Lecture Notes by Jürgen Vollmer

Stochastic Processes

— Working Copy –

— 2020-04-12 03:32:16+02:00—

Fakultät für Physik und Geowissenschaften, Universität Leipzig

LECTURES DELIVERED AT FAKULTÄT FÜR PHYSIK UND GEOWISSENSCHAFTEN, UNIVERSITÄT LEIPZIG

http://www.uni-leipzig.de/physik/~vollmer

Die Philosophie steht in diesem großen Buch geschrieben, dem Universum, das unserem Blick ständig offen liegt. Aber das Buch ist nicht zu verstehen, wenn man nicht zuvor die Sprache erlernt und sich mit den Buchstaben vetraut gemacht hat, in denen es geschrieben ist. Es ist in der Sprache der Mathematik geschrieben, und deren Buchstaben sind Kreise, Dreiecke und andere geometrische Figuren, ohne die es dem Menschen unmöglich ist, ein einziges Wort davon zu verstehen; ohne diese irrt man in einem dunklen Labyrinth herum.

Galileo Galilei, Il Saggiatore, 1623

Die Mathematik ist das Instrument, welches die Vermittlung bewirkt zwischen Theorie und Praxis, zwischen Denken und Beobachten: sie baut die verbindende Brücke und gestaltet sie immer tragfähiger. Daher kommt es, daß unsere ganze gegenwärtige Kultur, soweit sie auf der geistigen Durchdringung und Dienstbarmachung der Natur beruht, ihre Grundlage in der Mathematik findet.

David Hilbert, Ansprache "'Naturerkennen und Logik"' am 8.9.1930 während des Kongresses der Vereinigung deutscher Naturwissenschafter und Mediziner

Insofern sich die Sätze der Mathematik auf die Wirklichkeit beziehen, sind sie nicht sicher, und insofern sie sicher sind, beziehen sie sich nicht auf die Wirklichkeit.

Albert Einstein Festvortrag "'Geometrie und Erfahrung"' am 27.1.1921 vor der Preußischen Akademie der Wissenschaften

Preface

Die ganzen Zahlen hat der liebe Gott geschaffen,
alles andere ist Menschenwerk.
Leopold Kronecker

These notes are a draft of notes for my course "Stochastic Dynamics" delivered at the University of Leipzig. At the moment they comprise a random collection of notes that were produced for other courses and remarks that *might* turn out useful to follow the lecture. I hope that in the course of time they will grow into a proper set of lecture notes. For the time being I recommend to also consult the books by Haken (1983) and Feller (1968) that provide excellent introductions into the topic from a physical and mathematical perspective, respectively. A good introduction that is somewhere in between these extreme perspective is provided by Garcia-Palacios (2007)

acknowledge co-workers

I am eager to receive feedback. It is crucial for the development of this project to learn about typos, inconsistencies, confusing or incomplete explanations, and suggestions for additional material (contents as well as links to papers, books and internet resources) that should be added in forthcoming revisions. Everybody who is willing to provide feedback will be invited to a coffee in Café Corso.

Contents

1

Probabilities

For the purpose of this lecture the results of an experiment will be regarded as a set[1] X of mutually distinct outcomes $x \in X$. When the cardinality of the set X is either finite or countable one can characterize the likelihood to encounter a given outcome x by a probability p_x. When the outcome of the observation is a real number, for instance when looking for the position of a particle on the real line, then X is not countable, and the probability will be provided in terms of a probability density. In general, probabilities will be described in terms of a relation between elements of a sigma algebra of possible outcomes of the experiment and real numbers in the interval $[0, 1]$ that provide the associated probability to find such an event. We discuss the three approaches one after the other.

[1] The pertinent mathematical notion of sets and their properties are summarized in Appendix A.

1.1 A countable set of outcomes

Definition 1.1: Probability

Let X be a finite or countable set of distinct results of an experiment. It will be denoted *sample space*. Then we assign the *probability* p_x to an element $x \in X$ when the number n_x of encounters of x in N repetition of the experiment converges towards p_x, i.e. $\lim_{N\to\infty} \frac{n_x}{N} = p_x$.

Remark 1.1. By construction probabilities are normalized, $\sum_{x\in X} p_x = 1$.

1.1.1 Examples: Flipping Coins and Throwing Dice

Coin flipping refers to an experiment where a coin is flipped into the air, picked up, and it is recorded which side is recognized as the top side after the kick.

Example 1.1 (One-sided Coin). Some coins show head on both sides. In that case the outcome is always head, h, the set of outcomes has a single element, $X_1 = \{h\}$, and the probability for that outcome is one, $p_h = 1$. ♢

Example 1.2 (Fair Coin Toss). When one can distinguish the two sides of the coin, head h and tail t, the set of outcomes has two elements, $X_1 = \{h, t\}$, and their probabilities sum up to one, $p_h + p_t = 1$. Often the flipping is arranged in such a way that $p_h = p_t = 1/2$. The relation between frequencies of observations and probabilities for coin tossing is further elaborated in worksheet 1. ♢

A typical die comprises a cubic body with its six sides marked by the numbers $\{1, \ldots, 6\}$. Throwing dice refers to an experiment where one of several dice set in motion, and after they came to rest it is noted which numbers are shown on their top sides.

Example 1.3 (Throwing Dice). For a single die the set of outcomes has six elements, $X_D = \{1, \ldots, 6\}$, and their probabilities sum up to one, $\sum_{i=1}^{6} p_i = 1$. The die throwing is called fair, when $\forall i : p_i = 1/6$. Otherwise, we refer to the die as a *loaded die*.[2] ♢

[2] The notion of a loaded die refers to the observation that a certain outcome, is encountered with a higher probability than $1/6$ when the center of mass of the die is not located in the center but rather close to the side opposing the preferred outcome.

