

# **Stochastik SS 2019**

Dozent: Prof. Dr. ANITA BEHME

4. April 2019

# *Inhaltsverzeichnis*

<b>I</b>	<b>Grundbegriffe der Wahrscheinlichkeitstheorie</b>	<b>4</b>
1	Wahrscheinlichkeitsräume . . . . .	4
2	Zufallsvariablen . . . . .	8
<b>II</b>	<b>Test</b>	<b>11</b>

# *Vorwort*

# *Literatur*

- *Georgii*: Stochastik (5. Auflage)
- *Schilling*: Wahrscheinlichkeit (1. Auflage)
- *Bauer*: Wahrscheinlichkeitstheorie (5. Auflage) (sehr maßtheoretisch!)

Ohne Maßtheorie!

- *Krengel*: Einführung in die Wahrscheinlichkeitstheorie und Statistik
- *Dehling & Haupt*: Einführung in die Wahrscheinlichkeitstheorie und Statistik

# *Was ist Stochastik?*

Altgriechisch Stochastikos ( $\sigma\tau\omicron\chi\alpha\sigma\tau\iota\kappa\omicron\varsigma$ ) und bedeutet sinngemäß “scharfsinning in Vermuten”.

Fragestellung insbesondere aus Glückspiel, Versicherungs-/Finanzmathematik, überall da wo Zufall/Risiko / Chance auftauchen.

Was ist Stochastik?

- Beschreibt zufällige Phänomene in einer exakten Sprache!  
Beispiel: “Beim Würfeln erscheint jedes sechste Mal (im Schnitt) eine 6.”  $\rightarrow$  Gesetz der großen Zahlen ( $\nearrow$  später)
- Lässt sich mathematische Stochastik in zwei Teilgebiete unterteilen  
Wahrscheinlichkeitstheorie (W-Theorie) & Statistik
  - *W-Theorie*: Beschreibt und untersucht konkret gegebene Zufallssituationen.
  - *Statistik*: Zieht Schlussfolgerungen aus Beobachtungen.

Statistik benötigt Modelle der W-Theorie und W-Theorie benötigt die Bestätigung der Modelle durch Statistik.

In diesem Semester konzentrieren wir uns nur auf die Wahrscheinlichkeitstheorie!

## Kapitel I

# *Grundbegriffe der Wahrscheinlichkeitstheorie*

## 1. Wahrscheinlichkeitsräume

### Ergebnisraum

Welche der möglichen Ausgänge eines zufälligen Geschehens interessieren uns?

Würfeln? Augenzahl, nicht die Lage und die Fallhöhe

#### **Definition 1.1 (Ergebnisraum)**

Die Menge der relevanten Ergebnisse eines Zufallsgeschehens nennen wir Ergebnisraum und bezeichnen diesen mit  $\Omega$ .

#### ■ Beispiel

- Würfeln:  $\Omega = \{1, 2, \dots, 6\}$
- Wartezeiten:  $\Omega = \mathbb{R}_+ = [0, \infty)$  (überabzählbar!)

### Ereignisse

Oft interessieren wir uns gar nicht für das konkrete Ergebnis des Zufallsexperiments, sondern nur für das Eintreten gewisser Ereignisse.

#### ■ Beispiel

- Würfeln: Zahl ist  $\geq 3$
- Wartezeit: Wartezeit  $\leq 5$  Minuten

→ Teilmenge des Ereignisraums, also Element der Potenzmenge  $\mathcal{P}(\Omega)$ , denen eine Wahrscheinlichkeit zugeordnet werden kann, d.h. welche messbar (mb) sind.

#### **Definition 1.2 (Ereignisraum, messbarer Raum)**

Sei  $\Omega \neq \emptyset$  ein Ergebnisraum und  $\mathcal{F}$  eine  $\sigma$ -Algebra auf  $\Omega$ , d.h. eine Familie von Teilmenge von  $\Omega$ , sodass

1.  $\Omega \in \mathcal{F}$
2.  $A \in \mathcal{F} \Rightarrow A^C \in \mathcal{F}$
3.  $A_1, A_2, \dots \in \mathcal{F} \Rightarrow \bigcup_{i \geq 1} A_i \in \mathcal{F}$

Dann heißt  $(\Omega, \mathcal{F})$  Ereignisraum bzw. messbarer Raum.

