Vorlesung 7.5

Peter Nejjar

Hier wird der Inhalt der Vorlesung vom 7.5 wiedergegeben. Insofern sich die Vorlesung an [1] orientierte, werden die Inhalte anhand der dortigen Bezeichnungen/Nummern nur kurz genannt.

Satz 2.2.1. - allgemeine Version

Wir hatten im Satz 2.2.1 gesehen, dass eine stetige Funktion $f : [a,b] \to [0,+\infty)$ ein Wahrscheinlichkeitsmaß $\mathbb{P}$ auf den Borelmengen von $[a,b]$ produziert, falls $\int_a^b \mathrm{d}x f(x) = 1$ gilt: Das Maß $\mathbb{P}$ ist dadurch charakterisiert, dass für $[c,d] \subseteq [a,b]$ gilt:

$$\mathbb{P}([c,d]) = \int_c^d \mathrm{d}x f(x). \tag{1}$$

Es ist aber nicht nötig, dass f stetig ist. Dies ist der Inhalt des folgenden Satzes:

Satz (Satz 2.2.1 - allgemeine Version)**.** *Sei $f : \mathbb{R} \to [0,+\infty)$ vorgegeben. Wir setzen voraus, dass das Integral $\int_{\mathbb{R}} \mathrm{d}x f(x)$ existiert, und $\int_{\mathbb{R}} \mathrm{d}x f(x) = 1$ gilt. Dann gibt es ein eindeutig bestimmtes Wahrscheinlichkeitsmaß $\mathbb{P}$ auf den Borelmengen von $\mathbb{R}$, das durch die Festlegung*

$$\mathbb{P}([c,d]) = \int_c^d \mathrm{d}x f(x) \tag{2}$$

für beliebige $c, d \in \mathbb{R}$ mit $c < d$ charakterisiert ist.

Die folgende, erste Anwendung dieses allgemeineren Satzes ist technischer Natur.
Dazu führen wir folgende Notation ein:

Definition (Indikatorfunktion)**.** *Sei E eine beliebige Menge und $A \subseteq E$ eine Teilmenge. Die sogenannte Indikatorfunktion von A, geschrieben $\mathbf{1}_A$, ist definiert durch*

$$\mathbf{1}_A : E \to \{0,1\} \tag{3}$$

$$\mathbf{1}_A(x) = \begin{cases} 1, \text{ falls } x \in A \\ 0 \ \text{ sonst.} \end{cases} \tag{4}$$

Ist z.B. $E = \mathbb{R}$ und $A = [5,6]$, dann ist $\mathbf{1}_A(x)$ gleich 1, wenn x zwischen 5 und 6 liegt, und 0 für jede andere reelle Zahl x. Ist $A = [0,+\infty)$, ist auch die Notation $\mathbf{1}_{\{x \geq 0\}} := \mathbf{1}_{[0,+\infty)}$ üblich.

Sei nun A eine Teilmenge von $\mathbb{R}$, es ist entweder $A = [a,b]$ oder $A = [0,+\infty)$. Mit der Indikatorfunktion können wir ein Wahrscheinlichkeitsmaß $\mathbb{P}$, das mithilfe einer Dichtefunktion $f : A \to [0,+\infty)$ auf den Borelmengen von A definiert sind, zu einem Wahrscheinlichkeitsmaß $\tilde{\mathbb{P}}$ auf den Borelmengen $\mathbb{R}$ erweitern: Dazu betrachten wir einfach die Dichtefunktion

$$g(x) := \mathbf{1}_A(x) f(x). \tag{5}$$

Es ist also $g(x) = f(x)$ für alle $x \in A$, und $g(x) = 0$, falls $x \notin A$. Diese Funktion g erfüllt dann die Voraussetzungen des obigen Satzes, daher wird so ein Wahrscheinlichkeitsmaß $\tilde{\mathbb{P}}$ auf den Borelmengen von $\mathbb{R}$ definiert. Das neue Wahrscheinlichkeitsmaß $\tilde{\mathbb{P}}$ stimmt mit $\mathbb{P}$ auf allen Teilmengen von A überein: Stets ist $\mathbb{P}(B) = \tilde{\mathbb{P}}(B)$, für $B \subseteq A$. Allgemeiner ist für beliebiges $B \subseteq \mathbb{R}$

$$\tilde{\mathbb{P}}(B) = \mathbb{P}(B \cap A), \tag{6}$$

insbesondere, falls $B \subseteq \mathbb{R} \setminus A$ gilt, d.h. $B \cap A = \emptyset$, ist $\tilde{\mathbb{P}}(B) = 0$.