Example 1.4 (Board games). In German board games a stone is commonly moved by trowing a die, moving forward by the encountered number of steps, and being allowed to throw again and move again if one encountered a six. In principle, by thowing n times six and then $m \in \{1, \ldots, 5\}$ one can reach any number of steps $N = 6n + m \in \mathbb{N}$. This example is discussed in worksheet 3. ♢

1.1.2 *Joint probabilities and conditional probabilities*

In some games we through two or even three dice at a time, and in some circumstances we might be interested in the outcomes of N successive coin flips.

Definition 1.2: Joint Probability

Let $X = X_1, \ldots, X_N$ be the sample spaces of N joint observations. Then the *joint probability* $P(x_1, \ldots, x_N)$ provides the probability to encounter the values $x_1 \in X_1, \ldots, x_N \in X_N$ in a given observation.

For throwing dice the different die behave *independently* from each other, In such a case the joint probability amounts to the product of the probabilities of the individual events.

Definition 1.3: Independent Probabilities

Let X and Y be the possible outcomes of two experiments with probability distribution $P_X(x)$ and $P_Y(y)$, respectively. Then the experiments are *independent* iff the joint distribution $P_{X,Y}(x,y)$ factorizes, i.e., when $P_{X,Y}(x,y) = P_X(x)\, P_Y(y)$.

Counterexample 1.5 (Colors of Balls selected from a bag). There are N_R red balls and N_G green balls in a bag, i.e. $N = N_R + N_G$ balls in total. When we draw a ball at random its color will be a red with probability $P_1(r) = N_R/N$ and green with probability $P_1(g) = N_G/N$. Let now $P_2(xy)$ with $x, y \in \{r, g\}$ be the joint probability of two balls drawn from the bag. When the first ball is red the probability to also draw another red ball is $(N_R - 1)/(N - 1)$. Otherwise, the probability that it is red is $N_R/(N - 1)$. In any case the second probability is different from the first one because there is a different number of balls in the bag. It is different from the product of the probability The problem is further discussed in Problem 1.1. ♢

The probability for the color of the second ball drawn from the bag is most conveniently described by a conditional probability.

Definition 1.4: Conditional Probabilities

The *conditional probability* $P(x|y)$ describes the probability to encounter y under the condition that we also observe x in another experiment. Therefore,

$$P(x,y) = P(x)\, P(x|y) \tag{1.1.1}$$

When the outcome of y does not depend on the condition x then events are independent and $P(x|y) = P(y)$. Consequently, $P(x,y) = P(x)\, P(x|y) = P(x)\, P(y)$ as stated in Definition 1.3.

Summing over x and y provides 1 by normalization of the probability,

$$\sum_{x \in X} \sum_{y \in Y} P(x,y) = 1$$

Summing over only x or y recovers the probability of the other observable

Definition 1.5: Marginal Probability Distribution

A *marginal probability distribution* is obtained from a joint distribution $P_{X\times Y}(x,y)$ by summing over all events that observe a certain constraint. In particular, the sum over all values (x,y) with a given value of x comes down to

$$\sum_{y\in Y} P_{X\times Y}(x,y) = P_X(x)$$

and for (x,y) with a given value of y

$$\sum_{x\in X} P_{X\times Y}(x,y) = P_Y(y)$$

Remark 1.2. Often the probability distribution of a some events can be calculated most straightforwardly by formulating them as a marginal probability distribution of a joint probability that is more easily accessible. For instance the probability to have a sum s as a the result of throwing two dice can easily be obtained from their joint probability distribution $P(n_1,n_2)$ that the dice show n_1 and n_2 eyes, respectively,

$$P(s) = \sum_{n_1=1}^{6} P(n_1, s-n_1) = \sum_{n_1=1}^{6} P_1(n_1)\, P_1(s-n_1)$$

where $P_1(n)$ takes the value $1/6$ when $1 \leq n \leq 6$ and otherwise it is zero.

1.1.3 *Moments and Cumulants of a distribution*

For the setting of Example 1.4 one might be interested in how far one can move on average in each round of the game, whether one will typically move more or less the same distance, and whether outliers lie rather towards small or large distances. These questions are addressed by the moments and cumulants of the distribution. In all these settings x must characterize a distance. Typically it is sampled from a set $X \subset \mathbb{R}$.

Definition 1.6: Moments

Let $P_X(x)$ be a probabilitiy distribution with values $x \in X \subset \mathbb{R}$. Then the ν^{th} *moment* M_ν of the distribution is defined as

$$M_\nu = \sum_{x \in X} x^\nu \, P_X(x) \equiv \langle x^\nu \rangle_X$$

where $\langle . \rangle_X$ introduces a shorthand for the expression with the sum. The first moment is also called *expectation* or *mean value* of the distribution.

Remark 1.3. The zeros moment M_0 always takes the values 1 due to the normalization of the probability distribution.

Remark 1.4. This definition can be generalized in straightforward manner to metric spaces, i.e. settings where the absolute value of x is defined with $|x| \in \mathbb{R}$,

$$M_\nu = \sum_{x \in X} |x|^\nu \, P_X(x) = \langle |x|^\nu \rangle_X$$

Example 1.6 (Moments of tossing dice). For throwing a die the first moments take the values[3]

$$\begin{aligned}
M_1 &= \frac{1}{6}\,[1+2+3+4+5+6] = \frac{1}{6}\frac{6^2+6}{2} = \frac{7}{2} \\
M_2 &= \frac{1}{6}\,\left[1^2+2^2+3^2+4^2+5^2+6^2\right] = \frac{1}{6}\frac{2\cdot 6^3+3\cdot 6^2+6}{6} = \frac{91}{6} \\
M_3 &= \frac{1}{6}\,\left[1^3+2^3+3^3+4^3+5^3+6^3\right] = \frac{1}{6}\frac{6^4+2\cdot 6^3+6^2}{4} = \frac{147}{2} \\
M_4 &= \frac{1}{6}\,\left[1^4+2^4+3^4+4^4+5^4+6^4\right] = \frac{2275}{6}
\end{aligned}$$

At least in principle, all integer moments can be calculated. ♢

[3] The results of the sums are provided by Faulhaber's formula.

