Statistics and probability theory

WS 2002 – 2003 : BA UNI Mannheim

Dan Fulea

Internetbegleitung
http://www.mathi.uni-heidelberg.de/~dan/Teaching/WS2002-2003/

Inhaltsverzeichnis

I Theoretical prerequisities **2**

1 Introduction : Generalities **3**
1.1 Probability spaces . . . 3
1.2 Examples . . . 3
1.2.1 Finite probability space . . . 3
1.2.2 Countable probability space . . . 4
1.2.3 Continous probability spaces . . . 4

2 Random variables **6**
2.1 Definition . . . 6
2.2 Examples . . . 6
2.3 The mean value (expectation) $\mathbb{E}$ of random variables . . . 7

3 Independence **8**
3.1 Independence of events . . . 8
3.2 Independence of σ–algebras . . . 8
3.3 Independence of random variables . . . 8

4 Processes **9**
4.1 Martingales . . . 9
4.2 Markov processes . . . 10
4.3 Brownian motion . . . 10

II Problems session **11**

5 Discrete probability spaces **12**

6 Exam preparation **13**

Teil I

Theoretical prerequisities

Kapitel 1

Introduction : Generalities

1.1 Probability spaces

Definition 1 (Probability space) *A triple* $(\Omega, \mathcal{F}, \mathbb{P})$ *is called a* **probability space**, *iff the following axioms hold:*

- Ω *is a set.*
- $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ *is a subset of the power set of* Ω, *which contains as elements the empty set* $\emptyset$ *and the whole space* Ω *and which is closed under:*
  - ▷ *taking complements and*
  - ▷ *taking (finite or countable) unions.*
- $\mathbb{P} : \mathcal{F} \to [0,1]$ *is a map satisfying:*
  - ▷ $\mathbb{P}(\Omega) = 1$.
  - ▷ *For pairwise disjoint elements of* $\mathcal{F}$, $(A_n)_{n\in\mathbb{N}}$, *one has:*

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mathbb{P}(A_n) \ .$$

1.2 Examples

1.2.1 Finite probability space

A probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called finite, iff the underlying set Ω of the space is a finite set. Then the smallest non empty events in $\mathcal{F}$ are called the atoms of $\mathcal{F}$.
Example: $\Omega = \{1,2,3,A,B,C,D,x,y\}$ and $\mathcal{F}$ contains the events:
$\emptyset$,
$\Omega_1 := \{1,2,3\}$, $\Omega_2 := \{A,B,C,D\}$, $\Omega_3 := \{x,y\}$,
$\Omega_1 \sqcup \Omega_2 = \{1,2,3,A,B,C,D\}$, $\Omega_1 \sqcup \Omega_3 = \{1,2,3,x,y\}$, $\Omega_2 \sqcup \Omega_3 = \{A,B,C,D,x,y\}$,
$\Omega_1 \sqcup \Omega_2 \sqcup \Omega_3 = \{1,2,3,A,B,C,D,x,y\} = \Omega$.
The atoms are of course the sets Ω_i, $i = 1,2,3$.
The probability $\mathbb{P}$ is then determined, when one fixes $p_i := \mathbb{P}(\Omega_i) \geqslant 0$, $i = 1,2,3$. The relation $p_1 + p_2 + p_3 = 1$ holds.
For the purposes of the probability theory one can suppose, that all atoms contain exactly one element.
In the above case one could replace the space $(\Omega, \mathcal{F}, \mathbb{P})$ by the space $(\Omega', \mathcal{F}', \mathbb{P}')$ with $\Omega' := \{1', 2', 3'\}$, $\mathcal{F}' := \mathcal{P}(\Omega')$ and $\mathbb{P}'$ determined by $\mathbb{P}'(\ \{i'\}\ ) := p_i := \mathbb{P}(\Omega_i)$, $i = 1,2,3$.
So we can reduce ourselves to the case:

$$\begin{aligned} (\Omega, \mathcal{F}, \mathbb{P}) \qquad & \Omega := \{1,2,\dots,N\} \ , \\ & \mathcal{F} := \mathcal{P}(\Omega) \ , \\ & \mathbb{P} \text{ determined by } \mathbb{P}(\ \{i\}\ ) := p_i \geqslant 0 \\ & \qquad \text{for given } p_1, p_2, \dots, p_N \ , \ p_1 + p_2 + \dots + p_N = 1 \ . \end{aligned}$$

So for each finite sum with N terms having the value equal to one we can consider the corresponding probability space. There is of course a difference between writing the terms of a sum and writing the sum. We will but write a sum with value one as an allusion to the corresponding probability space with the distribution of probability 1 given by the terms of this sum as they are written. We obtain the following important examples of finite probability spaces:

• LAGRANGE experiments, uniform probability repartition.

$$\underbrace{\frac{1}{N} + \frac{1}{N} + \cdots + \frac{1}{N}}_{N \text{ times}} = 1 \ .$$

• Binomial repartition. Let $N \in \mathbb{N}$ ba a natural number. Consider two positive numbers p, q with $p + q = 1$. Then the binomial repartition of probability corresponds to the decomposition of one:

$$\binom{N}{0} p^0 q^N + \binom{N}{1} p^1 q^{N-1} + \binom{N}{2} p^2 q^{N-2} + \cdots = \sum_{0 \leqslant k \leqslant N} \binom{N}{k} p^k q^{N-k} = (p+q)^N = 1 \ .$$

