Stochastik SS 2019

Dozent: Prof. Dr. Anita Behme

7. April 2019

Inhaltsverzeichnis

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 ($\sigma\tau o\chi\alpha\sigma\tau\iota\kappa\grave{o}\zeta$) 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 Spache!
  Beispiel: "Beim Würfeln erscheint jedes sechste Mal (im Schnitt) eine 6." $\longrightarrow$ 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 Ω.

■ **Beispiel**

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

Ereignisse

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

■ **Beispiel**

- Würfeln: Zahl ist ≥ 3
- Wartezeit: Wartezeit ≤ 5 Minuten

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

Definition 1.2 (Ereignisraum, messbarer Raum)
Sei $\Omega \neq \varnothing$ ein Ergebnisraum und $\mathscr{F}$ eine σ-Algebra auf Ω, d.h. eine Familie von Teilmenge von Ω, sodass

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

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

Wahrscheinlichkeiten

Ordne Ereignissen Wahrscheinlichkeiten zu mittels der Abbildung

$$\mathbb{P} : \mathscr{F} \to [0, 1]$$

sodass

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

$$\sigma\text{-Additivität für paarweise disjunkte Ereignisse} \tag{A}$$

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

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

Definition 1.3 (Wahrscheinlichkeitsmaß, Wahrscheinlichkeitsverteilung)
Sei $(\Omega, \mathscr{F})$ ein Ereignisraum und $\mathbb{P} : \mathscr{F} \to [0, 1]$ eine Abbildung mit Eigenschaften Gleichung (N) und Gleichung (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 $(\Omega, \mathscr{F}), A, B, A_1, A_2, \dots \in \mathscr{F}$. Dann gelten:

1. $\mathbb{P}(\varnothing) = 0$
2. Monotonie: $A \subseteq B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$
3. endliche σ-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. σ-Subadditivität:
$$\mathbb{P}\left(\bigcup_{i\geq 1} A_i\right) \leq \sum_{1\geq 1} \mathbb{P}(A_i)$$
5. σ-Stetigkeit: Wenn $A_n \uparrow A$ (d.h. $A_1 \subseteq A_2 \subseteq \cdots$ und $A = \bigcup_{i=1}^{\infty}(A_i)$) oder $A_n \downarrow A$, so gilt:
$$\mathbb{P}(A_n) \longrightarrow \mathbb{P}(A), n \to \infty$$

Beweis. In der Vorlesung wurde nur auf Schillings MINT Vorlesung verwiesen. Der folgende Beweis wurde ergänzt.
Beweise erst folgende Aussage: $A \cap B = \varnothing \Longrightarrow \mathbb{P}(A \uplus B) = \mathbb{P}(A) + \mathbb{P}(B)$.
Es kann σ-Additivität verwendet werden, indem "fehlende" Mengen durch $\varnothing$ ergänzt werden:

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

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

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

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

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

   Im letzten Schritt wurde Gleichung (*) verwendet.
4. Folgt aus endlicher σ-Additivität, da $\mathbb{P}\left(\bigcap_{i \geq 1} A_i\right) \geq 0$.
5. Definiere $F_1 := A_1, F_2 := A_2 \setminus A_1, \ldots, F_{i+1} := A_{i+1} \setminus A_n$. Die F_i Mengen sind paarweise disjunkt und damit folgt für $m \to \infty$

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

   und

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

□

■ Beispiel 1.5

Für ein beliebigen Ereignisraum $(\Omega, \mathscr{F})$ $(\Omega \neq \varnothing)$ und eine 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 $(\Omega, \mathscr{F})$, welches wir als Dirac-Maß oder Dirac-Verteilung bezeichnen.

■ Beispiel 1.6

Würfeln mit fairem, 6-(gleich)seitigem Würfel mit Ergebnismenge $\Omega = \{1, \ldots, 6\}$ und Ereignisraum $\mathscr{F} = \mathscr{P}(\Omega)$ 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 σ-Algebra $\mathscr{B}(\mathbb{R}_+) = \mathscr{F}$. Eine mögliches W-Maß können wir dann durch

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

für einen Parameter $\lambda > 0$ festlegen. (Offenbar gilt $\mathbb{P}(\Omega) = 1$ und die σ-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 $(\Omega, \mathscr{F})$ ein Eriegnisraum.

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

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

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

- Umgekehrt definiert jedes WMaß $\mathbb{P}$ auf $(\Omega, \mathscr{F})$ definiert Folge $\rho(\omega) = \mathbb{P}(\{\omega\}), \omega \in \Omega$ eine Folge ρ mit den obigen Eigenschaften.

