,\bar{z}z_3 \ \geq\ 0$$

with $y_1 := z, y_2 := z_3$ it follows $\sum\limits_{ij=1}^{2} a_{ij}y_i\overline{y_j} \geq 0$

$$\Rightarrow \qquad \det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \geq 0$$

$$\text{hence} \quad \det\begin{pmatrix} G_{11} + G_{22} - 2\text{Re}G_{12} & \overline{G_{13} - G_{23}} \\ G_{13} - G_{23} & G_{33} \end{pmatrix} \geq 0$$

$\Rightarrow$ assertion 4.

Definition 7.3

1. *Let $\xi_1(x), \xi_2(x)$ be Hilbert random functions with $\mathbf{E}\,\xi_p(x) = a_p(x), p = 1,2$. Then*

$$R_{\xi_1\xi_2}(x,y) \ = \ \mathbf{E}\,(\xi_1(x) - a_1(x))\overline{(\xi_2(y) - a_2(y))}$$

is called **cross correlation function** *of ξ_1 and ξ_2.*

2. *Let $\xi(x) = (\xi_1(x), \dots, \xi_m(x))^\tau$ be a vector-valued random function such that $\xi(x) \in \mathcal{L}_2^m, \forall x \in \mathcal{X}$, where $\mathcal{L}_2^m$ is the Hilbert space of random vectors with values in $\mathbb{C}^m$. For $z = (z_1, \dots, z_m)^\tau \in \mathbb{C}^m$ the adjoint vector is denoted by $z^* = \overline{z}^\tau = (\overline{z_1}, \dots, \overline{z_m})$.*

   Then $a(x) = \mathbf{E}\,\xi(x) = (\mathbf{E}\,\xi_1(x), \dots, \mathbf{E}\,\xi_m(x))^\tau$ is called **mean value** *and*

$$\begin{aligned} R(x,y) &= \mathbf{E}\,(\xi(x) - a(x))(\xi(y) - a(y)^* = (R_{pq}(x,y))_{1\le p,q\le m} \\ \text{where} \quad R_{pq}(x,y) &= \mathbf{E}\,(\xi_p(x) - a_p(x))(\overline{\xi_q(y) - a_q(y)}) \end{aligned}$$

   is called **matrix correlation function** *of ξ.*

Remark 7.2 If in 1. it holds $\xi_1 = \xi_2 = \xi$ then $R_{\xi\xi}(x_1, x_2)$ sometimes is called **autocorrelation function**. The matrix correlation function $R(x,y)$ in 2. contains the autocorrelation functions of ξ_k as diagonal entries and the cross correlation functions of ξ_k and ξ_l as non diagonal entries, $k, l = 1, \dots, m, k \neq l$.

Definition 7.4 *A matrix function $G(x,y) = (G_{pq}(x,y))_{1\le p,q\le m}$ is called a* **positive semi-definite matrix kernel** *on $\mathcal{X}^2$, if*

$$\sum_{p,q=1}^{n} z_p^* G(x_p, x_q) z_q \ \geq \ 0, \quad \forall\ x_p \in \mathcal{X},\ \forall z_p \in \mathbb{C}^m, p = 1, \dots, n,\ \forall n \in \mathbb{N}$$

where $z_p^* G(x_p, x_q) z_q = \sum\limits_{j,k=1}^{m} \overline{z_{pj}} G_{jk}(x_p, x_q) z_{qk}$

Theorem 7.3 *The matrix correlation function is a positive semi-definite matrix kernel.*

Proof:

$$\begin{aligned} 0 &\leq \mathbf{E}\left|\sum_{p=1}^{n} z_p^*(\xi(x_p) - a(x_p))\right|^2 \\ &= \mathbf{E}\left\{\sum_{p,q=1}^{n} z_p^*(\xi(x_p) - a(x_p))(\xi(x_q) - a(x_q))^* z_q\right\} \\ &= \sum_{p,q=1}^{n} z_p^* R(x_p, x_q) z_q \Rightarrow \quad \text{assertion} \end{aligned}$$

Theorem 7.4 *There hold the following properties of positive semi-definite matrix kernels $G(x,y)$:*

1. *the matrix $G(x,x)$ is positive semi-definite $\forall\ x \in \mathcal{X}$, i.e.*
$$z^* G(x,x) z = \sum_{k,l=1}^{m} G_{kl}(x,x) \bar{z}_k z_l \geq 0 \quad \forall \quad z = (z_1, \dots, z_m)^\tau \in C^m,$$
2. $G_{kl}(x,y) = \overline{G_{lk}(y,x)}$, *i.e.* $G(x,y) = G^*(y,x) = \overline{G^\tau(y,x)}$
3. $|G_{kl}(x,y)|^2 \ \leq \ G_{kk}(x,x) G_{ll}(y,y)$.

Proof:

1. Definition 7.4 with $n = 1$: $\qquad z_1^* G(x_1, x_1) z_1 \geq 0 \qquad \Rightarrow$ assertion 1.
2. Definition 7.4 with $n = 2$:

$$\Rightarrow \quad \underbrace{z_1^* G(x_1, x_1) z_1}_{\geq 0} + \underbrace{z_2^* G(x_2, x_2) z_2}_{\geq 0} + \underbrace{z_2^* G(x_2, x_1) z_1 + z_1^* G(x_1, x_2) z_2}_{\text{real}} \geq 0$$

$$\Downarrow$$

$$z_2^* G(x_2, x_1) z_1 + z_1^* G(x_1, x_2) z_2 = z_1^\tau \overline{G(x_1, x_2)}\,\overline{z_2} + z_2^\tau \overline{G(x_2, x_1)}\,\overline{z_1}$$

$z_1 = (i, 0, \ldots, 0)^\tau, \; z_2 = (1, 0, \ldots, 0)^\tau$

$$\Rightarrow \quad -G_{11}(x_1, x_2) + G_{11}(x_2, x_1) = \overline{G_{11}(x_1, x_2)} - \overline{G_{11}(x_2, x_1)} \quad (1)$$

$z_1 = (1, 0, \ldots, 0)^\tau, \; z_2 = (1, 0, \ldots, 0)^\tau$

$$\Rightarrow \quad G_{11}(x_1, x_2) + G_{11}(x_2, x_1) = \overline{G_{11}(x_1, x_2)} + \overline{G_{11}(x_2, x_1)} \quad (2)$$

$$\stackrel{(1)}{\Rightarrow} \quad \overline{G_{11}(x_1, x_2)} - G_{11}(x_2, x_1) = -(G_{11}(x_1, x_2) - \overline{G_{11}(x_2, x_1)})$$

$$\stackrel{(2)}{\Rightarrow} \quad \overline{G_{11}(x_1, x_2)} - G_{11}(x_2, x_1) = G_{11}(x_1, x_2) - \overline{G_{11}(x_2, x_1)}$$

$$\Rightarrow \quad \overline{G_{11}(x_2, x_1)} = G_{11}(x_1, x_2) \qquad \Rightarrow \qquad \overline{G_{kk}(x_1, x_2)} = G_{kk}(x_2, x_1)$$

$$\left.\begin{array}{ll} z_1 = (i, 0, \ldots, 0); & z_2 = (0, 1, 0, \ldots, 0) \\ z_1 = (1, 0, \ldots, 0); & z_2 = (0, 1, 0, \ldots, 0) \end{array}\right\} \quad \Rightarrow \quad G_{12}(x_1, x_2) = \overline{G_{21}(x_2, x_1)}$$

$\Rightarrow$ assertion 2.