1.2.2 Countable probability space

A probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called countable, iff the underlying set Ω of the space is a countable set. Then the smallest non empty events in $\mathcal{F}$ are called the atoms of $\mathcal{F}$. It may happen, that Ω has finitely many atoms. In this case one can - for the purposes of the probability theory - replace tne space by a finite space. As in the last case, one can suppose, that all atoms are containing exactly one element.
So we can reduce ourselves to the case:

$$\begin{aligned} (\Omega, \mathcal{F}, \mathbb{P}) \qquad & \Omega := \{1, 2, \dots, N, \dots\} \ , \\ & \mathcal{F} := \mathcal{P}(\Omega) \ , \\ & \mathbb{P} \text{ determined by } \mathbb{P}(\ \{i\}\ ) := p_i \geqslant 0 \\ & \qquad \text{for given } p_1, p_2, \dots, p_N, \dots \ , \ p_1 + p_2 + \cdots + p_N + \cdots = 1 \ . \end{aligned}$$

So for each series with the sum equal to one we can consider the corresponding probability space.
There is of course a difference between writing the terms of a series and writing the sum of the series. We will but write a series to have value one and consider the corresponding summands/terms to be the values of $p_1, p_2, \dots$. Sometimes it is preferable to start with the index zero, i.e. give $p_0, p_1, \dots$.
We obtain the following important examples of finite probability spaces:
• Geometric repartition. Fix a $q \in (0, 1)$. Then one has

$$\sum_{k \geqslant 0} \frac{q^k}{1-q} = 1 \ .$$

• POISSON repartition or exponential repartition with parameter $\lambda \geqslant 0$.

$$\sum_{k \geqslant 0} e^{-\lambda} \frac{\lambda^k}{k!} = 1 \ .$$

1.2.3 Continous probability spaces

Such probability spaces are modelling experiments, where the number of possible experiment results is no longer finite or countable. Such situations are more complicated. We resume ourselves to give only the main examples of such a situation. More delicate aspects are covered by the measure theory on general and/or euclidean spaces.
Consider the set $\mathbb{R}$ of real numbers. The standard σ-algebra on $\mathbb{R}$ is the σ-algebra $\mathcal{B} = \mathcal{B}(\mathbb{R})$ of the so called BORELian subsets of $\mathbb{R}$. It is the smallest σ-algebra on $\mathbb{R}$ containing all intervals. An interval is a subset of $\mathbb{R}$ of the form ($\emptyset$ or) $[a, b]$ or $[a, b)$ or $(a, b]$ or (a, b) for $a \leqslant b$ arbitrar elements of $\mathbb{R}$. (The interval $[a, a]$ reduces to the (set with one) point a.)
One can show, that each BORELian subset can be written as a disjoint union of intervals. Reciprocally, each (disjoint) union of intervals is a BORELian (sub)set.
In the measure theory (on $\mathbb{R}$ and/or $\mathbb{R}^k$) one learns how to integrate functions $f : \mathbb{R} \to \mathbb{R}$ or more general $f : \mathbb{R}^k \to \mathbb{R}$

▷ (i) which are positive and measurable. (A measurable function $f : \mathbb{R}^k \to \mathbb{R}$ is characterized by the fact, that it "turns back through the preimage map f^{-1} measurable sets of $\mathbb{R}$ to measurable sets of $\mathbb{R}^k$.) The result $\int_{\mathbb{R}^k} f(x)\ dx \geqslant 0$ is a positive number or ("the symbol) infinity ∞. In the first case $\int_{\mathbb{R}^k} f(x)\ dx < \infty$ the function f is called integrable.

▷ (ii) which are measurable, such that the positive function $|f|$ is integrable in the sense of (i).

Especially, continous functions $f : \mathbb{R}^k \to \mathbb{R}$ are measurable.
Let $\Omega \subset \mathbb{R}^k$ be a measurable subset. Then a function $f : \Omega \subset \mathbb{R}^k \to \mathbb{R}$ is measurable/integrable iff the function $\tilde{f} : \mathbb{R}^k \to \mathbb{R}$, which coincides with f on Ω and is zero outside Ω, is. Our standard example of a "continous probability space" is the following one:

Fix $\Omega \subseteq \mathbb{R}^k$ measurable. Fix a "density function $f : \Omega \to \mathbb{R}$. This means

$$f \geqslant 0 \qquad \text{and} \qquad \int_\Omega f(x)\ dx = 1\ .$$

Set $\mathcal{F}$ to be the set of measurable subsets of Ω. Define the probability $\mathbb{P}$ on $(\Omega, \mathcal{F})$ by

$$\mathbb{P}(A) := \int_A f(x)\ dx \qquad \text{for all measurables } A \subseteq \Omega\ .$$

Then is $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space. We can place in this framework the following special probability spaces, obtained by special choices of the space $\Omega \subseteq \mathbb{R}^k$ and of the density f on Ω.

The uniform distribution on the intervall $[a, b]$

$\Omega := [a, b]$, $a < b$. The density is the function $f(x) := (x - a)/(b - a)$. Then the line:

$$\int_{[a,b]} \frac{x-a}{b-a}\ dx = 1$$

is analogous to the (finite or countable) sumes in the examples for finite or countable probability spaces in the last subsections. The sums are replaced by an integral. The finite or countable index set is replaced by an interval (more general, a measurable subset of $\mathbb{R}$ or $\mathbb{R}^k$).
We willalso in the following very allusive write down an integral with value 1 over a measurable subset Ω of a function f and mean by this the corresponding probability space Ω to the density f.