### Wahrscheinlichkeiten

Ordne Ereignissen Wahrscheinlichkeiten zu mittels der Abbildung

$$\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$$

sodass

$$\text{Normierung } \mathbb{P}(\Omega) = 1 \quad (\text{N})$$

$$\sigma\text{-Additivität für paarweise disjunkte Ereignisse} \quad (\text{A})$$

$$A_1, A_2, \dots \in \mathcal{F} \Rightarrow \mathbb{P}\left(\bigcup_{i \geq 1} A_i\right) = \sum_{i \geq 1} \mathbb{P}(A_i)$$

(N), (A) und die Nichtnegativität von  $\mathbb{P}$  werden als KOLMOGOROVsche Axiome bezeichnet (nach Kolmogorov: Grundbegriffe der Wahrscheinlichkeitstheorie, 1933)

**Definition 1.3 (Wahrscheinlichkeitsmaß, Wahrscheinlichkeitsverteilung)**

Sei  $(\Omega, \mathcal{F})$  ein Ereignisraum und  $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$  eine Abbildung mit Eigenschaften (N) und (A).

Dann heißt  $\mathbb{P}$  Wahrscheinlichkeitsmaß oder auch Wahrscheinlichkeitsverteilung.

Aus der Definition folgen direkt:

**Satz 1.4 (Rechenregeln für W-Maße)**

Sei  $\mathbb{P}$  ein W-Maß, Ereignisse  $(\Omega, \mathcal{F})$ ,  $A, B, A_1, A_2, \dots \in \mathcal{F}$ . Dann gelten:

1.  $\mathbb{P}(\emptyset) = 0$
2. Monotonie:  $A \subseteq B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$
3. endliche  $\sigma$ -Additivität:  $\mathbb{P}(A \cup B) + \mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B)$  und insbesondere  $\mathbb{P}(A) + \mathbb{P}(A^C) = 1$
4.  $\sigma$ -Subadditivität:

$$\mathbb{P}\left(\bigcup_{i \geq 1} A_i\right) \leq \sum_{i \geq 1} \mathbb{P}(A_i)$$

5.  $\sigma$ -Stetigkeit: Wenn  $A_n \uparrow A$  (d.h.  $A_1 \subseteq A_2 \subseteq \dots$  und  $A = \bigcup_{i=1}^{\infty} (A_i)$ ) oder  $A_n \downarrow A$ , so gilt:

$$\mathbb{P}(A_n) \longrightarrow \mathbb{P}(A), n \rightarrow \infty$$

*Beweis.* In der Vorlesung wurde nur auf Schillings MINT Vorlesung verwiesen. Der folgende Beweis wurde ergänzt.

Beweise erst folgende Aussage:  $A \cap B = \emptyset \implies \mathbb{P}(A \uplus B) = \mathbb{P}(A) + \mathbb{P}(B)$ .

Es kann  $\sigma$ -Additivität verwendet werden, indem "fehlende" Mengen durch  $\emptyset$  ergänzt werden:

$$\mathbb{P}(A \uplus B) = \mathbb{P}(A \uplus B \uplus \emptyset \uplus \emptyset \dots) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(\emptyset) + \dots = \mathbb{P}(A) + \mathbb{P}(B),$$

wobei Maßeigenschaften verwendet wurden.