3. Definition 7.4 with $n = 2$:

$$\Rightarrow \quad z_1^* G(x_1, x_1) z_1 + z_2^* G(x_2, x_2) z_2 + z_1^* G(x_1, x_2) z_2 + z_2^* G(x_2, x_1) z_1 \geq 0$$

$z_1 = (a, 0, \ldots, 0)^\tau, \; z_2 = (0, b, 0, \ldots, 0)^\tau$

$$\Rightarrow \quad G_{11}(x_1, x_1)|a|^2 + G_{22}(x_2, x_2)|b|^2 + G_{12}(x_1, x_2)\bar{a}b + G_{21}(x_2, x_1)a\bar{b} \geq 0$$

$$\Rightarrow \quad \begin{vmatrix} G_{11}(x_1, x_1) & G_{12}(x_1, x_2) \\ G_{21}(x_2, x_1) & G_{22}(x_2, x_2) \end{vmatrix} \geq 0$$

$$\Rightarrow \quad G_{11}(x_1, x_1) G_{22}(x_2, x_2) \geq G_{12}(x_1, x_2) G_{21}(x_2, x_1)$$

$$= G_{12}(x_1, x_2)\overline{G_{12}(x_1, x_2)} = |G_{12}(x_1, x_2)|^2$$

analogously for arbitrary $k, l \in \{1, \ldots, m\}$ $\qquad \Rightarrow$ assertion 3.

Theorem 7.5 *A function $R(x, y),\; x, y \in \mathcal{X}$ is a (matrix) correlation function if and only if it is a positive semi-definite (matrix) kernel.*

Proof:

necessity: follows from Theorem 7.1 and Theorem 7.3

sufficiency: let $R(x, y)$ be positve semi-definite, Theorems 3.3, 3.4, 3.8 $\Rightarrow$
$\exists$ Gaussian random function with R as correlation function

assume now that $\mathcal{X}$ is a metric space with the metric ρ

Definition 7.5 *A Hilbert random function $\xi(x),\; x \in \mathcal{X}$ is called* **mean-square (m.-s.) continuous** *in $x_0 \in \mathcal{X}$, if*

$$\mathbf{E}\, |\xi(x) - \xi(x_0)|^2 \;\to\; 0 \quad \textit{for} \quad \rho(x, x_0) \to 0.$$

Remark 7.3 Because of Chebychev - inequality:

$$\mathbf{P}(|\xi(x)-\xi(x_0)|>\varepsilon) \leq \tfrac{1}{\varepsilon^2}\mathbf{E}\,|\xi(x)-\xi(x_0)|^2$$

m.-s. continuity in x_0 implies that ξ is stochastically continuous in x_0 (see Definition 4.3).

Remark 7.4 The m.-s. continuity on $\mathcal{X}$ (i.e. for all points in $\mathcal{X}$) does not imply the continuity of the paths.
see: let $\xi(t)$ be a Poisson process (see Definition 4.7)
$\Rightarrow\ \xi(t+h)-\xi(t),\ h>0$, is Poisson distributed with the parameter $qh, q>0$

$$\begin{aligned}\Rightarrow\quad P\big(\xi(t+h)-\xi(t)=k\big) &= \frac{1}{k!}(hq)^k e^{-hq}\\ \mathbf{E}\,|\xi(t+h)-\xi(t)|^2 &= \sum_{k=0}^{\infty} k^2\frac{1}{k!}(hq)^k e^{-hq} = hq+(hq)^2 \quad \overset{h\to 0}{\longrightarrow} \quad 0\end{aligned}$$

$\Rightarrow\quad \xi(t)$ is m.-s. continuous $\forall\ \ t\in[0,\infty)$
but values of $\xi(t)$ are integers $\quad\Rightarrow$ paths are not continuous

Theorem 7.6 *A Hilbert random function $\xi(x)$ is m.-s.-continuous in $x_0\in\mathcal{X}$ if and only if the covariance function $B(x,y)=\mathbf{E}\,\xi(x)\overline{\xi(y)}$ is continuous at the point (x_0,x_0).*

Proof: apply Theorem 1.9 with $\psi(x)=\rho(x,x_0)$

Stationary random processes

Definition 7.6 *A random process $\xi(t),\ t\in\mathcal{X}\subset\mathbb{R}^1$ is called* **stationary (in the strict sense)**, *if the joint distribution of*

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

does not depend on τ, $\forall\ n\in\mathbb{N},\ \forall\ t_1,\ldots,t_n\in\mathcal{X}$ and $\forall\ \tau\in\mathbb{R}^1$ with $\tau+t_k\in\mathcal{X}$ (i.e. f.d.d.'s of ξ are invariant w.r.t. time shifts)

let $\xi(t)$ be a stationary Hilbert random function on $\mathcal{X}=(-\infty,+\infty)$ then

$$\begin{aligned} a(t) &= \mathbf{E}\,\xi(t)=a=\text{const}\ .\\ R(t_1+\tau,t_2+\tau) &= \mathbf{E}\,(\xi(t_1+\tau)-a)(\overline{\xi(t_2+\tau)-a})\ \ \text{is independent of }\tau\\ \Rightarrow\quad R(t_1,t_2) &= R((t_1-t_2)+t_2,0+t_2)=R(t_1-t_2,0)=R(t_1-t_2)\end{aligned}$$

Definition 7.7 *A mean-square continuous Hilbert random process $\xi(t)$ on $\mathcal{X}=(-\infty,\infty)$ with values in $\mathbb{C}^m$ is called* **wide-sense stationary** *or* **weakly stationary**, *if*

1. $\mathbf{E}\,\xi(t)=a=\text{const}$
2. $\mathbf{E}\,(\xi(t_1)-a)(\xi(t_2)-a)^*=R(t_1-t_2)$.