Example 1.7 (Flip and go). We consider a process very similar to the board games. We move as follows based on tosses of a fair coin: Stay when we hit tail; with probability 1/2. Move one step and flip again when we hit head; with probability 1/2 in each step. The probability to take n steps is therefore $P(n) = 2^{-n-1}$ with $n \in \mathbb{N}_0$. The ν^{th} moments of this distribution take the values

$$M_\nu = \sum_{n \in \mathbb{N}_0} n^\nu \, 2^{-n-1}$$

One can look up the first few values, $M_1 = 1$, $M_2 = 3$, $M_3 = 13$, $M_4 = 75$, but the sum can not readily be avaluated in a closed form. ♢

The evaluation of moments can greatly be simplified by moment generating functions and characteristic functions.

Definition 1.7: Moment generating functions

The moment generating function $G(q)$ and characteristic function $F(k)$ for a probability distribution $P(x)$ are defined as

$$G(q) = \sum_{x \in X} e^{q\,x}\, P(x) = \langle e^{q\,x} \rangle_X \tag{1.1.2}$$

$$F(k) = \sum_{x \in X} e^{i\,k\,x}\, P(x) = \langle e^{i\,k\,x} \rangle_X \tag{1.1.3}$$

The ν^{th} moments of the distribution $P(x)$ is related to the ν^{th} derivative of these functions,

$$M_\nu = \frac{d^\nu}{dq^\nu} G(q) \bigg|_{q=0} = i^{-n} \frac{d^\nu}{dk^\nu} F(k) \bigg|_{q=0}$$

Remark 1.5. The two functions are related by $F(k) = G(ik)$ such that the choice is a matter of taste and convenience.

When the distribution has a non-trivial first moment, one is often more interested in the deviation from this value than in the moments themself.

Definition 1.8: Centered Moments

Let $P_X(x)$ be a probabilitiy distribution with values $x \in X \subset \mathbb{R}$ and expectation $\bar{x}$. Then the ν^{th} *centered moment* M_ν^c of the distribution is defined as

$$M_\nu^c = \sum_{x \in X} (x - \bar{x})^\nu\, P_X(x) = \langle (x - \bar{x})^\nu \rangle_X$$

The second centered moment of the distribution is called *variance*, Var_X.

Remark 1.6. The first centered moment vanishes by construction, and the second centered moment is related to the moments by

$$\text{Var}_X = \langle (x - \bar{x})^2 \rangle_X = \langle x^2 \rangle - \bar{x}^2 = M_2 - M_1^2$$

Further the distribution can also be characterized by its cumulants. That are best introduced as derivatives of the cumulant generating function.

Definition 1.9: Cumulants and cumulant generating functions

The cumulant generating function $C(q)$ for a probability distribution $P(x)$ is defined as the logarithm of the moment generating function $M(q)$ of the distribution

$$C(q) = \ln G(q) \tag{1.1.4}$$

Its ν^{th} derivatives evaluated at $q = 0$ define the cumulants of the distribution,

$$\mu_\nu = \frac{\mathrm{d}^\nu}{\mathrm{d}q^\nu} C(q) \bigg|_{q=0}$$

Remark 1.7. The first cumlants take the values

$$\begin{aligned}
\mu_0 &= C(0) = 0 \\
\mu_1 &= \frac{G'(q)}{G(q)} \bigg|_{q=0} = M_1 \\
\mu_2 &= \frac{G(q)\, G''(q) - \big(G'(q)\big)^2}{\big(G(q)\big)^2} \Bigg|_{q=0} = M_2 - M_1^2 = \mathrm{Var}_X
\end{aligned}$$

1.2 *Probability Distributions*

One can no longer assign probabilities to specific values when x is sampling the positions in the continuum, taking real values in an interval or area. In that case the attributions of probabilities is commonly based on a probability density $P(x)$.

Definition 1.10: Probability densities

Let the sample space X be a compact set of distinct results of an experiment. Then we assign the *probability density* $P(x)$ assings the probability

$$p_{[a,b]} = \int_a^b \mathrm{d}x\, P(x)$$

to the interval $[a, b]$ when the number $n_{[a,b]}$ of results that lie in the interval in N repetitions of the experiment converges to $p_{[a,b]}$, i.e.

$$\lim_{N\to\infty} \frac{n_{[a,b]}}{N} = p_{[a,b]} \,.$$

Remark 1.8. Normalization of the probability entails that

$$\int_X \mathrm{d}x\, P(x) = 1$$

Remark 1.9. Rather than as a functions, the densities $P(x)$ should be considered as distributions. The probability distribution of the die can then be written in terms of six delta function,

$$P(x) = \frac{1}{6} \sum_{n=1}^{6} \delta(x-n)$$

With this understanding the definitions of the moments, cumulants and their generating functions can immediately be inferred,

$$M_\nu = \int_X \mathrm{d}x\, x^\nu\, P(x) = \langle x^\nu \rangle_X \tag{1.2.1}$$

$$G(q) = \int_X \mathrm{d}x\, \mathrm{e}^{q\,x}\, P(x) = \langle \mathrm{e}^{q\,x} \rangle_X \tag{1.2.2}$$

$$C(q) = \ln G(q) = \ln \langle \mathrm{e}^{q\,x} \rangle_X \tag{1.2.3}$$

Marginal densities can be evaluated by introducing the constraint by means of a delta function. In particular, we will have

$$P(x) = \int \mathrm{d}x'\, \mathrm{d}y'\, \delta(x-x')\, P(x',y')$$

$$P(y) = \int \mathrm{d}x'\, \mathrm{d}y'\, \delta(y-y')\, P(x',y')$$

$$P(R) = \int \mathrm{d}x\, \mathrm{d}y\, \delta\left(R - \sqrt{x^2+y^2}\right)\, P(x,y)$$