1. Definition des Maßes.
2. Da  $A \subseteq B$  ist auch  $B = A \uplus (B \setminus A) = A \uplus (B \setminus (A \cap B))$ . Wende wieder Aussage von oben an, damit folgt

$$\mathbb{P}(B) = \mathbb{P}(A \uplus (B \setminus A)) = \mathbb{P}(A) + \mathbb{P}(B \setminus A) \geq \mathbb{P}(A) \quad (*)$$

3. Zerlege  $A \cup B$  geschickt, dann sieht man mit oben gezeigter Aussage und  $(*)$

$$\begin{aligned} \mathbb{P}(A \cup B) + \mathbb{P}(A \cap B) &= \mathbb{P}(A \uplus (B \setminus (A \cap B))) + \mathbb{P}(A \cap B) \\ &= \mathbb{P}(A) + \mathbb{P}(B \setminus (A \cap B)) + \mathbb{P}(A \cap B) \\ &= \mathbb{P}(A) + \mathbb{P}(B). \end{aligned}$$

Im letzten Schritt wurde  $(*)$  verwendet.

4. Folgt aus endlicher  $\sigma$ -Additivität, da  $\mathbb{P}\left(\bigcap_{i \geq 1} A_i\right) \geq 0$ .
5. Definiere  $F_1 := A_1, F_2 := A_2 \setminus A_1, \dots, F_{i+1} := A_{i+1} \setminus A_i$ . Die  $F_i$  Mengen sind paarweise disjunkt und damit folgt für  $m \rightarrow \infty$

$$A_m = \biguplus_{i=1}^m F_i \Rightarrow A = \biguplus_{i=1}^{\infty} F_i = \biguplus_{i=1}^{\infty} A_i$$

und

$$\mathbb{P}(A) = \mathbb{P}\left(\biguplus_{i=1}^{\infty} F_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(F_i) = \lim_{m \rightarrow \infty} \mathbb{P}\left(\biguplus_{i=1}^m F_i\right) = \lim_{m \rightarrow \infty} \mathbb{P}(A_m).$$

Siehe Schillings Buch. □

### ■ Beispiel 1.5

Für ein beliebigen Ereignisraum  $(\Omega, \mathcal{F})$  ( $\Omega \neq \emptyset$ ) und ein beliebiges Element  $\xi \in \Omega$  definiere

$$\delta_{\xi}(A) := \begin{cases} 1 & \xi \in A \\ 0 & \text{sonst} \end{cases}$$

eine (degeneriertes) W-Maß auf  $(\Omega, \mathcal{F})$ , welches wir als DIRAC-Maß oder DIRAC-Verteilung bezeichnen.

### ■ Beispiel 1.6

Würfeln mit fairem, 6-(gleich)seitigem Würfel mit Ergebnismenge  $\Omega = \{1, \dots, 6\}$  und Ereignisraum  $\mathcal{F} = \mathcal{P}(\Omega)$  setzen wir als Symmetriegründen

$$\mathbb{P}(A) = \frac{\#A}{6}.$$

(Wobei  $\#A$  oder auch  $|A|$  die Kardinalität von  $A$  ist.) Das definiert ein W-Maß.

### ■ Beispiel 1.7

Wartezeit an der Bushaltestelle mit Ergebnisraum  $\Omega = \mathbb{R}_+$  und Ereignisraum BORELSche  $\sigma$ -Algebra



$\mathcal{B}(\mathbb{R}_+) = \mathcal{F}$ . Eine mögliches W-Maß können wir dann durch

$$\mathbb{P}(A) = \int_A \lambda e^{-\lambda x} dx$$

für einen Parameter  $\lambda > 0$  festlegen. (Offenbar gilt  $\mathbb{P}(\Omega) = 1$  und die  $\sigma$ -Additivität aufgrund der Additivität des Integrals.) Wir bezeichnen diese Maß als Exponentialverteilung. (Warum gerade dieses Maß für Wartezeiten gut geeignet ist  $\nearrow$  später)

### Satz 1.8 (Konstruktion von WMaßen durch Dichten)

Sei  $(\Omega, \mathcal{F})$  ein Ereignisraum.