The standard normal distribution or the standard gaussian distribution $N(0, 1)$ on $\mathbb{R}$

This corresponds to the integral:

$$\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\ dx = 1\ .$$

The GAUSSian density function $\rho = \rho_{0,1}$ is very important, so that we will display it once more:

$$\rho(x) = \rho_{0,1}(x) := \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\ .$$

The normal distribution or the gaussian distribution $N(\mu, \sigma^2)$ on $\mathbb{R}$

The parameters have names: $\mu \in \mathbb{R}$ is called the mean and $\sigma^2 > 0$ is called the variance. One usually chooses σ to be positive: $\sigma > 0$.
This distribution corresponds to the integral:

$$\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\ dx = 1\ .$$

The GAUSSian density function ρ_{μ,σ^2} is very important, so that we will display it once more:

$$\rho_{\mu,\sigma^2}(x) := \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\ .$$

Please remark the two possibilities of writing $\sqrt{2\pi\sigma^2} = \sqrt{2\pi}\ \sigma$, that often lead to confusions for novices.

Kapitel 2

Random variables

2.1 Definition

Definition 2 (Random variable) *Let $(\Omega, \mathcal{F})$ be a σ–algebra structure on the set Ω. Let $(\Omega', \mathcal{F}')$ be a σ–algebra structure on the set Ω'. A map*

$$X : \Omega \to \Omega'$$

is called **measurable** *or also a* **random variable**, *iff for all $\mathcal{F}'$–measurable events A' in $\mathcal{F}'$ the set theoretical preimage $X^{-1}(A')$ is an $\mathcal{F}$–measurable event:*

$$X^{-1}(A') \in \mathcal{F} \ .$$

A **measurable** *map or* **random variable** *X between two probability spaces $(\Omega, \mathcal{F}, \mathbb{P})$ and $(\Omega', \mathcal{F}', \mathbb{P}')$ is a measurable map betwwen the underlying σ–algebras $(\Omega, \mathcal{F})$ and $(\Omega', \mathcal{F}')$.*

Definition 3 *Let $X : \Omega \to \Omega'$ be a map between sets. Let us suppose, that there is a given σ–algebra $\mathcal{F}'$ on the set Ω'. Then the smallest σ–algebra $\mathcal{F}$ on Ω making the map X a measurable map between the spaces $(\Omega, \mathcal{F})$ and $(\Omega', \mathcal{F}')$ is called the σ-***algebra generated by** *X and denoted by $\sigma(X)$. One has:*

$$\sigma(X) = \{ \ X^{-1}(A') \ : \ A' \in \mathcal{F}' \ \} \ .$$

We will always consider the σ–algebra of **borel**ians $\mathcal{B}$ on the spaces $\mathbb{R}$ or $\mathbb{R}^k$ or on (**borel**ian) subsets of them. Then a map $X : \mathbb{R} \to \mathbb{R}$ is measurable iff

$$X^{-1}(\ (-infty, a] \ ) \subset \mathbb{R}$$

is a measurable (i.e. **borel**ian) set.

2.2 Examples

• Let $(\Omega, \mathcal{F})$ be a σ–algebra. Let $A \in \mathcal{F}$ be a measurable set. The **characteristic function** of the set A, denoted by χ_A or also by 1_A in the literature is the map:

$$1_A(\omega) := \begin{cases} 1 & \text{for } \omega \in A \ , \\ 0 & \text{for } \omega \notin A \ . \end{cases}$$

Then $1_A : \Omega \to \mathbb{R}$ is measurable. This is so because one has:

$$1_A^{-1}(\ (-\infty, a] \ ) = \begin{cases} \emptyset \in \mathcal{F} & \text{for } a < 0 \ , \\ A \in \mathcal{F} & \text{for } 0 \leqslant a < 1 \ , \\ \emptyset \in \mathcal{F} & \text{for } 1 \leqslant a \ . \end{cases}$$

• A linear combination of charactersitic functions

$$X := \sum_{i \in I} a_i 1_{A_i} \ , \qquad a_i \in \mathbb{R} \ , \ A_i \in \mathcal{F} \qquad \text{for all } i \in I \ ,$$

is called a step function (with steps A_i, $i \in I$). Step functions are measurable.

• Let (X_n) be a monotone increasing sequence of step functions on $(\Omega, \mathcal{F}, \mathbb{P})$ with value in $\mathbb{R}$: $X_1 \leqslant X_2 \leqslant \ldots \leqslant X_n \leqslant \ldots$

Then one can pointwise define

$$X := \lim_{n\to\infty} X_n \ , \qquad i.e. \qquad X(\omega) := \lim_{n\to\infty} X_n(\omega) \ .$$

The funcion X is then defined on Ω and takes values in $\mathbb{R} \cup \{\infty\}$. One can define the σ–algebra of BOREL sets in $\mathbb{R}\cup\{\infty\}$ or even in $\overline{\mathbb{R}} := \mathbb{R}\cup\{\infty\}$ to be the smallest σ–algebra containing all intervals $[a,b]$ with $-\infty \leqslant a \leqslant b \leqslant +\infty$. Then $X := \lim X_n$ is measurable as a function $\Omega \to \overline{\mathbb{R}}$. One can reciprocally show: Each measurable function $\Omega \to \mathbb{R}$ can be written as the limit of a monotone sequence of step functions.