Definition 7.8 *Let $\mathcal{X}$ be a linear space. A complex-valued function $f(x)$, $x \in \mathcal{X}$, is called* **positive semi-definite**, *if*

$$\sum_{p,q=1}^{n} f(x_p - x_q)\overline{z_p} z_q \geq 0, \quad \forall\, n \in \mathbb{N},\ x_p \in \mathcal{X},\ z_p \in \mathbb{C},\, p = 1, \ldots, n.$$

Example 7.1 The correlation function of a wide-sense stationary random process is positve semi-definite.

Theorem 7.7 *Let f be positive semi-definite, then it holds*

1. $f(0) \geq 0$
2. $\overline{f(x)} = f(-x)$
3. $|f(x)| \leq f(0)$ *(i.e. f is bounded)*
4. $|f(x_1) - f(x_2)| \leq 2f(0)[f(0) - \operatorname{Re} f(x_2 - x_1)]$
   (i.e. if f is conitinuous at 0 then it is uniformly conitinuous on $\mathcal{X}$.

Proof: apply Theorem 7.2 with $G(x_1, x_2) = f(x_1 - x_2)$

Consider vector-valued wide-sense stationary processes $\xi(t) \in \mathbb{C}^m$, Def. 7.7 $\Rightarrow$

$$\begin{aligned} \mathbf{E}\, \xi_k(t) &= a_k = \text{const} \\ \mathbf{E}\, (\xi_k(t_1) - a_k)\overline{(\xi_l(t_2) - a_l)} &= R_{kl}(t_1, t_2) = R_{kl}(t_1 - t_2) \end{aligned}$$

i.e. the components of ξ are also wide-sense stationary $(k = l)$

Definition 7.9 *If $\xi_1(t)$ and $\xi_2(t)$ are wide-sense stationary and the vector process $(\xi_1(t), \xi_2(t))$ is also wide-sense stationary then the processes ξ_1 and ξ_2 are called* **wide-sense stationarily related** *or* **stationarily correlated**.

Example 7.2 *Let $\eta(t)$ be centered wide-sense stationary, then $\xi_1(t) = \eta(t)$ and $\xi_2(t) = \eta(2t)$ are wide-sense stationary but not stationarily related since*

$$R_{12}(t_1, t_2) = \mathbf{E}\, \xi_1(t_1)\overline{\xi_2(t_2)} = \mathbf{E}\, \eta(t_1)\overline{\eta(2t_2)} = R_{\eta\eta}(t_1 - 2t_2).$$

Extension to positive semi-definite complex-valued matrix functions:

Definition 7.10 *Let $\mathcal{X}$ be a linear space. The* **matrix function** $G(x) = (G_{jk}(x))_{1 \leq j,k \leq m}$, $x \in \mathcal{X}$ *is called* **positive semi-definite**,*if*

$$\sum_{p,q=1}^{n} z_p^* G(x_p - x_q) z_q \ \geq\ 0, \qquad \forall\, n \in \mathbb{N},\ \forall x_p \in \mathcal{X}, \forall\ z_p \in \mathbb{C}^m,\ p = 1, \ldots, n.$$

Example 7.3 Matrix correlation functions of wide-sense stationary vector processes are positive semi-definite.

Theorem 7.8 *Let G be a positive semi-definite matrix function, then it holds*

1. $G(0)$ *is a positive semi-definite matrix,*
2. $G_{kl}(x) = \overline{G_{kl}(-x)}$, *i.e.* $G(x) = G^*(-x)$

3. $|G_{kl}(x)|^2 \ \le\ G_{kk}(0)G_{ll}(0).$

Proof: apply Theorem 7.4 with $G(x_1, x_2) = G(x_1 - x_2)$

Examples of wide-sense stationary random processes

Example 7.4 Standardized non-correlated sequence, $\mathcal{X} = \mathbb{Z}$:

$$\{\xi(n),\ n = 0, \pm 1, \dots\}, \quad \xi(n) \in \mathbb{C}^m$$

with $\quad a = \mathbf{E}\,\xi(n) = 0 \qquad$ and $\qquad R(n) = \mathbf{E}\,\xi(k+n)\xi(k)^* = \begin{cases} I & n = 0 \\ 0 & n \neq 0 \end{cases}$;

(i.e. $R_{ij}(n) = \mathbf{E}\,\xi_i(k+n)\overline{\xi_j(k)} = \delta_{ij}\,\delta_{0n}$)

Example 7.5 Process of moving average

Let $\{\xi(n),\ n = 0, \pm 1, \dots\}$ as above and $\{A_n, n = 0, 1, \dots\}$ with $A_n \in \mathbb{C}^{m\times m}$ be a sequence of matrices:

$$\zeta(n) := \sum_{k=0}^{\infty} A_k \xi(n-k)$$

- $\{\eta_k = A_k\xi(n-k),\ k = 0, 1, \dots\}$ are orthogonal vectors in $\mathcal{L}_2^m(\Omega, \mathcal{A}, \mathbf{P})$

  i.e. $\qquad \mathbf{E}\,\eta_k^*\eta_l = 0 \qquad$ for $\quad k \neq l$ since

$$\begin{aligned}
\mathbf{E}\,\eta_k^*\eta_l &= \mathbf{E}\,\xi^*(n-k)\underbrace{A_k A_l}_{B_{kl}}\xi(n-l) \\
&= \sum_{p,q=1}^{m} B_{kl,pq}\underbrace{\mathbf{E}\,\overline{\xi}_p(n-k)\xi_q(n-l)}_{\delta pq\cdot\delta_{kl}} \\
&= \begin{cases} \sum\limits_{p=1}^{m} B_{kk,pp} = \mathrm{tr}[A_k^*A_k] = ||A_k||^2 & \text{for } k = l \\ 0 & \text{for } k \neq l \end{cases} \qquad (*)
\end{aligned}$$

where $||A||^2 \ = \ \mathrm{tr}[A^*A] = \sum\limits_{p,q=1}^{m} A_{pq}\overline{A}_{pq} = \sum\limits_{p,q=1}^{m} |A_{pq}|^2 \quad$ (Frobenius norm)

- series converges in $\mathcal{L}_2^m$ if $\ \sum\limits_{k=0}^{\infty} \|A_k\|^2 < \infty$

  see $\quad \|\zeta(n)\|^2 = \mathbf{E}\,\zeta(n)^*\zeta(n) = \mathbf{E}\left(\sum\limits_{k=0}^{\infty}\eta_k^*\right)\left(\sum\limits_{l=0}^{\infty}\eta_l\right)$

$$= \sum_{k,l=0}^{\infty} \mathbf{E}\,\eta_k^*\eta_l \ \overset{(*)}{=}\ \sum_{k=0}^{\infty} \|A_k\|^2 < \infty$$