Wir veranschaulichen das am Beispiel der

- Exponentialverteilung: Diese ist durch die Dichtefunktion $f : [0, +\infty) \to [0, +\infty)$ mit $f(x) = \lambda e^{-\lambda x}$ charakterisiert. Durch f wird ein Wahrscheinlichkeitsmaß $\mathbb{P}$ auf den Borelmengen von $[0, +\infty)$ definiert. Wir definieren nun die Funktion g aus (5) mittels $A = [0, +\infty)$, d.h.

$$g : \mathbb{R} \to [0, +\infty), \quad g(x) = \mathbf{1}_{[0+\infty)}(x)\lambda e^{-\lambda x}.$$

  Dadurch wird ein Wahrscheinlichkeitsmaß $\tilde{\mathbb{P}}$ auf *allen* Borelmengen von $\mathbb{R}$ definiert. Es ist $\tilde{\mathbb{P}}(B) = 0$ sobald $B \subseteq (-\infty, 0)$ ist, und $\tilde{\mathbb{P}}(B) = \mathbb{P}(B)$ für alle $B \subseteq [0, +\infty)$.

Als weitere, interessantere Anwendung von Satz 2.2.1 (allg. Version) können wir etwa die unstetige Funktion

$$f : \mathbb{R} \to [0, +\infty), \quad f(x) = \mathbf{1}_{[0,1/2]}(x) + \mathbf{1}_{[1,3/2]}(x) \tag{7}$$

betrachten. Dann erfüllt f erfüllt die Voraussetzungen von Satz 2.2.1 (allg. Version) und definiert also ein Wahrscheinlichkeitsmaß auf $\mathbb{R}$.

Die Funktion f ist die *Dichtefunktion* des Wahrscheinlichkeitsmaßes $\mathbb{P}$. Dies halten wir in folgender Definition fest:

Definition. *Sei f wie im obigen Satz 2.2.1 (allgemeine Version). Dann heißt f die* ***Dichtefunktion*** *des dazugehörigen Wahrscheinlichkeitsmaßes $\mathbb{P}$.*

Also ist etwa $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ die Dichtefunktion der Normalverteilung. Gibt es eine Teilmenge $A \subseteq \mathbb{R}$, sodass die Einschränkung von f auf A, definiert durch

$$f_{|A}(x) : A \to [0, +\infty), f_{|A}(x) = f(x),$$

bereits $\int_A \mathrm{d}x f_{|A}(x) = 1$ erfüllt, wird auch häufig $f_{|A}$ als Dichtefunktion von $\mathbb{P}$ bezeichnet, obwohl $f_{|A}$ nur ein Wahrscheinlichkeitsmaß auf den Borelmengen in A definiert.

Zur Dichtefunktion gehört noch die Verteilungsfunktion F:

Definition. *Sei f wie im obigen Satz 2.2.1 (allgemeine Version). Dann heißt*

$$F(x) := \int_{-\infty}^{x} \mathrm{d}y f(y) \tag{8}$$

die ***Verteilungsfunktion*** *des dazugehörigen Wahrscheinlichkeitsmaßes $\mathbb{P}$.*

Es ist also gerade $F(x) = \mathbb{P}((-\infty, x])$. Wir beweisen folgendes Lemma.

Lemma 1. *Sei F eine Verteilungsfunktion . Es gilt:*

(i) Für alle $x \in \mathbb{R}$ ist $F(x) \in [0, 1]$.

(ii) Ist $x \leq y$, so ist auch $F(x) \leq F(y)$.

(iii) Es ist $\lim_{x\to+\infty} F(x) = 1, \lim_{x\to-\infty} F(x) = 0$.

Proof. (i) folgt direkt aus $F(x) = \mathbb{P}((-\infty, x])$. Da $(-\infty, x] \subseteq (-\infty, y]$ für $x \leq y$, folgt auch (ii), indem wir Satz 1.3.2 (iii) anwenden. Für (iii) sei x_n eine Folge mit $x_n \to +\infty$, z. B. $x_n = n$. Sei dann $E_n = (-\infty, x_n]$. Es ist $\cup_n E_n = \mathbb{R}$, und wegen der Stetigkeit von oben (Satz 1.3.2 (v)) folgt

$$\lim_{x\to+\infty} F(x) = \lim_{n\to+\infty} F(x_n) = \lim_{n\to+\infty} \mathbb{P}(E_n) = \mathbb{P}(\mathbb{R}) = 1. \tag{9}$$

Analog zeigt man $\lim_{x\to-\infty} F(x) = 0$, indem man die Stetigkeit nach unten (Satz 1.3.2 (vi)) ausnutzt. □

Kapitel 3 - Zufallsvariablen

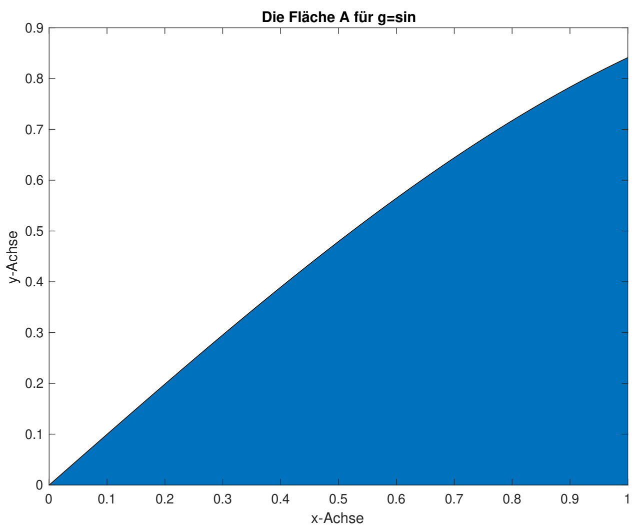
Zwei einfache Beispiele von Zufallsvariablen (vgl. S 72 in [1]):

- Augensumme beim zweifachen Würfeln
- Anzahl der Richtigen beim Lotto

Ein komplizierteres Beispiel ist das Folgende: Es geht darum, den Zufall zu nutzen, um Integrale auszurechnen. Wir betrachten die Gleichverteilung $\mathbb{P}$ auf dem Quadrat $\Omega = [0,1]^2$, die Wahrscheinlichkeit einer Borelmenge E ist also gerade der Flächeninhalt von E. Sei nun $g : [0,1] \to [0,1]$ stetig. Was ist dann anschaulich die Menge

$$A := \{(c,d) \in [0,1]^2 : d \leq g(c)\}\,?$$

Das folgende Bild zeigt A für $g(x) = \sin(x)$ in blau:



A ist die Menge aller Punkte im Quadrat, die 'unter' der Kurve von g liegen. Die Fläche von A ist dann gerade das Integral $\int_0^1 \mathrm{d}x g(x)$, und per Konstruktion ist damit $\mathbb{P}(A) = \int_0^1 \mathrm{d}x g(x)$. Wir definieren eine Zufallsvariable X wie folgt:

$$X : [0,1]^2 \to \{0,1\}$$
$$X((c,d)) = \mathbf{1}_A((c,d)).$$

Wir können uns X als das Ergebnis eines Münzwurfs vorstellen, bei dem 1 rauskommt, wenn ein Punkt (c,d) zur Fläche A gehört, und 0 sonst. Das Wahrscheinlichkeitsmaß $\mathbb{P}_X$ auf $\{0,1\}$, das durch

$$\mathbb{P}_X(\{1\}) := \mathbb{P}(\{(c,d) : X(c,d) = 1\}) = \mathbb{P}(A) = \int_0^1 \mathrm{d}x g(x)$$

eindeutig definiert ist, ist also eine Bernoulliverteilung mit $p = \int_0^1 \mathrm{d}x g(x)$. Jetzt nehmen wir an, unser Computer wäre in der Lage, wiederholt zufällig gleichverteilt Punkte in Ω zu erzeugen - z.B. in Matlab geschieht das mit dem Befehl $rand(2,1)$. Das Erzeugen der Punkte soll unabhängig voneinander geschehen - was das mathematisch genau heißt, wird bald in der Vorlesung geklärt werden. Dann sollte doch, wenn $n-$mal Punkte im Quadrat erzeugt werden, in etwa $p*n$ häufig der Punkt in A landen und $(1-p)*n$ häufig nicht in A. Wir gehen nun wie folgt vor: Wir ezeugen n solche Punkte im Quadrat. Wir zählen, wieviele dieser n Punkte in A gelandet sind. Sei M diese Zahl (es sind also $n - M$ Punkte nicht in A gelandet). Dann sollte *ungefähr* gelten

$$\frac{M}{n} \approx p \tag{10}$$

mit $p = \int_0^1 \mathrm{d}x g(x)$, und wir berechnen p numerisch als $\frac{M}{n}$. Was 'ungefähr' hier genau heißt, wird in den kommenden Vorlesungen geklärt werden.

3.1 Was ist eine Zufallsvariable?

Behandelte Themen: Definition 3.1.1, Satz 3.1.2, Korollar 3.1.3, Satz 3.1.4.

Literatur

[1] E. Behrends. Elementare Stochastik. *Vieweg+Teubner Verlag, 2013,* https://link.springer.com/book/10.1007/978-3-8348-2331-1.