• Each continous function $X : \mathbb{R} \to \mathbb{R}$ is measurable. (The preimage $X^{-1}(\ (-infty, a]\ )$ of the closed set $(-infty, a]$ is closed, thus measurable.)
• The composition of measurable functions is measurable.
• Sums and products of real valued measurable maps are also measurable.

2.3 The mean value (expectation) $\mathbb{E}$ of random variables

One defines for positive measurable functions $X : \Omega \to \mathbb{R}_{\geqslant 0}$ the integral of X, also called mean value of X, also called expectation of X to be the positive "number" $\mathbb{E}[X] \in [0, +\infty]$

$$\mathbb{E}[X] \qquad \text{also denoted by} \qquad \int_\Omega X \; d\mathbb{P} = \int_\Omega X(\omega) \; d\mathbb{P}(\omega)$$

such that the map $\mathbb{E} : \ X \to \mathbb{E}[X]$ has the following properties:
• $\mathbb{E}$ is linear: $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$ for all $a, b \in \mathbb{R}_{\geqslant 0}$ and positive measurable functions $X, Y : \Omega \to \mathbb{R}$.
• $\mathbb{E}$ is monotone: $\mathbb{E}[X] \leqslant \mathbb{E}[Y]$ for positive measurable functions $X, Y : \Omega \to \mathbb{R}$ with $X \leqslant Y$.
• $\mathbb{E}$ is given on charactersitic functions by $\mathbb{E}[1_A] := \mathbb{P}(A)$ for all $A \in \mathcal{F}$.
• $\mathbb{E}$ is compatible with monotone limits: Let $X_1 \leqslant X_2 \leqslant \ldots \leqslant X_n \leqslant \ldots$ be a monotone sequence of measurable functions (for instance step functions with measurable steps). Then the monotone sequence $\mathbb{E}[X_1] \leqslant \mathbb{E}[X_2] \leqslant \ldots \leqslant \mathbb{E}[X_n] \leqslant \ldots$ converges in $\overline{\mathbb{R}}$ and one has:

$$\mathbb{E}\left[\lim_{n\to\infty} X_n\right] = \lim_{n\to\infty} \mathbb{E}\left[X_n\right] \ .$$

Definition 4 *A measurable function $X : \Omega \to \mathbb{R}$ is called* **integrable** *iff the positive measurable function $|X|$ has a finite integral:*

$$\mathbb{E}[|X|] = \int_\Omega |X| \; d\mathbb{P} < \infty \ .$$

The space of integrable functions (modulo functions which are zero on a set of probability 1) is denoted by

$$L^1(\Omega, \mathcal{F}, \mathbb{P}) \qquad \text{or simply} \qquad L^1(\mathbb{P}) \ .$$

Then one has $\mathbb{E}[|X|] = 0$ iff $X = 0$ on a set of probability one. Terminology: One also says $X = 0$ almost everywhere (with respect to $\mathbb{P}$) in this case.

Definition 5 *For an integrable function $X : \Omega \to \mathbb{R}$ one has the positive part of X, $X_+ := \max(X, 0) : \Omega \to \mathbb{R}$ and the (minus) negative part of X, $X_- := \min(-X, 0) : \Omega \to \mathbb{R}$:*

$$X = X_+ - X_- \qquad \text{with integrable functions } 0 \leqslant X_+, X_- \leqslant |X| \ .$$

Then one defines:

$$E[X] := \mathbb{E}[X_+] - \mathbb{E}[X_-] \in \mathbb{R} \ .$$

The expectation $\mathbb{E}$ is then a map

$$\mathbb{E} : Ł^1(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$$

having the properties:
• $\mathbb{E}$ is *linear* $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$ for all $a, b \in \mathbb{R}_{\geqslant 0}$ and all $X, Y \in Ł^1(\mathbb{P})$.
• $\mathbb{E}$ is monotone: $\mathbb{E}[X] \leqslant \mathbb{E}[Y]$ for all $X, Y \in Ł^1(\mathbb{P})$ with $X \leqslant Y$.
• $\mathbb{E}$ is given on charactersitic functions 1_A (which always are integrable) by $\mathbb{E}[1_A] := \mathbb{P}(A)$ for all $A \in \mathcal{F}$.
• $\mathbb{E}$ is compatible with limits: Let $X_1, X_2, \ldots, X_n, \ldots$ be an arbitrary sequence of measurable functions, which are *dominated* by an integrable function $Y \in L^1(\mathbb{P})$:

$$|X_n| \leqslant Y \qquad \text{for all } n \geqslant 1$$

and such that the poinwise defined limit function X, $X(\omega) := \lim X_n(\omega)$ exists with the exception of a set of probability zero (i.e. it exists almost everywhere).
Then the limit X is integrable (because it satisfies $|X| \leqslant Y$ almost everywhere) and one has:

$$\mathbb{E}\left[\lim_{n\to\infty} X_n\right] = \lim_{n\to\infty} \mathbb{E}\left[X_n\right] \ .$$

This property is called "LEBESGUE dominated convergence".