1.3 Formal introduction of probabilities

1.4 Problems

1.4.1 Rehearsing Concepts

Problem 1.1. **Colors of Balls selected from a bag**

For the setting of Example 1.5:

a) Determine the conditional probability $P(x|y)$ that the second ball takes the color y when initially we draw a ball with color x. Employ Equation (1.1.1) to show that the joint probability takes the values

$$P_2(rr) = \frac{N_R\,(N_R-1)}{N\,(N-1)}, \quad P_2(gg) = \frac{N_G\,(N_G-1)}{N\,(N-1)}$$
$$P_2(rg) = P(gr) = \frac{N_R\,N_G}{N\,(N-1)}\,.$$

b) Verify that the probabilities are normalized.

c) Compare $P_2(xy)$ and $P_1(x)\,P_1(y)$.

Problem 1.2. **Uncorrelated vs. independent variables**

We consider the following probability distribution

$$p(x,y) = c\,(x^2+y^2) \qquad \text{for} \quad (x,y) \in [-1,1]\times[-1,1]\,.$$

a) Determine the normalization constant.

b) Determine the marginal probabilities $p_1(x) = \int \mathrm{d}y\, p(x,y)$ and $p_2(y) = \int \mathrm{d}x\, p(x,y)$.

c) Determine the conditional probabilities $p(x|y)$ and $p(y|x)$.

d) Determine the expectation values $\langle x\rangle$, $\langle y\rangle$, and $\langle x\,y\rangle$.

Problem 1.3. **Cumulants and centered moments**

a) Verify that the third cumulant of a distribution always agrees with the third centered moment.

b) Show that the centered moments and cumulants differ for orders higher than and equal to four.

c) Determine the centered moments and the cumulants for the normal distribution.

1.4.2 Practicing Concepts

Problem 1.4. **Marginal Probabilities for the 2d Gaussian Distribution**

We consider a Gaussian distribution in the two variables,

$$P(x,y) = \frac{1}{2\pi\sigma^2}\, \mathrm{e}^{-(x^2+y^2)/2\sigma^2} \qquad \text{with} \quad (x,y) \in \mathbb{R}^2 .$$

a) Determnine the marginal probability densities $p_1(x) = \int \mathrm{d}y\, P(x,y)$ and $p_2(y) = \int \mathrm{d}x\, P(x,y)$.

b) Demonstrate that the conditional probability density $P(x|y) := P(x,y)/p_2(y)$ amounts to $p_1(x)$. Is there a faster way to see that x and y are independent variables?

c) Determine the probability density $p_3(R)$ to find a value (x,y) with modulus $R = \sqrt{x^2+y^2}$.

d) Determine the probability density $p_4(\phi)$ to find a value (x,y) in direction ϕ with respect to the positive x axis.

Problem 1.5. **The characteristic function of the Lorentz-Cauchy distribution**

The Lorentz-Cauchy distribution is defined as

$$p_{LC}(x) = \frac{1}{\pi}\, \frac{\Gamma}{(x-m)^2+\Gamma^2} \qquad \text{with parameters } m,\Gamma \in \mathbb{R} \tag{1.4.1a}$$

a) Verify that the distribution is normalized.

b) Determine the expectation value of the distribution.
What about the variance?

c) Show that the characteristic function of $p_{LC}(x)$ is

$$\chi_{LC}(t) = \left\langle \mathrm{e}^{\mathrm{i}\,t x} \right\rangle = \mathrm{e}^{\mathrm{i}\,m\,t - \Gamma\,|t|} \tag{1.4.1b}$$

What does this tell about the normalization, expectation and variance?

(bonus) Consider the distribution

$$p_4(x) = \frac{4}{\pi}\, \frac{\Gamma^3}{(x-m)^4+4\,\Gamma^4} \qquad \text{with parameters } m,\Gamma \in \mathbb{R} \tag{1.4.1c}$$

Determine the characteristic function.

Show that it provides m and $2\Gamma^2$ for the expectation and the variance, respectively.
What happens for higher moments?

1.4.3 Mathematical Background

Problem 1.6. σ **algebras for throwing dice**
We consider the outcomes of rolling a die with faces $\Omega = \{1, 2, 3, 4, 5, 6\}$. The die is a cube with three independent axes that correspond to the sides $X = \{1, 6\}$, $Y = \{2, 5\}$, and $Z = \{3, 4\}$, respectively.

a) Construct the σ-algebra that admits distinction *only* of the axes.

b) Provide a generating set for this σ algebra.

c) We throw two dice and are interested in the overall sum of points. What are the mutually exclusive sets of events that can be used to generate the algebra for this problem? What are probabilities of the events?

d) We again throw two dice, but we are only interested in doubles, i.e., outcomes where both dice show the same number of points. How is this problem related to Russion roulette?

Problem 1.7. **Probabiliy of non-exclusive events**
The events X, Y and Z in exercise 1.6 are mutually exclusive, and for a fair die they appear with equal probability.

a) What is the probability for the events $A = \{X, Y\}$ that either X or Y is encountered, and for $B = \{X, Z\}$ that either X or Z is encountered. Provide an intuitive and a formal argument.

b) What are the probabilities for $\Pi(A \cup B)$ and $\Pi(A \cap B)$?

c) Use the axioms for probabilities to derive a general relation between the probabilities $\Pi(M_1)$, $\Pi(M_2)$, $\Pi(M_1 \cup M_2)$, and $\Pi(M_1 \cap M_2)$.