- $\zeta(n)\quad$ is wide-sense stationary

  see $\ \mathbf{E}\,\zeta(n+p)\zeta^*(n) = \mathbf{E}\left(\sum\limits_{k=0}^{\infty} A_k\xi(n+p-k)\right)\left(\sum\limits_{l=0}^{\infty}\xi^*(n-l)\,A_l^*\right)$

$$\begin{aligned}
&= \sum_{k,l=0}^{\infty} A_k \underbrace{\mathbf{E}\,\xi(n+p-k)\xi^*(n-l)}_{\delta_{n+p-k,n-l}\cdot I = \delta_{p+l,k}\cdot I} A_l^* \\
&= \sum_{l=0}^{\infty} A_{p+l}A_l^* \quad =: \quad R(p)
\end{aligned}$$

Remark 7.5 generalization of example 7.4 to continuos time processes:
let $\xi(t), t \in \mathbb{R}$, be a process with $\mathbf{E}\,\xi(t) = 0$, $\mathbf{E}\,|\xi(t)|^2 = 1$, $\mathbf{E}\,\xi(t)\overline{\xi(s)} = 0$ for $t \neq s$
ξ is not m.-s.-continuous since its covariance function
$B(t,s) = \begin{cases} 0 & t \neq s \\ 1 & t = s \end{cases}$ is not continuous at (t,t) (see Theorem 7.6).

$\Rightarrow$ ξ is called a process of white noise type

Example 7.6 Random oscillations

$$\xi(t) = \sum_k \gamma_k e^{iu_k t} = \sum_k \gamma_k(\cos(i_k t) + i\sin(u_k t))$$

$\gamma_k e^{iu_k t}$: harmonic oscillations with frequency $\frac{u_k}{2\pi}$ and power $|\gamma_k|^2$
$\{\gamma_k\}$: orthogonal centered r.v.; i.e. $\mathbf{E}\,\gamma_k = 0$; $\mathbf{E}\,\gamma_k\overline{\gamma_j} = \delta_{kj}c_k^2$
$\{u_k\}$: (frequency) spectrum of the process $\xi(t)$

$$\begin{aligned} \Rightarrow \quad R(t_1,t_2) &= \mathbf{E}\,\xi(t_1)\overline{\xi(t_2)} = \mathbf{E}\sum_{j,k}\gamma_j\overline{\gamma_k}\exp(iu_j t_1 - iu_k t_2) \\ &= \sum_k c_k^2 \exp(iu_k(t_1-t_2)) =: R(t_1-t_2) \end{aligned}$$

$\Rightarrow$ $\xi(t)$ is a wide-sense stationary process

motivates the definition of a new characteristic of a random process:
$F(u) := \sum\limits_{\substack{k \\ u_k<u}} c_k^2$: "spectral function" of ξ;
average power carried by the harmonics of $\xi(t)$ with frequencies smaller than u
$F(u)$ completely characterizes the average powers $\mathbf{E}\,|\gamma_k|^2 = c_k^2$
since $c_k^2 = F(u_k+0) - F(u_k)$; $\sum\limits_{u_1 \le u < u_2} c_k^2 = F(u_2) - F(u_1)$
correlation function of ξ can be written in terms of the spectral function:

$$R(t) = \sum_k c_k^2\, e^{iu_k t} = \underbrace{\int_{-\infty}^{\infty} e^{itu} F(du)}_{\text{Lebesgue-Stieltjes integral}}$$

7.2 Spectral representation of correlation functions

Definition 7.11 *A random function $\{\xi(x),\ x \in \mathbb{R}^m\}$ with values in $\mathbb{C}$ is called* **homogeneous**, *if*

1. $\mathbf{E}\,\xi(x) = a = $const
2. $R(x_1, x_2) = \mathbf{E}\,(\xi(x_1) - a)(\overline{\xi(x_2) - a}) = R(x_1 - x_2)$.

The correlation function of a homogeneous random function depends only on the vector $x_1 - x_2$ and it holds:

$$\sum_{j,k=1}^{n} R(x_j - x_k)\bar{z}_j z_k \geq 0 \quad \forall\ n,\ \ \forall\ x_j \in \mathbb{R}^m,\ \ \forall\ z_j \in \mathbb{C},\ j = 1, \ldots, n.$$

Theorem 7.9 *Let ξ be a homogeneous random function on $\mathbb{R}^m$ with the correlation function $R(x)$. If $R(x)$ is continuous at $x = 0$ then $\xi(x)$ is mean-square continuous for all $x \in \mathbb{R}^m$.*

Proof:
$$\begin{aligned}\mathbf{E}\,|\xi(x+h) - \xi(x)|^2 &= \mathbf{E}\,(\xi(x+h) - \xi(x))(\overline{\xi(x+h)} - \overline{\xi(x)}) \\ &= R(0) - R(h) - \underbrace{R(-h)}_{=\overline{R(h)}} + R(0) \\ &= 2(R(0) - \mathrm{Re}\, R(h)) \qquad \Rightarrow \quad \text{assertion}\end{aligned}$$

Theorem 7.10 *The function $R(x)$, $x \in \mathbb{R}^m$, is the correlation function of a mean-square continuous homogeneous random function $\{\xi(x),\ x \in \mathbb{R}^m\}$ if and only if $R(x)$ admits a representation*

$$R(x) = \int_{\mathbb{R}^m} e^{i(x,u)} F(du) \qquad (*)$$

where F is a finite measure on the borel sets $\mathcal{B}^m$ of $\mathbb{R}^m$. Moreover the measure F is uniquely determined on $\mathcal{B}^m$.

Proof:
sufficiency ($\Leftarrow$): let $R(x)$ be given by $(*)$, then

- $R(x)$ is positive semi-definite, since

$$\begin{aligned}\sum_{j,k=1}^{n} R(x_j - x_k)\bar{z}_j z_k &= \int_{\mathbb{R}^m} \sum_{j,k=1}^{n} e^{i(x_j - x_k, u)} \bar{z}_j z_k F(du) \\ &= \int_{\mathbb{R}^m} \left| \sum_{k=1}^{n} e^{-i(x_k,u)} z_k \right|^2 F(du) \ \geq\ 0\end{aligned}$$

- $R(x)$ is continuous at $x = 0$, since
$$\lim_{h\to 0} R(h) = \lim_{h \to 0} \int_{\mathbb{R}^m} e^{i(h,u)} F(du) = \int_{\mathbb{R}^m} 1 \cdot F(du) = R(0)$$
follows from Lebesgues theorem on dominated convergence (see basic course Theorem 3.12, $|e^{i(h,u)}| \leq 1 = |e^{i(0,u)}|$ and F is a finite measure)