Kapitel 3

Independence

3.1 Independence of events

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Two events A, B are called **independent** iff the following relation holds:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B) \ .$$

More general:

Definition 6 *A collection $\{A_i \ : \ i \in I\}$ of events, indexed by an index set I of arbitrary cardinality, is called* **independent** *iff for each finite subcollection $A_{i_1}, A_{i_2}, \dots, A_{i_n}$ one has the equality:*

$$\mathbb{P}\left(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_n}\right) = \mathbb{P}\left(A_{i_1}\right) \mathbb{P}\left(\cap A_{i_2}\right) \dots \mathbb{P}\left(A_{i_n}\right) \ .$$

3.2 Independence of σ–algebras

Definition 7 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A collection $\{\mathcal{F}_i \ : \ i \in I\}$ of subalgebras $\mathcal{F}_i \subseteq \mathcal{F}$, indexed by an index set I of arbitrary cardinality, is called* **independent** *iff for each finite family of indices $i_1, i_2, \dots, i_n$ and for each choice of events $A_{i_1} \in F_{i_1}, A_{i_2} \in F_{i_1}, \dots, A_{i_n} \in F_{i_1}$, one has the equality:*

$$\mathbb{P}\left(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_n}\right) = \mathbb{P}\left(A_{i_1}\right) \mathbb{P}\left(\cap A_{i_2}\right) \dots \mathbb{P}\left(A_{i_n}\right) \ .$$

3.3 Independence of random variables

Definition 8 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A collection $\{X_i \ : \ i \in I\}$ of random variables $X_i : \Omega \to \Omega'$ into a common values domain Ω' is called* **independent** *iff the family of σ–algebras $\sigma(X_i)$ is independent.*

Kapitel 4

Processes

Definition 9 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A process X indexed by the ordered (time) index set $T = (T, \leqslant)$ is a family of $\mathcal{F}$-measurable random variables*

$$X := (X_t)_{t \in T} \ .$$

We will often neglect in notations the time index set T.

Because of the very general definition, process are too general objects to give a rich theory. We will in the following study special classes of processes. There are but some common features:

Definition 10 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.*
A **filtration** *on $\mathcal{F}$ is a family $\mathbb{F} := (\mathcal{F}_t)_{t \in T}$ of σ–subalgebras of $\mathcal{F}$, such that it holds:*

$$\mathcal{F}_s \subseteq \mathcal{F}_t \qquad \text{for all } s, t \in T \ , \ s \leqslant t \ .$$

A process X indexed by the ordered by the same index set T is called $\mathbb{F}$–adapted, iff for all $t \in T$ the random variable X_t is $\mathcal{F}_t$-measurable. One often imposes the condition, that the filtration $\mathbb{F}$ is **exhaustive**, *i.e. the union of all $\mathcal{F}_t$, $t \in T$, equals $\mathcal{F}$.*

Definition 11 *Let (X_t) be a processes on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.*
We define the σ-algebra $\mathcal{F}_{\cdot \leqslant t} \subseteq \mathcal{F}$ to be the σ-algebra generated by the random variables $(X_s)_{s \leqslant t}$. It is the smallest σ-algebra (included in $\mathcal{F}$ containing all σ-algebras $\sigma(X_s)$, $s \in T$, $s \leqslant t$. We sometimes humanly call $\mathcal{F}_{\cdot \leqslant t}$ to be the **the σ-algebra of the past of t associated to the process X**.
We define the σ-algebra $\mathcal{F}_{\cdot < t} \subseteq \mathcal{F}$ to be the σ-algebra generated by the random variables $(X_s)_{s \leqslant t}$. It is the smallest σ-algebra (included in $\mathcal{F}$ containing all σ-algebras $\sigma(X_s)$, $s \in T$, $s \leqslant t$. We sometimes humanly call $\mathcal{F}_{\cdot < t}$ to be the **the σ-algebra of the strict past of t associated to the process X**.
We define the σ-algebra $\mathcal{F}_{\cdot = t} \subseteq \mathcal{F}$ to be the σ-algebra generated by the random variable X_t. It is exactly the σ-algebra $\sigma(X_s)$. We sometimes humanly call $\mathcal{F}_{\cdot = t}$ to be the **the σ-algebra of the present of t associated to the process X**.
We define the σ-algebra $\mathcal{F}_{\cdot \geqslant t} \subseteq \mathcal{F}$ to be the σ-algebra generated by the random variables $(X_s)_{s \geqslant t}$. It is the smallest σ-algebra (included in $\mathcal{F}$ containing all σ-algebras $\sigma(X_s)$, $s \in T$, $s \geqslant t$. We sometimes humanly call $\mathcal{F}_{\cdot \geqslant t}$ to be the **the σ-algebra of the future of t associated to the process X**.
We define the σ-algebra $\mathcal{F}_{\cdot > t} \subseteq \mathcal{F}$ to be the σ-algebra generated by the random variables $(X_s)_{s \geqslant t}$. It is the smallest σ-algebra (included in $\mathcal{F}$ containing all σ-algebras $\sigma(X_s)$, $s \in T$, $s \geqslant t$. We sometimes humanly call $\mathcal{F}_{\cdot > t}$ to be the **the σ-algebra of the strict future of t associated to the process X**.