1.4.4 Transfer and Bonus Problems, Riddles

Problem 1.8. **Bertrand's paradox**
A line is dropped randomly on a circle. The intersection will be a chord. What is the probability that the length of the chord is larger than that of a side of the inscribed equilateral triangle?

Bertrand's paradox states that the resulting probabiliy depends on the rule how the lines are selected. To understand this observation we consider intersections with the unit circle at the origin.

a) Specify the line by two points: The first point is $P_1 = (0, 1)$ on the circle and a second point $P_2 = (x, y)$ is selected at random in $\mathbb{R}^2$. Determine the probability distribution for the length of the chords, and its expectation value.
   Hint: What is the probability density to find the second point in a direction ϕ with respect to the x-axis, when looking from P_1?

b) Pick as first point $P_1 = (0, 1)$ on the circle, as before. However, now the second point $P_2 = (x, y)$ is selected at random in $[-L, L] \times [-L, L]$ for some $L \in \mathbb{R}^+$.
   Determine the probability distribution for the length of the chords and its expectation value.
   How does the result depend on L?

c) Employ rotational symmetry and only consider horizontal lines (x, y_h) with $x \in \mathbb{R}$ and fixed y_h. Let y_h be uniformly distributed in $[-1, 1]$.
   Determine the probability distribution for the length of the chords and its expectation value.

d) Write a Python program to support your findings and explore other settings.

Problem 1.9. **Total sum on N dice**

We throw N dice and explore the distribution $P_N(n)$ that the sum of points of their point is n. For sake of simplifying notations we assume $n \in \mathbb{N}$ and assign zero probabilities where appropriate.

a) In Example 1.3 we obtained that

$$P_1(n) = \begin{cases} 1/6 & \text{for } \; n \leq 7 \\ 0 & \text{else} \end{cases}$$

   Proof the recursion

$$P_{N+1}(n) = \sum_{m=1}^{n} P_1(m)\, P_N(n-m)$$

   and show that the resulting distributions are normalized.

b) Determine $P_2(n)$ and compare to the data in worksheet 2. Expand the worksheet by the recursion relation, and explore how

$P_N(n)$ approaches a normal distribution for large N.

c) Calculate the first and second moment of distribution.
Hint: Use the recursion relation. Do you see how the derivation can be expanded to calculate moments of higher orders?

A

Basic notions of set theory

In this chapter we provide a brief introduction into relevant aspects of set theory, and how relations, and functions are defined on this background. These mathematical notions will be employed when providing a formal definition of *probability* and *probability distributions*.

A.1 Sets

In mathematics and physics we often wish to make statements about a collection of objects, numbers, or other distinct entities.

Definition A.1: Set

A *set* is a gathering of well-defined, distinct objects of our perception or thoughts.
An object a that is part of a set A is an *element* of A; we write $a \in A$.
If a set M has a finite number n of elements we say that its *cardinality* is n. We write $|M| = n$.

Remark A.1. Notations and additional properties:

a) When a set M has a finite number of elements, e.g., $+1$ and -1, one can specify the elements by explicitly stating the elements, $M = \{+1, -1\}$. The order does not play a role and it does not make a difference when elements are provided several times. In other words the set M of cardinality two can be specified by any of the following statements

$$M = \{-1, +1\} = \{+1, -1\} = \{-1, 1, 1, 1, \} = \{-1, 1, +1, -1\}$$

b) If e is not an element of a set M, we write $e \notin M$. For instance

$-1 \in M$ and $2 \notin M$.

c) There is only one set with no elements, i.e., with cardinality zero. It is denoted as $\emptyset$.

Example A.1 (Sets).

- Set of capitals of German states:
  $A_2 = \{$Berlin, Bremen, Hamburg, Stuttgart, Mainz, Wiesbaden, München, Magdeburg, Saarbrücken, Potsdam, Kiel, Hannover, Dresden, Schwerin, Düsseldorf, Erfurt$\}$
- Set of small letters in German:
  $A_3 = \{$a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z, ä, ö, ü, ß$\}$
- Set of month with more than 28 days:[1]
  $A_1 = \{$January, February, March, April, May, June, July, August, September, October, November, December$\}$
  The cardinalities of these sets are
  $|A_1| = 12$, $|A_2| = 16$, and $|A_3 = 30|$. ◇

[1] Most of them even have more days.

Example A.2 (Sets of sets). A set can be an element of a set. For instance the set

$$M = \{1, 3, \{1, 2\}\}$$

has three elements 1, 3 and $\{1, 2\}$ such that $|M| = 3$, and

$$1 \in M\,, \qquad \{1, 2\} \in M\,, \qquad 2 \notin M \qquad \{1\} \notin M\,.$$

◇

Often it is bulky to list all elements of a set. In obvious cases we use ellipses such as $A_3 = \{$a, b, c, $\dots$, z, ä, ö, ü, ß$\}$ for the set given in Example A.1. Alternatively, one can provide a set M by specifying the properties of its elements x in the following form

$$\underbrace{M}_{\text{The set } M} \underbrace{=}_{\text{contains}} \underbrace{\{}_{\text{all elements}} \underbrace{x}_{x,} \underbrace{:}_{\text{with :}} \underbrace{A(x)}_{\text{properties}\dots} \}.$$

where the properties specify one of several properties of the elements. The properties are separated by commas, and must all be true for all elements of the set.