Theorem 3.8 $\Rightarrow$ $\exists$ complex-valued Gaussian random function $\xi(x)$ possessing $R(x)$ as correletion function

ξ is m.-s.-continuous since $R(x)$ is continuous at $x = 0$ (see Theorem 7.9)

neccessity ($\Rightarrow$): follows from the next Theorem 7.11

if $\xi(x)$ is a m.-s.-continuous homogeneous random function then its correlation function $R(x)$ is positive semi-definite and continuous at $x = 0$

$\Rightarrow$ $R(x)$ is continuous $\forall\, x$(exercise) Theorem 7.11 $\Rightarrow$ representation $(*)$

Remark 7.6 A simple example of a homogeneous random function possessing a correlation function given in $(*)$ is the following:
let η be a random vector in $\mathbb{R}^m$ with the distribution

$$\mathbf{P}(\eta \in A) = \frac{1}{F_0} F(A), \quad \forall\, A \in \mathcal{B}^m, \quad F_0 := F(\mathbb{R}^m),$$

and ϕ be r.v. which is uniformly distributed on $[-\pi, \pi]$ and independent of η, set

$$\xi(x) := \sqrt{F_0}\; e^{i\,[(\eta,x)+\phi\,]}.$$

Then it follows

$$\begin{aligned} \mathbf{E}\,\xi(x) &= \sqrt{F_0}\mathbf{E}\,e^{i(\eta,x)}\mathbf{E}\,e^{i\phi} = 0 \quad \text{since} \quad \mathbf{E}\,e^{i\phi} = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{iv}dv = 0 \\ R(x,y) &= \mathbf{E}\,\xi(x)\overline{\xi(y)} = F_0\mathbf{E}\,e^{i(\eta,x-y)} = \int_{\mathbb{R}^m} e^{i(u,x-y)}F(du) = R(x-y) \end{aligned}$$

$\Rightarrow$ assertion

Theorem 7.11 *Every continuous and positive semi-definite function R on $\mathbb{R}^m$ admits a representation $(*)$ (with a uniquely determined F).*

Proof: idea: construct the measure F using the Fourier transform of

$$R_N(x) := R(x)\underbrace{\exp\left(-\frac{|x|^2}{2N}\right)}_{\to 1\ \forall x},$$

taking the limit $N \to \infty$
remarks:

- $R(x)$ may be not absolutely integrable, but for its "correction" $R_N(x)$ it holds:
$$\int_{-\infty}^{\infty} |R_N(x)|dx \le R(0)\int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2N}\right)dx < \infty$$
- Fourier transform of $f(x)$: $\tilde{f}(z) := \frac{1}{(\sqrt{2\pi})^m}\int_{\mathbb{R}^m} f(x)e^{-i(x,z)}dx$

  inversion formula: $f(x) = \frac{1}{(\sqrt{2\pi})^m}\int_{\mathbb{R}^m} \tilde{f}(z)e^{i(x,z)}dz$

consider the Fourier transform of $R_N(x)$:

$$\tilde{R}_N(z) = \frac{1}{(\sqrt{2\pi})^m} \int\limits_{\mathbb{R}^m} R(x) \exp\left(-\frac{|x|^2}{2N} - i(x,z)\right) dx$$

and prove (see below) (1): $\tilde{R}_N(z) \geq 0$
(2): $\tilde{R}_N(z)$ is integrable
(3): $\tilde{R}_N(z)$ is differentiable

then it folllows

- (1),(2) $\Rightarrow$ $F_N(A) := \frac{1}{(\sqrt{2\pi})^m} \int\limits_A \tilde{R}_N(z)dz$, $A \in \mathcal{B}^m$ is a finite measure on $(\mathbb{R}^m, \mathcal{B}^m)$
- $F_N^*(A) := \frac{F_N(A)}{F_N(\mathbb{R}^m)}$ is a probability measure with the characteristic function:

$$\begin{aligned} J_N(u) &= \int\limits_{\mathbb{R}^m} e^{i(u,z)} F_N^*(dz) = \int\limits_{\mathbb{R}^m} e^{i(u,z)} \frac{1}{F_N(\mathbb{R}^m)} \frac{1}{(\sqrt{2\pi})^m} \tilde{R}_N(z)dz \\ &= \underbrace{\frac{1}{F_N(\mathbb{R}^m)}}_{\frac{1}{R(0)}} \underbrace{\frac{1}{(\sqrt{2\pi})^m} \int\limits_{\mathbb{R}^m} e^{i(u,z)} \tilde{R}_N(z)dz}_{R_N(u)=R(u)\exp\left(-\frac{|u|^2}{2N}\right)} \end{aligned}$$

(because of (2) and (3) the inversion formula can be applied)

$$F_N(\mathbb{R}^m) = \frac{1}{(\sqrt{2\pi})^m} \int\limits_{\mathbb{R}^m} \tilde{R}_N(z) \underbrace{e^{i(0,z)}}_{=1} dz \;=\; R_N(0) \;=\; R(0)$$

- $\Rightarrow$ $J_N(u) = \frac{1}{R(0)} \cdot R(u) \cdot e^{-\frac{|u|^2}{2N}}$
- $\Rightarrow$ $\forall\, u \in \mathbb{R}^m$ it holds: $J_N(u) = \frac{R(u)}{R(0)} e^{\frac{-|u|^2}{2N}} \xrightarrow{N\to\infty} \frac{R(u)}{R(0)} =: J(u);$ $J(u)$ is continuous at $u = 0$;
- Theorem 1.15 $\Rightarrow$ $J(u)$ is the characteristic function of a probability measure F^*, moreover F_N^* converges weakly to F^*
$\Rightarrow$ $\exists$ finite measure $F = R(0)F^*$
- Theorem 1.14 $\Rightarrow$ measure F^* is uniquely determinded (by $J(u)$)

proof of the assertions (1), (2), (3):

(1) to prove: $\tilde{R}_N(z) \;\geq\; 0$

$R(x)$ is positive semi-definite $\Rightarrow$ $\int\limits_{\mathbb{R}^m} \int\limits_{\mathbb{R}^m} R(x-y)\overline{g(x)}g(y)dxdy \geq 0$

for any integrable function $g(x)$ on $\mathbb{R}^m$, since the defining property of positive semi-definiteness implies the assertion for the integral sums, moreover $R(x)$ is bounded choose $g(x) = \exp(-\frac{|x|^2}{N} + i(x,z)), \quad N > 0, \quad z \in \mathbb{R}^m$

$$\Rightarrow \quad 0 \le \int_{\mathbb{R}^m}\int_{\mathbb{R}^m} R(x-y)\exp(-\tfrac{|x|^2+|y|^2}{N} - i(x-y,z))dxdy$$

$$= \int_{\mathbb{R}^m}\int_{\mathbb{R}^m} R(\sqrt{2}\,u)\exp(-\tfrac{|u|^2+|v|^2}{N} - i(\sqrt{2}u,z))dudv$$

see $x - y = \sqrt{2}\,u,\ x + y = \sqrt{2}\,v,\ x = \frac{1}{\sqrt{2}}(u+v),\ y = \frac{1}{\sqrt{2}}(v-u),$