4.1 Martingales

Definition 12 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $\mathbb{F} = (\mathcal{F}_t)_{t \in T}$ be a filtration on it. Let X be an $\mathbb{F}$–adapted process.*
Then X is called a **martingale**, *iff the following condition holds:*

$$\mathbb{E}[X_t | \mathcal{F}_s] = X_s \qquad \text{for all } s, t \in T \text{ with } s \leqslant t \ .$$

For a martingale with the discrete time index set $T = \{ \ 0, 1, \ldots, (n-1), n, (n+1), \ldots \ \}$ one has the following diagram:

$$X_0 \xleftarrow{\mathbb{E}[\ |\mathcal{F}_0]} X_1 \xleftarrow{\mathbb{E}[\ |\mathcal{F}_1]} X_2 \longleftarrow \cdots \longleftarrow X_{n-1} \xleftarrow{\mathbb{E}[\ |\mathcal{F}_{n-1}]} X_n \xleftarrow{\mathbb{E}[\ |\mathcal{F}_n]} X_{n+1} \longleftarrow \cdots$$

4.2 Markov processes

4.3 Brownian motion

Teil II

Problems session

Kapitel 5

Discrete probability spaces

Aufgabe 1 *In a class room there are $N = 22$ persons. Which is the probability, that at least two persons have the same birthday ? (We assume, that no person is borned on the 29th of february.)*

Aufgabe 2 *In a class room there are $N = 22$ persons. Which is the probability, that at least two persons have the same birthday ? (We assume, that persons are eventually also borned on the 29th of february.)*

Aufgabe 3 *On the dance floor there are $2n$ dancers, n women numbered $1, 2, \dots, n$ and n men numbered $1', 2', \dots, n'$. In real life, 1 and $1'$ are married, this also happens for 2 and $2'$ and so on. A person having no knowledge of the marriage correspondence of the n pairs in real life, builds n dancing pairs for a performace following artistical criteria. Which is the probability, that none of the n dance pairs is a married pair in real life.*

Aufgabe 4 *An urn contains N balls numbered from 1 to N. A person extracts k times a ball, each time putting it back into the urn. Which is the probability, that the sum of the numbers on the k extracted balls equals S. (S is of course a natural number between $k \times 1$ and $k \times N$.)*

Aufgabe 5 *In a box there are 1000 cards. On these cards there are written some real numbers. A person with no knowledge of the written numbers plays the following game:*
(S)he takes out of the box one by one some cards out of the 1000 cards and records the written numbers up to some point, when (s)he decides to stop. Then (s)he wins some (un)important amount of money, when the **last** *extracted card is the biggest one among all 1000 cards.*
The players imagines now the following strategy to play the game:
He fixes a number N between 1 and 1000 and randomly extracts N cards from the box. Then he continues to extract cards up to the point, that he extracts a first number bigger then the (first N) already extracted cards.
Which is the "optimal" choice of N, so that the above strategy has a maximal probability to lead to a win. Which is this probability ?

Aufgabe 6 *A train has four (passenger) waggons. At some railway station a number of N persons randomly get in. There is no magnetic (attraction or rejection) forces between the passengers. Which is the probability, that in each of the four waggons at least one passenger gets in ?*

Aufgabe 7 *Two dices are repeatedly dropped on a table N times. Which is the probability p_N to obtain at least once the double of six ? (Which is the approximative value of p_N for N equal to 20 and 36 and 60 and 100 ?)*

Kapitel 6

Exam preparation

Aufgabe 8 *Let X be a random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with the standard normal distribution $N(0,1)$.*
Which are the moments M_0, M_1, M_2, M_3 of X ?

$$M_k := \int_\Omega X^k \, d\mathbb{P}$$

Find a recursive formula for the computation of these moments.
Find the mean and the variance of X by an explicit computation.

LÖSUNG: The density of a standard normally distributed random variance is the function:

$$\rho(x) := \frac{1}{\sqrt{2\pi}} e^{-\dfrac{x^2}{2}} \ .$$

Then we have:

$$\begin{aligned} M_k &:= \mathbb{E}[\ X^k\ ] \\ &= \int_\Omega X^k \, d\mathbb{P} \\ &= \int_\mathbb{R} x^k \ \rho(x)\ dx = \int_\mathbb{R} x^k \ \frac{1}{\sqrt{2\pi}} e^{-\dfrac{x^2}{2}}\ dx \ . \end{aligned}$$

Especially:

$$M_0 = \mathbb{E}[\ X^0\ ] = \mathbb{E}[\ 1\ ] = \mathbb{P}(\Omega) = 1 \ ,$$

$$M_1 = \mathbb{E}[\ X^1\ ] = \mathbb{E}[\ X\ ] = \int_\mathbb{R} x\ \rho(x)\ dx = \int_\mathbb{R} x\ \frac{1}{\sqrt{2\pi}} e^{-\dfrac{x^2}{2}}\ dx = 0 \ .$$

We used the fact, that the integral of an odd function is zero. The same reason shows directly, that $M_1 = M_3 = M_5 = \dots = 0$, all moments of odd degree are zero. For general (even) $k \geqslant 2$ we develop the recursion formula:

$$M_k = \mathbb{E}[\ X^k\ ] = \int_\mathbb{R} x^k\ \rho(x)\ dx = \int_\mathbb{R} x^k\ \frac{1}{\sqrt{2\pi}} e^{-\dfrac{x^2}{2}}\ dx$$

We will use partial integration:

$$
\begin{aligned}
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} x^{k-1}\, x e^{-\dfrac{x^2}{2}}\, dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} x^{k-1} \left(-e^{-\dfrac{x^2}{2}} \right)'\, dx \\
&= \underbrace{\frac{1}{\sqrt{2\pi}} \left[x^{k-1} \left(-e^{-\dfrac{x^2}{2}} \right) \right]_{-\infty}^{+\infty}}_{=0} - \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} (\, x^{k-1}\,)' \left(-e^{-\dfrac{x^2}{2}} \right) dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} (k-1) x^{k-2}\, e^{-\dfrac{x^2}{2}}\, dx \\
&= (k-1) M_{k-2}\ .
\end{aligned}
$$

It quickly follows:

$$
\begin{aligned}
M_{2k} &= (2k-1)(2k-3)\dots(5)(3)(1)M_0 \\
&= (2k-1)(2k-3)\dots(5)(3)(1) = \frac{(2k)!}{2^k\ k!}\ .
\end{aligned}
$$

Some `maple` experiments confirm our formula:

```
rho := x -> exp( -x^2/2 ) / sqrt(2*Pi) ;
M   := k -> int( x^k * rho(x) , x=-infinity..+infinity );
for k from 0 to 10 do
  print( k, "    " , M(k) );
od;
```

■

Aufgabe 9 *Let X, Y be two independent variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.*
Find a formula for the third centered moment $\overline{M}_3(X+Y)$ of $(X+Y)$ in terms of the centered moments of the variables X and Y.
Generalization.