Example A.3 (Set definition by property). The set of digits $D = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ can also be defined as follows $D = \{x : 0 < x \leq 9, x \in \mathbb{Z}\}$ or even $D = \{1, \dots, 9\}$. ◇

In order to specify the properties in a compact form we use logical junctors as short hand notation. In the present course we adopt the notations not $\neg$, and $\wedge$, or $\vee$, implies $\Rightarrow$, and is equivalent $\Leftrightarrow$ for the relations indicated in A.1.

A	0	0	1	1	
B	0	1	0	1	
$\neg A$	1	1	0	0	not A
$\neg B$	1	0	1	0	not B
$A \vee B$	0	1	1	1	A or B
$A \wedge B$	0	0	0	1	A and B
$A \Rightarrow B$	1	1	0	1	A implies B
$A \Leftrightarrow B$	1	0	0	1	A is equivalent to B
$A \vee \neg B$	1	0	1	1	A or not B
$\neg A \wedge B$	0	1	0	0	not A or B
$A \wedge \neg B$	0	0	1	0	A and not B

Table A.1: List of the results of different junctors acting on two statements A and B. Here 0 and 1 indicate that a statement is wrong or right, respectively. In the rightmost column we state the contents of the expression in the left column in words. The final three lines provide examples of more complicated expressions.

The definition of the digits in Example A.3 entails that all elements of D are also numbers in $\mathbb{Z}$: we say that D is a subset of $\mathbb{Z}$.

Definition A.2: Subsets and Supersets

The set M_1 is a *subset* of M_2, if all elements of M_1 are also contained in M_2. We write[2] $M_1 \subseteq M_2$. We denote M_2 then as *superset* of M_1, writing $M_2 \supseteq M_1$.
The set M_1 is a *proper subset* of M_2 when at least one of its elements is not contained in M_2. In this case $|M_1| < |M_2|$ and we write $M_1 \subset M_2$, or $M_2 \supset M_1$.

[2] Some authors use $\subset$ instead of $\subseteq$, and $\subsetneq$ to denote proper subsets.

Example A.4 (Subsets).
- The set of month with names that end with "ber" is a subset of the set A_2 of Example A.1

$$\{\text{September, October, November, December}\} \subseteq A_3$$

- For the set M of Example A.2 one has

$$\{1\} \subseteq M\,, \qquad \{1, 3\} \subseteq M\,, \qquad \{1, 2\} \not\subseteq M\,, \qquad \{2, \{1, 2\}\} \not\subseteq M\,.$$

Note that $\{1, 2\}$ is an elements of M. However, it is not a subset. The last two sets are no subsets because $2 \notin M$.

$\diamondsuit$

Two sets are the same when they are subsets of each other.

Theorem A.1: Equivalence of Sets

Two sets A and B are *equal* or *equivalent*, iff

$$(A \subseteq B) \wedge (B \subseteq A).$$

Proof. $A \subseteq B$ implies that $a \in A \Rightarrow a \in B$.
$B \subseteq A$ implies $b \in B \Rightarrow b \in A$.
If $A \subseteq B$ and $B \subseteq A$, then we also have $a \in A \Leftrightarrow a \in B$. □

Remark A.2. In the logical $\in$ and $\notin$ are always evaluated with higher priority than junctors like $\Rightarrow$, $\Leftrightarrow$, $\wedge$, and $\vee$.

The description of sets by properties of its members, Example A.3, suggests that one will often be interested in operations on sets. For instance the odd and even numbers are subsets of the natural numbers. Together odd and even numbers form the set of natural numbers. One is left with the even numbers when removing the odd numbers from the natural numbers. We define the following operations on sets to formally express these statements.

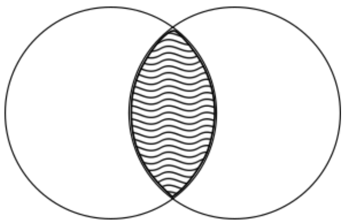


Figure A.1: Intersection of two sets.

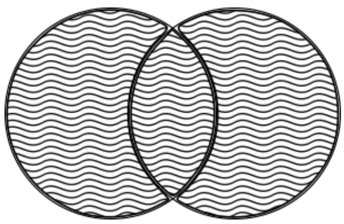


Figure A.2: Union of two sets.

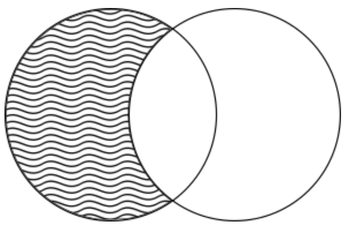


Figure A.3: Difference of two sets.

Figure A.4: Complement of a set.

Definition A.3: Set Operations

For two sets M_1 and M_2 we define the following operations:

- *Intersection*: $M_1 \bigcap M_2 = \{m \, : \, m \in M_1 \wedge m \in M_2\}$,
- *Union*: $M_1 \bigcup M_2 = \{m \, : \, m \in M_1 \vee m \in M_2\}$,
- *Difference*: $M_1 \backslash M_2 = \{m \, : \, m \in M_1 \wedge m \notin M_2\}$,
- The *complement* of a set M in a *universe* U is defined for subsets $M \subseteq U$ as follows $M^C = \{m \in U \, : \, m \notin M\}$.
- The *cartesian product of two sets* M_1 *and* M_2 is defined as the set of ordered pairs (a, b) of elements $a \in M_1$ and $b \in M_2$,
$$M_1 \times M_2 = \{(a, b) \, : \, a \in M_1, b \in M_2\}.$$

The oprations are graphically illustrated in Figures A.1 to A.4.