$|x|^2 + |y|^2 = |u|^2 + |v|^2; \quad |\frac{\partial(x_1,\ldots,x_m,y_1,\ldots,y_m)}{\partial(u_1,\ldots,u_m,v_1,\ldots,v_m)}| = 1$

$$= \int_{\mathbb{R}^m} R(\sqrt{2}\,u)\exp(-\tfrac{|u|^2}{N} - i(\sqrt{2}\,u,z))du \int_{\mathbb{R}^m}\exp(-\tfrac{|v|^2}{N})dv$$

vgl. $\int_{\mathbb{R}^m}\exp(-\frac{|v|^2}{N})dv = (\int_{-\infty}^{\infty}\exp(-\frac{s^2}{N})ds)^m = \sqrt{\pi N}^{\,m}$

$$= \sqrt{\pi N}^{\,m}\int_{\mathbb{R}^m} R(\sqrt{2}\,u)\exp(-\tfrac{|u|^2}{N} - i(\sqrt{2}\,u,z))du \qquad (w = \sqrt{2}u)$$

$$= \sqrt{\pi N}^{\,m}\,\frac{1}{\sqrt{2}^m}\sqrt{2\pi}^{\,m}\underbrace{\frac{1}{\sqrt{2\pi}^{\,m}}\int_{\mathbb{R}^m} R(w)\exp(-\frac{|w|^2}{2N} - i(w,z))dw}_{\tilde{R}_N(z)}$$

$$= \sqrt{\pi N}^{\,m}\sqrt{\pi}^{\,m}\tilde{R}_N(z)$$

(2) to prove : $\tilde{R}_N(z)$ is integrable

Parseval equation for Fourier integrals: $\quad (\tilde{f},\tilde{g})_{L_2} = (f,g)_{L_2}$

$f(x) = R_N(x), \qquad \tilde{f}(z) = \tilde{R}_N(z);$

$g(x) = \sqrt{\varepsilon}^{\,-m}\exp(-\frac{|x|^2}{2\varepsilon}), \quad \tilde{g}(z) = \exp(-\frac{1}{2}\varepsilon|z|^2)$

$$\Rightarrow \quad |(\tilde{f},\tilde{g})| = |\int_{\mathbb{R}^m}\tilde{R}_N(z)\exp(-\tfrac{1}{2}\varepsilon|z|^2)dz|$$

$$= |(f,g)| = |\int_{\mathbb{R}^m} R(x)\exp(-\tfrac{|x|^2}{2N} - \tfrac{|x|^2}{2\varepsilon})\sqrt{\varepsilon}^{\,-m}dx|$$

$$\le R(0)\frac{1}{\sqrt{\varepsilon}^{\,m}}\int_{\mathbb{R}^m}\exp(-\tfrac{|x|^2}{2\varepsilon})dx \ = \ \sqrt{2\pi}^{\,m}R(0)$$

for $\varepsilon \to 0$ follows from Fatou's Theorem:

$$|\int_{\mathbb{R}^m}\tilde{R}_N(z)dz| \le \sqrt{2\pi}^{\,m}R(0) \qquad \Rightarrow \quad \text{integrability}$$

recall Fatou: $\quad f_n \to f$ f"ur $f_n \ge 0, \quad \int f_n(x)d\mu \le K$

$\Rightarrow \quad f$ ist integrable and $\int f(x)d\mu \le K;$

(3) to prove : $\tilde{R}_N(z)$ is differentiable

holds, since $\tilde{R}_N(z)$ is the Fourier transform of the continuous function $R_N(x)$ with $|\int_{\mathbb{R}^m}|x|^2R_N(x)dx| < \infty$:

see $m = 1$: $\quad \tilde{R}_N(z) = \int_{-\infty}^{\infty}H(x)e^{-ixz}dx \qquad$ with $\quad H(x) := \frac{1}{\sqrt{2\pi}}R_N(x)$

$$\Rightarrow \quad \left|\frac{1}{h}(\tilde{R}_N(z+h) - \tilde{R}_N(z)) - \int_{-\infty}^{\infty}H(x)(-ix)e^{-ixz}dx\right|$$

$$= \left|\int_{-\infty}^{\infty}H(x)e^{-ixz}(\tfrac{e^{-ihx}-1}{h} + ix)dx\right|$$

see $\frac{1}{h}(e^{-ihx}-1) = -ix - \frac{1}{2}hx^2 e^{-\vartheta ixh}$

$$
\begin{aligned}
&= \left| \int_{-\infty}^{\infty} H(x) e^{-ixz} \tfrac{1}{2} h x^2 e^{-\vartheta ixh} dx \right| \\
&\leq \tfrac{h}{2} \int_{-\infty}^{\infty} |H(x)| \cdot x^2 \cdot 1 \; dx \\
&= \tfrac{h}{2\sqrt{2\pi}} \int_{-\infty}^{\infty} |R(x)| e^{-\frac{x^2}{2N}} x^2 dx \quad \leq \quad \tfrac{h}{2\sqrt{2\pi}} |R(0)| \int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2N}} dx \quad \xrightarrow{h\to 0} \; 0
\end{aligned}
$$

$\Rightarrow$ assertion

Definition 7.12 *The measure F in representation $(*)$ of Theorem 7.10 is called* **spectral measure** *(to the continuous and positive semi-definite function $R(x)$). The corresponding distribution function*

$$F(u) = F(I_u) \qquad where \qquad I_u = \{x \in \mathbb{R}^m : \; x < u\}$$

is called **spectral function**. *If $F(A)$ is absolutely continuous with respect to Lebesgue measure, i.e. $F(A) = \int_A f(u)du$, then $f(u)$ is called* **spectral density**.