LÖSUNG: We compute under the circumstances of the exercise:

$$
\begin{aligned}
\overline{M}_3(X+Y) &= M_3(X+Y-\mathbb{E}(X+Y)) \\
&= \mathbb{E}[\ (X+Y-\mathbb{E}(X+Y))^3\ ] \\
&= \mathbb{E}[\ ((X-\mathbb{E}X)+(Y-\mathbb{E}Y))^3\ ]
\end{aligned}
$$

We now use the binomial formula...

$$
= \mathbb{E}[\ (X-\mathbb{E}X)^3 + 3(X-\mathbb{E}X)^2(Y-\mathbb{E}Y)) + 3(X-\mathbb{E}X)(Y-\mathbb{E}Y))^2 + (Y-\mathbb{E}Y))^3\ ]
$$

... and the linearity of the expectation $\mathbb{E}$

$$
= \mathbb{E}[\ (X-\mathbb{E}X)^3\ ] + 3\mathbb{E}[\ (X-\mathbb{E}X)^2(Y-\mathbb{E}Y))\ ] + 3\mathbb{E}[\ (X-\mathbb{E}X)(Y-\mathbb{E}Y))^2\ ] + \mathbb{E}[\ (Y-\mathbb{E}Y))^3\ ]
$$

Now we use the independence of X, Y in the form $\mathbb{E}[\ f(X)g(Y)\ ] = \mathbb{E}[\ X\ ]\mathbb{E}[\ Y\ ]$ for suitable measurable functions f, g. In our cases these functions are power functions...

$$
\begin{aligned}
&= \mathbb{E}[\ (X-\mathbb{E}X)^3\ ] + 3\mathbb{E}[\ (X-\mathbb{E}X)^2\ ]\ ;\mathbb{E}[\ (Y-\mathbb{E}Y))\ ] + 3\mathbb{E}[\ (X-\mathbb{E}X)\ ;\mathbb{E}[\ (Y-\mathbb{E}Y))^2\ ] + \mathbb{E}[\ (Y-\mathbb{E}Y))^3\ ] \\
&= \overline{M}_3(X) + 3\overline{M}_2(X)\overline{M}_1(Y) + 3\overline{M}_1(X)\overline{M}_2(Y) + \overline{M}_3(Y) \\
&= \overline{M}_3(X) + \overline{M}_3(Y)\ .
\end{aligned}
$$

The last equality holds beacuse of $\overline{M}_1(X) = \overline{M}_1(Y)$. The general formula follows the same lines:

$$
\overline{M}_n(X+Y) = \sum_{\substack{0 \leqslant p,q \leqslant n \\ p+q=n}} \binom{n}{p} \overline{M}_p(X)\overline{M}_q(Y)
$$

There are minor simplifications in this general formula, exploiting $\overline{M}_0(X) = \overline{M}_0(Y) = 1$ and $\overline{M}_1(X) = \overline{M}_1(Y) = 0$ ■

Aufgabe 10 *The stock price of some financial instrument at time $T = 1$ year is modelled by the random variable X on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.*
It is supposed that X has the density $\rho(x)$ given by:
(a) $\rho(x) := C\ (x-100)^2\ 1_{[99,101]}$.
(b) *$\rho(x)$ is the normal density with parameters $\mu := 100$ (EURO) and $\sigma := 3$ (EURO).*
Which is the associated option price at time $t = 0$ ("now") with maturity $T = 1$ when the exercise price is $A := 100.5$ (EURO). Use a value of the anual rate of 5 % on the market.

LÖSUNG: **(a)** Recall, that the function $1_{[99,101]}$ is the charactersitic function of the interval $[99, 100]$. The mean value at the time $T = 1$ of the option on the given stock is:

$$\mathbb{E}[\ (X-A)_+\ ] = \int_{\mathbb{R}} (x-A)_+\ \rho(x)\ dx\ .$$

Here we use the notation:

$$(x-A)_+ := \begin{cases} x-A & \text{for } x-A \geqslant 0\ , \\ 0 & \text{for } x-A \leqslant 0\ . \end{cases}$$

In case of a density as in **(a)** we have to compute the value of the constant C, so that the probability one condition for the whole space is satisfied:

$$\begin{aligned} 1 &= \int_{\mathbb{R}} \rho(x)\ dx \\ &= \int_{\mathbb{R}} C\ (x-100)^2\ 1_{[99,101]}\ dx = \int_{[99,101]} C\ (x-100)^2\ dx \\ &= \left[\ \frac{C}{3}(x-100)^3\ \right]_{99}^{101} = \frac{2C}{3}\ . \end{aligned}$$

It follows $C = 3/2$. `Maple`–Check:

```
rho := x-> piecewise(
             x>99 and x<101 , C*(x-100)^2 ,
                              0
           );
int( rho(x) , x=-infinity..+infinity );
```