Example A.5 (Set operations for participants in my class). Consider the set of participants P in my class. The sets of female F and male M participants of the class are proper subsets of P with an empty intersection $F \bigcap M$. The set of non-female participants is $P \backslash W$. The set of heterosexual couples in the class is a subset of the Cartesian

name	symbol	description
natural numbers	$\mathbb{N}$	$\{1, 2, 3, \ldots\}$
natural numbers with 0	$\mathbb{N}_0$	$\mathbb{N} \bigcup \{0\}$
negative numbers	$-\mathbb{N}$	$\{-n \,:\, n \in \mathbb{N}\}$
even numbers	$2\mathbb{N}$	$\{2n \,:\, n \in \mathbb{N}\}$
odd numbers	$2\,\mathbb{N} - 1$	$\{2n - 1 \,:\, n \in \mathbb{N}\}$
integer numbers	$\mathbb{Z}$	$(-\mathbb{N}) \bigcup \mathbb{N}_0$
rational numbers	$\mathbb{Q}$	$\left\{\frac{p}{q} \,:\, p \in \mathbb{Z}, q \in \mathbb{N}\right\}$
real numbers	$\mathbb{R}$	see below
complex numbers	$\mathbb{C}$	$\mathbb{R} + \mathrm{i}\mathbb{R}$, where $\mathrm{i} = \sqrt{-1}$

Table A.2: Summary of important sets of numbers.

product $F \times M$. Furthermore, the union of the union of $W \bigcup M$ is a proper subset of P, when there is at least one participant who is neither female nor male. ♢

A.1.1 Sets of Numbers

Many sets of numbers that are of interest in physics have infinitely many elements. We construct them in Table A.2 based on the natural numbers

$$\mathbb{N} = \{1, 2, 3, \ldots\}$$

or the natural numbers with zero

$$\mathbb{N}_0 = \mathbb{N} \bigcup \{0\} \; .$$

Remark A.3. Some authors adopt the convention that zero is included in the natural numbers $\mathbb{N}$. In case of doubt one must check which convention is adopted.

There are many more sets of numbers. For instance, in mathematics the set of *constructable numbers* is relevant for certain proofs in geometry, and in physics we occasionally use *quaternions*. In any case one needs intervals of numbers.

Definition A.4: Intervals of real numbers $\mathbb{R}$

An *interval* is a continuous subset of a set of numbers. We distinguish *open, closed, and half-open subsets*.

- closed interval: $[a, b] = \{x : x \geq a, x \leq b\}$,
- open interval: $(a, b) =]a, b[= \{x : x > a, x < b\}$,
- right open interval: $[a, b) = [a, b[= \{x : x \geq a, x > b\}$,
- left open interval: $(a, b[=]a, b] = \{x : x > a, x \geq b\}$.

Subsets of $\mathbb{R}$ will be denoted as real intervals.

A.2 Problems

A.2.1 Rehearsing Concepts

Problem A.1. Es gibt vier paarweise verschiedenen Elemente A, B, C, D. Verwenden Sie die Symbole $\in, \notin, \ni, \not\ni, \subset, \not\subset, \supset, \not\supset, =$ in den Kästchen, so dass wahre Aussagen entstehen.

(a) $\{A, B\} \ \square \ \{A, B, C\}$,

(b) $\{A\} \ \square \ B$,

(c) $\{\emptyset\} \ \square \ \emptyset$,

(d) $\{\{A\}\} \ \square \ \{\{A\}, \{B\}\}$,

(e) $A \ \square \ \{A, B, C\}$,

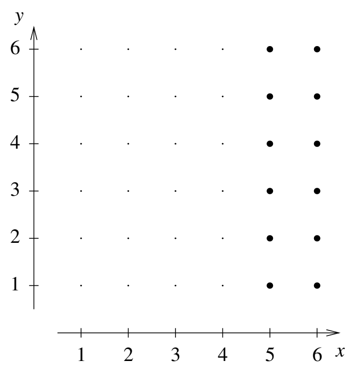
(f) $\{A, C, D\} \bigcap \{A, B\} \ \square \ \{A, B, C, D\}$

(g) $\{A, C, D\} \setminus \{A, B\} \ \square \ \{A, B, C\}$,

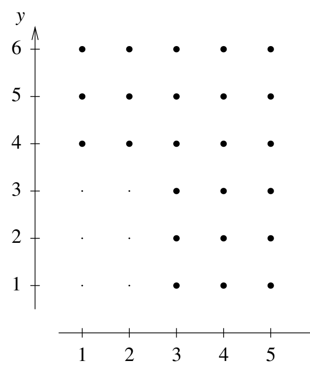
(h) $\{A, C, D\} \bigcup \{A, B\} \ \square \ A$.

Problem A.2. Wir betrachten hier die Menge $M := \{(x, y) \in \mathbb{Z} \times \mathbb{Z} : 1 \leq x, y \leq 6\}$. Bildlich kann man sie sich als Raster aus sechs mal sechs Punkten vorstellen. Die Abbildung unten zeigt die beiden Teilmengen M_1 und M_2 von M (dünne Punkte zählen nicht).

$M_1 \ \{(x, y) \in M : x \geq 5\}$

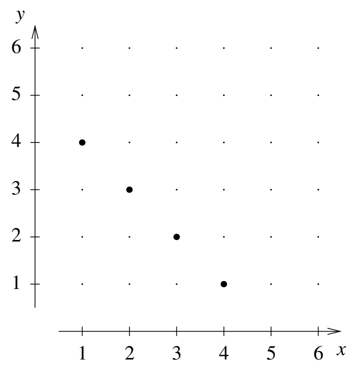


$M_2 \ \{(x, y) \in M : x \geq 3\} \bigcup \{(x, y) \in M : y \geq$

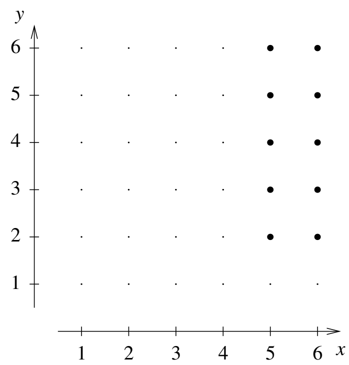


(a) Beschreiben Sie die in der Abbildung unten definierten Mengen M_3 und M_4:

M_3



M_4



(b) Zeichnen Sie die Menge $M_5 = \{(x,y) \in M \;:\quad 3 \mid x+y\ \}$.[3]

[3] Hier bedeutet $a \mid b$, dass *a ein Teiler ist von b*.

A.2.2 *Practicing Concepts*

Problem A.3. Schnittmengen.

Beschreiben Sie die folgenden Schnittmengen:

(a) $32\mathbb{Z} \bigcap 8\mathbb{Z}$. (b) $25\mathbb{Z} \bigcap 4\mathbb{Z}$. (c) $6\mathbb{Z} \bigcap 14\mathbb{Z}$.

(d) Stellen Sie eine allgemeine Regel für die Beschreibung der Schnittmenge $m\mathbb{Z} \bigcap n\mathbb{Z}$ auf.

Problem A.4. Grundlagen der Mengenlehre.

(a) Man zeige: $\{6z \,:\, z \in \mathbb{Z}\} \subset \{2z \,:\, z \in \mathbb{Z}\}$.

(b) Stellen Sie $[1;17] \bigcap]0;5[$ als Intervall dar. Begründen Sie Ihre Aussage!

(c) Stellen Sie $\bigcap_{n\in\mathbb{N}} [-\frac{1}{n}; 1+\frac{1}{n}]$ als Intervall dar. Begründen Sie Ihre Aussage!

(d) Man zeige: $\{2z \,:\, z \in \mathbb{Z}\} \bigcap \{3z \,:\, z \in \mathbb{Z}\} = \{6z \,:\, z \in \mathbb{Z}\}$.

(e) Untersuchen Sie, ob für alle $a, b \in \mathbb{N}$ gilt $T(a) \bigcup T(b) = T(a \cdot b)$, wobei $T(n)$ die Menge der Teiler der Zahl n ist.

Problem A.5. Intervalle.

(a) Beschreiben Sie $[-1,4] \setminus [1,2[$ als Vereinigung disjunkter Intervalle.

(b) Beschreiben Sie $[2,4[\bigcup ([3,10] \setminus (]3,4[\bigcup [6,7]))$ als Vereinigung disjunkter Intervalle.

(c) Stellen Sie die Menge

$$([1,3] \setminus]1,2[) \bigcup ([4,11] \setminus (]5,10[\setminus [6,7])) \bigcup \{9\}$$

als disjunkte Vereinigung von Intervallen oder Elementen dar.

Problem A.6. Teilmengen des $\mathbb{R}^2$ und $\mathbb{R}^3$.

(a) Schreiben Sie die Strecke zwischen den Punkten $(1;2)$ und $(4;0)$ als Teilmenge des $\mathbb{R}^2$.
(b) Man zeige: $\{(s,0) \,:\, s \in \mathbb{R}\} = \{(x,y) \in \mathbb{R}^2 \,:\, y = 0\}$.
(c) Man zeige: $\left\{ r \cdot \begin{pmatrix} 2 \\ 3 \end{pmatrix} \,:\, r \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \,:\, 3x - 2y = 0 \right\}$.
(d) Schreiben Sie die x-z-Ebene im $\mathbb{R}^3$ auf (mindestens) drei verschiedene Arten.
(e) Definieren Sie die quadratische Fläche mit den Eckpunkten $(2;3)$, $(2;5)$, $(4;3)$, $(4;5)$.
(f) Definieren Sie die Kreisfläche des Kreises mit Radius 3 um den Mittelpunkt $(1;-2)$.

A.2.3 Proofs

Problem A.7. Beweise mit Mengen.

(a) Es seien A, B Mengen. Man zeige: $A \setminus B \subset A$.
(b) Es seien A, B, C Mengen. Man zeige: $A \setminus (B \bigcup C) = A \setminus B \bigcap A \setminus C$.
(c) Es seien A, B, C Mengen. Formulieren Sie einen Beweis des Distributivitätsgesetzes

$$A \bigcap (B \bigcup C) = (A \bigcap B) \bigcup (A \bigcap C)$$

Überführen Sie dazu die Aussage $x \in A \bigcap (B \bigcup C)$ schrittweise, durch sukzessives Einsetzen der Definitionen und Verwendung elementarer logischer Schlüsse, in die Aussage $x \in (A \bigcap B) \bigcup (A \bigcap C)$. Tun Sie danach dasselbe in die entgegengesetzte Richtung, d.h. formen Sie die Aussage $x \in (A \bigcap B) \bigcup (A \bigcap C)$ schrittweise in die Aussage $x \in A \bigcap (B \bigcup C)$ um.

A.2.4 Transfer and Bonus Problems

Problem A.8. Wir betrachten die Menge $M = \{(x,y) \in M \,:\, y = (x-7)^2 + 1\}$. Geben Sie für diese Menge eine Darstellung in der Form

$$M = \{(x,y) \in M \,:\, a \cdot x + b \cdot y = c\} ,$$

wobei a, b und c (möglichst hübsch gewählte) reelle Zahlen sein sollen.

B

Bibliography

Feller, W., 1968, *An Introduction to Probability Theory and Its Applications*, volume 1 (Wiley, New York), 3 edition, ISBN 0471257087.

Garcia-Palacios, J. L., 2007, Introduction to the theory of stochastic processes and brownian motion, arXiv:cond-mat/0701242, URL `https://arxiv.org/abs/cond-mat/0701242`.

Haken, H., 1983, *Synergetics. An Introduction* (Springer), an excellent introduction to key ideas from a physics perspective.

Index