Theorem 7.12 *If $R(x)$ is absolutely integrable (i.e. $\int_{\mathbb{R}^m} |R(x)|dx < \infty$) then the spectral density $f(u)$ exists and it holds*

$$R(x) = \int_{\mathbb{R}^m} e^{i(x,u)} f(u) du.$$

Proof: apply Parseval equation for Fourier integrals: $(\tilde{f}, \tilde{g})_{L_2} = (f,g)_{L_2}$ with

$\tilde{f}(z) = \tilde{R}_N(z), \quad f(x) = R_N(x) = R(x)\exp(-\frac{1}{2N}|x|^2)$ (as in Theorem 7.11)

$\tilde{g}(z) = I_K(z)$ with $K = (y-h, y+h)$

$= \{z : \; y_k - h_k < z_k < y_k + h_k, \; k = 1, \ldots, m\}$

$$
\begin{aligned}
\Rightarrow \quad g(u) \quad &= \quad \frac{1}{\sqrt{2\pi}^m} \int_{\mathrm{R^m}} e^{i(u,z)} \tilde{g}(z) dz \quad = \quad \frac{1}{\sqrt{2\pi}^m} \int_{\mathrm{R^m}} e^{i(u,z)} I_K(z) dz \\
&= \quad \frac{1}{\sqrt{2\pi}^m} \prod_{k=1}^{m} \int_{y_k-h_k}^{y_k+h_k} e^{iu_k z_k} dz_k \\
&= \quad \frac{1}{\sqrt{2\pi}^m} \prod_{k=1}^{m} \frac{1}{iu_k} \big(\exp(i(y_k+h_k)u_k) - \exp(-i(y_k-h_k)u_k) \big)
\end{aligned}
$$

$$
\begin{aligned}
\Rightarrow \quad \int_K \tilde{R}_N(z) dz \quad &= \quad \int_{\mathrm{R^m}} \tilde{R}_N(z) \tilde{g}(z) dz = \int_{\mathrm{R^m}} R_N(u) g(u) dz \\
&= \quad \int_{\mathrm{R^m}} \underbrace{R(u) \exp\left(-\frac{|u|^2}{2N}\right)}_{|.| \leq |R(u)| \cdot 1} \; \underbrace{g(u)}_{|.| \leq \sqrt{2\pi}^m 2^m h_1 \ldots h_m} \; du \\
&\leq \quad \frac{1}{\sqrt{2\pi}^m} V(K) \int_{\mathrm{R^m}} |R(u)| du
\end{aligned}
$$

see $|\frac{1}{iu_k}(e^{i(y_k+h_k)u_k} - e^{i(y_k-h_k)u_k})| \leq |\frac{2}{u_k}e^{iy_ku_k}\sin(h_ku_k)| \leq 2|h_k|$

$V(K) = \prod\limits_{k=1}^{m} 2h_k$

define $F_N(K) := \frac{1}{\sqrt{2\pi}^m}\int\limits_K \tilde{R}_N(z)dz$ as in the proof of Theorem 7.11

$\Rightarrow$ F_N is absolutely continuous w.r.t. Lebesgue measure $\forall\ N$

$N \to \infty \Rightarrow$ F is absolutely continuous w.r.t. Lebesgue measure

$\Rightarrow$ $\exists\ f$ such that $F(A) = \int\limits_A f(z)dz$

(see Radon-Nikodym theorem; measure theory Theorem 7.4)

$\Rightarrow$ $R(x) = \int\limits_{\mathbb{R}^m} e^{i(x,z)}F(dz) = \int\limits_{\mathbb{R}^m} e^{i(x,z)}f(z)dz$ $\Rightarrow$ assertion

Corollary 7.1 *A function $R(t)$, $t \in (-\infty,\infty)$, is the correlation function of a m.-s.-continuous wide-sense stationary process if and only if it admits a representation $R(t) = \int\limits_{-\infty}^{\infty} e^{itu}F(du)$, where $F(.)$ is a finite measure on $\mathcal{B}^1$.*

Proof: apply Theorem 7.10 with $m = 1$

Definition 7.13 *A random function is called* **isotropic** *if its correlation function satisfies the condition*

$$R(x_1,x_2) = R(x_1,\rho(x_1,x_2)) \quad \textit{where} \quad \rho(x_1,x_2) = |x_1 - x_2| = \sqrt{\sum_{p=1}^{m}(x_{1,p} - x_{2,p})^2},$$

i.e. it depends only on x_1 and the distance between x_1 and x_2 (but not on the direction of $x_1 - x_2$).

For a homogeneous and isotropic function it holds

$$R(x_1,x_2) = R(x_1 - x_2) = R(\rho(x_1,x_2)) = R(|x_1 - x_2|).$$

If $m = 1$ and R is real then $R(x) = R(-x) = R(|x|)$ $\Rightarrow$ real valued wide-sense stationary processes are isotropic.

Theorem 7.13 *The function $R(\rho)$, $0 \leq \rho < \infty$, is the correlation function of a homogeneneous, isotropic, m.-s.-continuous random function if and only if it admits a representation*

$$R(\rho) = 2^{\frac{m-2}{2}}\Gamma\left(\frac{m}{2}\right)\int_0^\infty I_{\frac{m-2}{2}}(\rho r)\frac{1}{(\rho r)^{\frac{m-2}{2}}}\,G(dr),$$

where G is a finite measure on $[0,\infty)$ and $I_p(x)$ is the Bessel functions of first kind, i.e.

$$I_p(x) = \sum_{k=0}^{\infty}(-1)^k\frac{1}{k!\,\Gamma(p+k+1)}\left(\frac{x}{2}\right)^{p+2k}$$

with the Gamma function $\Gamma(x) = \int\limits_0^\infty e^{-t}t^{x-1}dt.$

Remark 7.7 For the Gamma function and the Bessel function their hold the following relations

- $\Gamma(1) = 1, \quad \Gamma(\frac{1}{2}) = \sqrt{\pi}; \qquad \Gamma(x+1) = x\Gamma(x) \quad \Rightarrow \quad \Gamma(n+1) = n!$
- $I_{-\frac{1}{2}}(x) = \sqrt{\frac{2}{\pi x}} \cos x, \quad I_{\frac{1}{2}}(x) = \sqrt{\frac{2}{\pi x}} \sin x$
- $(x^p I_p(x))' = x^p I_{p-1}(x) \quad \Rightarrow \quad I_p' = I_{p-1} - \frac{p}{x} I_p \qquad (\circ)$

<u>**special cases:**</u>

$$m = 1: \quad R(\rho) = \frac{1}{\sqrt{2}}\Gamma(\tfrac{1}{2}) \int_0^\infty I_{-\frac{1}{2}}(\rho r) \frac{1}{(\rho r)^{-\frac{1}{2}}} G(dr) = \int_0^\infty \cos(\rho r)\, G(dr)$$

Theorem 7.10, $R(x)$ real $\Rightarrow R(\rho) = R(-\rho) = \int_{-\infty}^{\infty} e^{i\rho r} F(dr)$

(where $G(A) = F(A) + F(-A), \quad A \in \mathcal{B}([0,\infty)))$

$$m = 2: \quad R(\rho) = \int_0^\infty I_0(\rho r)\, G(dr)$$

$$m = 3: \quad R(\rho) = \sqrt{2}\Gamma(\tfrac{3}{2}) \int_0^\infty I_{\frac{1}{2}}(\rho r) \frac{1}{\sqrt{\rho r}} G(dr) \quad \Rightarrow \quad R(\rho) = \int_0^\infty \frac{\sin(\rho r)}{\rho r} G(dr)$$

see $\Gamma(\frac{3}{2}) = \frac{1}{2} \cdot \Gamma(\frac{1}{2}) = \frac{\sqrt{\pi}}{2}; \quad I_{\frac{1}{2}}(x) = \sqrt{\frac{2}{\pi x}} \sin x;$

Proof: of Theorem 7.13

consider a sphere with radius ρ : and the center at the origin:

$B_\rho = \{x \in \mathbb{R}^m : \ \rho(0,x) < \rho\}$ open sphere

$S_\rho = \{x \in \mathbb{R}^m : \ \rho(0,x) = \rho\}$ boundary

$V_\rho = \int_{B_\rho} dx$ volume

$O_\rho = \int_{S_\rho} s(dx) = \frac{2}{\Gamma(\frac{m}{2})} \rho^{m-1} \pi^{\frac{m}{2}} \quad (\overset{m=2}{=} \quad 2\pi\rho)$ surface

isotropy $\Rightarrow R(x) = R(|x|) \Rightarrow$ R is constant on S_ρ

$$\Rightarrow \int_{S_\rho} R(x) s(dx) = R(\rho)\, O_\rho = \int_{S_\rho} \int_{\mathbb{R}^m} e^{i(x,u)} F(du)\ s(dx)$$

$$\overset{\text{Fubini}}{=} \int_{\mathbb{R}^m} \int_{S_\rho} e^{i(x,u)} s(dx)\ F(du) \qquad (+)$$

it holds

$$\int_{S_\rho} f(x) s(dx) = \frac{d}{d\rho} \int_{B_\rho} f(x) dx \quad \text{and} \quad \int_{B_\rho} e^{i(x,u)} dx = \left(\frac{2\pi\rho}{|u|}\right)^{\frac{m}{2}} I_{\frac{m}{2}}(\rho |u|)$$

(can be proven using analytic results)

$$\Rightarrow \quad \int\limits_{S_\rho} e^{i(x,u)} s(dx) = \frac{d}{d\rho}\left(\left(\frac{2\pi\rho}{|u|}\right)^{\frac{m}{2}} I_{\frac{m}{2}}(\rho|u|)\right)$$

$$= \frac{m}{2}\left(\frac{2\pi\rho}{|u|}\right)^{\frac{m}{2}-1} \frac{2\pi}{|u|} I_{\frac{m}{2}}(\rho|u|) + \left(\frac{2\pi\rho}{|u|}\right)^{\frac{m}{2}} \cdot \underbrace{I'_{\frac{m}{2}}(\rho|u|)}_{\overset{(\circ)}{=} I_{\frac{m}{2}-1} - \frac{m}{2\rho|u|} I_{\frac{m}{2}}} \cdot |u|$$

$$= \left(\frac{2\pi\rho}{|u|}\right)^{\frac{m}{2}} |u| I_{\frac{m}{2}-1}(\rho|u|)$$

$$(+) \ \Rightarrow \ R(\rho) = \frac{1}{O_\rho} \int\limits_{\mathbb{R}^m} \left(\frac{2\pi\rho}{|u|}\right)^{\frac{m}{2}} |u| I_{\frac{m}{2}-1}(\rho|u|) F(du)$$

it holds

$$\int\limits_{\mathbb{R}^m} f(|u|)F(du) = \lim_{n\to\infty} \sum_{k=1}^{n} f(r_k^*) F(B_{r_k} \setminus B_{r_{k-1}})$$

$$= \lim_{n\to\infty} \sum_{k=1}^{n} f(r_k^*)\, G([r_{k-1}, r_k)) = \int\limits_0^\infty f(r) G(dr)$$

define $\quad G([a,b)) := F(B_b \setminus B_a) \quad$ for $0 \le a < b$;
$\quad r_1 < r_2 < \ldots; \qquad r_k^* \in [r_{k-1}, r_k) \quad$ (see integral sums)

$$\Rightarrow \quad R(\rho) = 2^{\frac{m-2}{2}} \Gamma\left(\frac{m}{2}\right) \int\limits_0^\infty I_{\frac{m-2}{2}}(\rho r) \frac{1}{(\rho r)^{\frac{m-2}{2}}}\, G(dr) \Rightarrow \quad \text{assertion}$$

Consider $\xi(t,x)$ where $-\infty < t < \infty, \ \ x \in \mathbb{R}^m$:
let $\xi(t,x)$ be homogeneous w.r.t. (t,x) and isotropic w.r.t. x
$\Rightarrow \ \mathbf{E}\left(\xi(t+s,x)\overline{\xi(s,y)}\right) \overset{\text{homogeneous}}{=} R(t, x-y) \overset{\text{isotropic}}{=} R(t, \rho(x,y))$

Theorem 7.14 *A function $R(t,\rho)$ is the correlation function of a m.-s.-continuous random function $\xi(t,x)$ which is homogeneous w.r.t. (t,x) and isotropic w.r.t. x if and only if*

$$R(t,\rho) = \int\limits_{-\infty}^{\infty}\int\limits_0^\infty e^{itv}\, \Omega_m(\rho r) G(dv \times dr) \ \textit{where} \ \ \Omega_m(y) = \left(\frac{2}{y}\right)^{\frac{m-2}{2}} \Gamma\left(\frac{m}{2}\right) I_{\frac{m-2}{2}}(y).$$

Here G is a finite meassure on the half-plane $\{(v,r) : r \in [0,\infty), v \in (-\infty, +\infty)\}$.

Proof: Theorem 7.10 $\Rightarrow$

$$R(t,x) = \int\limits_{-\infty}^{\infty} e^{itv} \int\limits_{\mathbb{R}^m} e^{i(x,u)} F(dv \times du)$$

$$\overset{\text{proof Theorem 7.13}}{=} \int\limits_{-\infty}^{\infty} e^{itv} \int\limits_0^\infty \Omega_m(\rho r) G(dv \times dr) \Rightarrow \quad \text{assertion}$$

8 References

Gihman, I.I.; Skorohod, A.V. The Theory of Stochastic Processes 1. Springer-Verlag, Berlin/Heidelberg/New York 1974

Todorovic, P. An Introduction to Stochastic Processes and Their Application. Springer-Verlag, Springer Series in Statistics 1992

Wentzell, A.D. Theorie zufälliger Prozesse. Akademie-Verlag, Berlin 1979

Doob, J.L. Stochastic Processes. John Wiley & Sons 1990 (1952)