- $\Omega$  abzählbar,  $\mathcal{F} = \mathcal{P}(\Omega)$ : Sei  $\rho = (\rho(\omega))_{\omega \in \Omega}$  eine Folge in  $[0, 1]$  in  $\sum_{\omega \in \Omega} \rho(\omega) = 1$ , dann definiert

$$\mathbb{P}(A) = \sum_{\omega \in \Omega} \rho(\omega), A \in \mathcal{F}$$

ein (diskretes) WMaß  $\mathbb{P}$  auf  $(\Omega, \mathcal{F})$ .  $\rho$  wird als Zähldichte bezeichnet.

- Umgekehrt definiert jedes WMaß  $\mathbb{P}$  auf  $(\Omega, \mathcal{F})$  definiert Folge  $\rho(\omega) = \mathbb{P}(\{\omega\}), \omega \in \Omega$  eine Folge  $\rho$  mit den obigen Eigenschaften.
- $\Omega \in \mathbb{R}^n, \mathcal{F} = \mathcal{B}(\Omega)$ : Sei  $\rho : \Omega \rightarrow [0, \infty)$  eine Funktion, sodass

$$1. \int_{\Omega} \rho(x) dx = 1$$

$$2. \{x \in \Omega : f(x) \leq c\} \in \mathcal{B}(\Omega) \text{ für alle } c > 0$$

dann definiert  $\rho$  ein WMaß  $\mathbb{P}$  auf  $(\Omega, \mathcal{F})$  durch

$$\mathbb{P}(A) = \int_A \rho(x) dx = \int_A \rho d\lambda, \quad A \in \mathcal{B}(\Omega). \quad (1)$$

Das Integral interpretieren wir stets als Lebesgue-Integral bzw. Lebesgue-Maß  $\lambda$ .  $\rho$  bezeichnet wir als Dichte, Dichtefunktion/Wahrscheinlichkeitsdichte von  $\mathbb{P}$  und nennen ein solches  $\mathbb{P}$  (absolut)stetig (bzgl. denn Lebesgue-Maß).

*Beweis.* • Der diskrete Fall ist klar.

- Im stetigen Fall folgt die Bahauptung aus den bekannten Eigenschaften des Lebesgue-Integrals ( $\nearrow$  Schilling MINT, Lemma 8.9)  $\square$

### ► Bemerkung

- Die Eineindeutige Beziehung zwischen Dichte und WMaß überträgt sich nicht auf den stetigen Fall.
  - Nicht jedes WMaß auf  $(\Omega, \mathcal{B}(\Omega)), \Omega \subset \mathbb{R}^n$  besitzt eine Dichte.
  - Zwei Dichtefunktionen definieren dasselbe WMaß, wenn sie sich nur auf einer Menge von Lebesgue-Maß 0 unterscheiden.
- Jede auf  $\Omega \subset \mathbb{R}^n$  definiert Dichtefunktion  $\rho$  lässt sich auf ganz  $\mathbb{R}^n$  fortsetzen durch  $\rho(x) =$

$0, x \notin \Omega$ . Das erzeugte WMaß auf  $(\mathbb{R}^n, \mathcal{B}(\Omega))$  lässt mit den WMaß auf  $(\Omega, \mathcal{F})$  identifizieren.

- Mittels Dirac-Maß  $\delta_x$  können auch jedes diskrete WMaß auf  $\Omega \subset \mathbb{R}^n$  als WMaß auf  $\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)$  interpretieren

$$\mathbb{P}(A) = \sum_{\omega \in A} \rho(\omega) = \int_A d\left(\sum_{\omega \in \Omega} \rho(\omega) \delta_\omega\right)$$

stetige und diskrete WMaße lassen sich kombinieren z.B.

$$\mathbb{P}(A) = \frac{1}{2} \delta_0 + \frac{1}{2} \int_A \mathbb{1}_{[0,1]}(x) dx, A \in \mathcal{B}(\mathbb{R})$$

ein WMaß auf  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ .

Abschließend erinnern wir uns an:

### Satz 1.9 (Eindeutigkeitssatz für WMaße)

Sei  $(\Omega, \mathcal{F})$  Ereignisraum und  $\mathbb{P}$  ein WMaß auf  $(\Omega, \mathcal{F})$ . Sei  $\mathcal{F} = \sigma(\mathcal{G})$  für ein  $\cap$ -stabiles Erzeugendensystem  $\mathcal{G} \subset \mathcal{P}(\Omega)$ . Dann ist  $\mathbb{P}$  bereits durch seine Einschränkung  $\mathbb{P}|_{\mathcal{G}}$  eindeutig bestimmt.

*Beweis.* ↗ Schilling MINT, Satz 4.5. □

Insbesondere definiert z.B.

$$\mathbb{P}([0, a]) = \int_0^a \lambda e^{-\lambda x} dx = 1 - e^{-\lambda a}, a > 0$$

bereits die Exponentialverteilung aus Beispiel 1.7.

### Definition 1.10 (Gleichverteilung)

Ist  $\Omega$  endlich, so heißt das WMaß mit konstanter Zähldichte  $\rho(\omega) = \frac{1}{|\Omega|}$  die (diskrete) Gleichverteilung auf  $\Omega$  und wird mit  $U(\Omega)$  notiert ( $U$  = Uniform). Ist  $\Omega \subset \mathbb{R}^n$  eine Borelmenge mit Lebesgue-Maß  $0 < \lambda^n(\Omega) < \infty$  so heißt das WMaß auf  $(\Omega, \mathcal{B}(\Omega))$  mit konstanter Dichtefunktion  $\rho(x) = 1/\lambda^n(\Omega)$  die (stetige) Gleichverteilung auf  $\Omega$ . Sie wird ebenso mit  $U(\Omega)$  notiert.

## WRäume

### Definition 1.11 (Wahrscheinlichkeitsraum)

Ein Tripel  $(\Omega, \mathcal{F}, \mathbb{P})$  mit  $\Omega, \mathcal{F}$  Ereignisraum und  $\mathbb{P}$  WMaß auf  $(\Omega, \mathcal{F})$ , nennen wir Wahrscheinlichkeitsraum.

## 2. Zufallsvariablen

Zufallsvariablen dienen dazu von einem gegebenen Ereignisraum  $(\Omega, \mathcal{F})$  zu einem Modellausschnitt  $(\Omega', \mathcal{F}')$  überzugehen. Es handelt sich also um Abbildungen  $X : \Omega \rightarrow \Omega'$ . Damit wir auch jedem Ereignis in  $\mathcal{F}'$  eine Wheat zuordnen können, benötigen wir

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

d.h.  $X$  sollte messbar sein.

**Definition 2.1 (Zufallsvariable)**

Seien  $(\Omega, \mathcal{F})$  und  $(\Omega', \mathcal{F}')$  Ereignisräume. Dann heißt jede messbare Abbildung

$$X : \Omega \rightarrow \Omega'$$

Zufallsvariable (von  $(\Omega, \mathcal{F})$ ) nach  $(\Omega', \mathcal{F}')$  / auf  $(\Omega', \mathcal{F}')$  oder Zufallselement.

■ **Beispiel 2.2**

1. Ist  $\Omega$  abzählbar und  $\mathcal{F} = \mathcal{P}(\Omega)$ , so ist jede Abbildung  $X : \Omega \rightarrow \Omega'$  messbar und damit eine Zufallsvariable.
2. Ist  $\Omega \subset \mathbb{R}^n$  und  $\mathcal{F} = \mathcal{B}(\Omega)$ , so ist jede stetige Funktion  $X : \Omega \rightarrow \mathbb{R}$  messbar und damit eine Zufallsvariable.

**Satz 2.3**

Sei  $(\Omega, \mathcal{F}, \mathbb{P})$  ein WRaum und  $X$  eine Zufallsvariable von  $(\Omega, \mathcal{F})$  nach  $(\Omega', \mathcal{F}')$ . Dann definiert

$$\mathbb{P}'(A') := \mathbb{P}(X^{-1}(A')) = \mathbb{P}(\{X \in A'\}), A' \in \mathcal{F}'$$

ein WMaß auf  $(\Omega', \mathcal{F}')$  auf  $(\Omega', \mathcal{F}')$ , welches wir als WVerteilung von X unter  $\mathbb{P}$  bezeichnen.

*Beweis.* Aufgrund der Messbarkeit von  $X$  ist die Definition sinnvoll. Zudem gelten

$$\mathbb{P}'(\Omega') = \mathbb{P}(X^{-1}(\Omega')) = \mathbb{P}(\Omega) = 1$$

und für  $A'_1, A'_2, \dots \in \mathcal{F}'$  paarweise disjunkt.

$$\begin{aligned} \mathbb{P}'\left(\bigcup_{i \geq 1} A'_i\right) &= \mathbb{P}\left(X^{-1}\left(\bigcup_{i \geq 1} A'_i\right)\right) \\ &= \mathbb{P}\left(\bigcup_{i \geq 1} X^{-1}(A'_i)\right) \\ &= \sum_{i \geq 1} \mathbb{P}(X^{-1} A'_i) \end{aligned}$$

da auch  $X^{-1} A'_1, X^{-1} A'_2, \dots$  paarweise disjunkt

$$= \sum_{i \geq 1} \mathbb{P}'(A'_i)$$

Also ist  $\mathbb{P}'$  ein WMaß. □

► **Bemerkung**

- Aus Gründen der Lesbarkeit schreiben wir in der Folge  $\mathbb{P}(X \in A) = \mathbb{P}(\{\omega : X(\omega) \in A\})$
- Ist  $X$  die Identität, so fallen die Begriffe WMaß und WVerteilung zusammen.
- In der (weiterführenden) Literatur zu WTheorie wird oft auf die Angabe eines zugrundeliegenden WRaumes verzichtet und stattdessen eine “Zufallsvariable mit Verteilung  $\mathbb{P}$  auf  $\Omega$ ”

eingeführt. Gemeint ist (fast) immer  $X$  als Identität auf  $(\Omega, \mathcal{F}, \mathbb{P})$  mit  $\mathcal{F} = \mathcal{P}(\omega)/\mathcal{B}(\Omega)$ .

- Für die Verteilung von  $X$  unter  $\mathbb{P}$  schreibe  $\mathbb{P}_X$  und  $X \sim \mathbb{P}_X$  für die Tatsache, dass  $X$  gemäß  $\mathbb{P}_X$  verteilt ist.

**Definition 2.4 (identisch verteilt, reellen Zufallsvariablen)**

Zwei Zufallsvariablen sind identisch verteilt, wenn sie dieselbe Verteilung haben. Von besonderem Interesse sind für uns die Zufallsvariablen, die nach  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  abbilden, sogenannten reellen Zufallsvariablen.

Da die halboffenen Intervalle  $\mathcal{B}(\mathbb{R})$  erzeugen, ist die Verteilung einer reellen Zufallsvariable durch die Werte  $(-\infty, c], c \in \mathbb{R}$  eindeutig festgelegt.

**Definition 2.5 ((kommutative) Verteilungsfunktion von  $\mathbb{P}$ )**

Sei  $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P})$  WRaum, so heißt

$$F : \mathbb{R} \rightarrow [0, 1] \text{ mit } x \mapsto \mathbb{P}((-\infty, x])$$

(kumulative) Verteilungsfunktion von  $\mathbb{P}$ .

Ist  $X$  eine reelle Zufallsvariable auf beliebigem WRaum  $(\Omega, \mathcal{F}, \mathbb{P})$ , so heißt

$$F : \mathbb{R} \rightarrow [0, 1] \text{ mit } x \mapsto \mathbb{P}(X \leq x) = \mathbb{P}(X \in (-\infty, x])$$

die (kumulative) Verteilungsfunktion von  $X$ .

## Kapitel II

### *Test*

# Anhang