- $\Omega \subset \mathbb{R}^n, \mathscr{F} = \mathscr{B}(\Omega)$: Sei $\rho : \Omega \to [0, \infty)$ eine Funktion, sodass

  1. $\int_\Omega \rho(x) \,\mathrm{d}x = 1$
  2. $\{x \in \Omega \colon f(x) \leq c\} \in \mathscr{B}(\Omega)$ für alle $c > 0$

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

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

  Das Integral interpretieren wir stets als Lebesgue-Integral bzw. Lebesgue-Maß λ. ρ 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 Bahuptung 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 $(\Omega, \mathscr{B}(\Omega)), \Omega \subset \mathbb{R}^n$ 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 ρ lässt sich auf ganz $\mathbb{R}^n$ fortsetzen durch $\rho(x) = 0, x \notin \Omega$. Das erzeugte WMaß auf $(\mathbb{R}^n, \mathscr{B}(\Omega))$ lässt mit den WMaß auf $(\Omega,)$ identifizieren.
- Mittels Dirac-Maß δ_x können auch jedes diskrete WMaß auf $\Omega \subset \mathbb{R}^n$ als WMaß auf $\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n)$

intepretieren

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

stetige und diskrete WMaße lassen sich kombiniere z.B.

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

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

Abschließend erinnern wir uns an:

Satz 1.9 (Eindeutigkeitssatz für WMaße)
Sei $(\Omega, \mathscr{F})$ Ereignisraum und $\mathbb{P}$ ein WMaß auf $(\Omega, \mathscr{F})$. Sei $\mathscr{F} = \omega(\mathscr{G})$ für ein $\cap$-stabiles Erzeugendensystem $\mathscr{G} \subset \mathscr{P}(\Omega)$. Dann ist $\mathbb{P}$ bereits durch seine Einschränkung $\mathbb{P}_{|\mathscr{G}}$ eindeutig bestimmt.

Beweis. ↗ Schhiling MINT, Satz 4.5. □

Insbesondere definiert z.B.

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

bereits die Exponentialverteilung aus Beispiel 1.7.

Definition 1.10 (Gleichverteilung)
Ist Ω endlich, so heißt das WMaß mit konstanter Zähldichte $\rho(\omega) = 1/|\Omega|$ die (diskrete) Gleichverteilung auf Ω 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 $(\Omega, \mathscr{B}(\Omega))$ mit konstanter Dichtefunktion $\rho(x) = 1/\lambda^n(x)$ die (stetige) Gleichverteilung auf Ω. Sie wird ebenso mit $U(\Omega)$ notiert.

WRäume

Definition 1.11 (Wahrscheinlichkeitsraum)
Ein Tripel $(\Omega, \mathscr{F}, \mathbb{P})$ mit $\Omega, \mathscr{F}$ Ereignisraum und $\mathbb{P}$ WMaß auf $(\Omega, \mathscr{F})$, nennen wir Wahrscheinlichkeitsraum.

2. Zufallsvariablen

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

$$A' \in \mathscr{F}' \Rightarrow X'A' \in \mathscr{F}$$

d.h. X sollte messbar sein.

Definition 2.1 (Zufallsvariable)
Seien $(\Omega, \mathscr{F})$ und $(\Omega', \mathscr{F}')$ Ereignisräume. Dann heißt jede messbare Abbildung

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

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

■ Beispiel 2.2

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

Satz 2.3
Sei $(\Omega, \mathscr{F}, \mathbb{P})$ ein WRaum und X eine Zufallsvariable von $(\Omega, \mathscr{F})$ nach $(\Omega', \mathscr{F}')$. Dann definiert

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

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

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 $A'_1, A'_2, \dots \in \mathscr{F}'$ paarweise disjunkt.

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

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

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

Also ist $\mathbb{P}'$ ein WMaß. □

▶ Bemerkung

- Aus Gründen der Lesbarkeit schreiben wir in der Folge $\mathbb{P}(X \in A) = \mathbb{P}(\{\omega\colon X(\omega) \in A\})$
- 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 "Zufalsvariable mit Verteilung $\mathbb{P}$ auf Ω" eingeführt. Gemeint ist (fast) immer X als Identität auf $(\Omega, \mathscr{F}, \mathbb{P})$ mit $\mathscr{F} = \mathcal{P}(\Omega)/\mathscr{B}(\Omega)$.

- 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 besonderen Interesse sind für uns die Zufallsvariablen, die nach $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ abbilden, sogenannte reelle Zufallsvariablen.

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

Definition 2.5 ((kommutative) Verteilungsfunktion von $\mathbb{P}$)
Sei $(\mathbb{R}, \mathscr{B}(\mathbb{R}), \mathbb{P})$ WRaum, so heißt

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

(kommulative) Verteilungsfunktion von $\mathbb{P}$.
Ist X eine reelle Zufallsvariable auf beliebigen WRaum $(\Omega, \mathscr{F}, \mathbb{P})$, so heißt

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

die (kommulative) Verteilungsfunktion von X.

Kapitel II

Test

Anhang