We can now compute the value in $T = 1$ of the option:

$$\begin{aligned} \mathbb{E}[\ (X-A)_+\ ] &= \int_{\mathbb{R}} (x-A)_+\ \rho(x)\ dx \\ &= \int_{\mathbb{R}} (x-A)_+\ \rho(x)\ dx \\ &= \int_{\mathbb{R}} (x-100.5)_+\ \frac{3}{2}\ (x-100)^2\ 1_{[99,101]}\ dx = \int_{[100.5,101]} (x-100.5)\ \frac{3}{2}\ (x-100)^2\ dx \\ &= \frac{3}{2}\ \int_{[100.5,101]} ((x-100)-0.5)\ (x-100)^2\ dx = \frac{3}{2}\ \left[\ \frac{1}{4}(x-100)^4 - 0.5\frac{1}{3}(x-100)^3\ \right]_{100.5}^{101} \\ &= \frac{3}{2}\ \left[\ \frac{1}{4}(1^4-0.5^4) - 0.5\frac{1}{3}(1^3-0.5^3)\ \right]\ . \end{aligned}$$

■

Aufgabe 11 *A person A (age $t_0 = 55$ years) monthly pays the following five years the amount of 200 (EURO) to an insurance company I in the time interval $[t_0, t_1) := [55, 60)$. (In case of death there is no contractual obligation of a continuation of payments by some other person.)*
The next five years, time interval $[t_1, t_2) = [60, 65)$, there is no payment of either part.
In the time interval $[t_2, \infty) = [65, \infty)$ there are monthly payments of the rent R to the person A, as long as A lives.
Which is the value of R ?
There is the following repartition $F(T)$ of the event
"Person A lives at t_0 and dies before T, $T \geqslant t_0 = 55$":

- $F(55) = 0$,
- $F(60) = 1/6$ *and F is linear in interval* $[55, 60)$,
- $F(65) = 1/2$, *and F is linear in interval* $[60, 65)$,
- *F is of the form* $F(T) = 1 - A\exp(-x^2/2)$ *on the interval* $[65, \infty)$

LÖSUNG:

Aufgabe 12 *There are $4 \times 13 = 52$ cards, of four colors ♣, ♢, ♡, ♠ and with the 13 labellings $2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A$. There are four players. Each player gets five cards. Which is the probability for each of the following events:*
(a) *The first player has two pairs, i.e. a constellation of cards of the shape $XXYYZ$ (after ordering) with different X, Y, Z among $2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A$.*
(b) *All players have five different cards.*
(c) *The second player has a flush (labellings 23456 or 34567 or $\dots$ or $10\ JQKA$)* **and** *the third has three in a color ($XXXYZ$ with different X, Y, Z). Are the events "The second player has a flush" and "The third player has three in a color" independent ?*

LÖSUNG:

Aufgabe 13 *Two dices roll on a table. Let A, B, C be the events:*
$A =$ the first dice shows an even number
$B =$ the second dice shows an odd number
$C =$ the sum of the two numbers on the two dices is an even number.
Are the events A, B, C pairwise independent ?
Are the events A, B, C independent ?

LÖSUNG: We model the given experiment by the space $\Omega := \{\ (i,j)\ :\ 1 \leqslant i,j \leqslant 6\ \}$ with $6 \times 6 = 36$ elements, each elementary event (consisitng of one element) having the same probability $1/36$.
We denote by $|X|$ the number of elements of the set X.
We compute for

$$
\begin{aligned}
A &:= \{\ (i,j) \in \Omega\ :\ i \text{ is even }\ \} \\
B &:= \{\ (i,j) \in \Omega\ :\ j \text{ is odd }\ \} \\
C &:= \{\ (i,j) \in \Omega\ :\ i+j \text{ is even }\ \}
\end{aligned}
$$

the following probabilities:

$$
\begin{aligned}
\mathbb{P}(A) &= \frac{|A|}{|\Omega|} = \frac{1}{2}\ , \\
\mathbb{P}(B) &= \frac{|B|}{|\Omega|} = \frac{1}{2}\ , \\
\mathbb{P}(C) &= \frac{|C|}{|\Omega|} = \frac{1}{2}\ , \\
\mathbb{P}(A \cap B) &= \frac{|A \cap B|}{|\Omega|} = \frac{1}{4}\ , \\
\mathbb{P}(A \cap C) &= \frac{|A \cap C|}{|\Omega|} = \frac{1}{4}\ , \\
\mathbb{P}(B \cap C) &= \frac{|A \cap C|}{|\Omega|} = \frac{1}{4}\ , \\
\mathbb{P}(A \cap B \cap C) &= \frac{|A \cap B \cap C|}{|\Omega|} = \frac{|\emptyset|}{|\Omega|} = 0\ .
\end{aligned}
$$

... ■

Aufgabe 14 *Let X_1, X_2 be two standard normally distributed random variables. Compute the* FOURIER*–transformations of (the repartitions) of X, Y:*

$$
F_{X_1}(t) := \mathbb{E}[\ e^{itX_1}\ ]\ , \qquad F_{X_2}(t) := \mathbb{E}[\ e^{itX_2}\ ]\ .
$$

Recall: If ρ is the density of a random variable X, then the FOURIER*–transformation of X is:*

$$
F_X(t) := \mathbb{E}[\ e^{itX}\ ] = \int_{\mathbb{R}} e^{itx}
$$

Lösung: