Stochastik SS 2019

Dozent: Prof. Dr. Anita Behme

10. 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 "scharfsinnig in Vermuten".
Fragestellung insbesondere aus Glücksspiel, 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 (Wahrscheinlichkeitstheorie) & Statistik
  - *Wahrscheinlichkeitstheorie*: Beschreibt und untersucht konkret gegebene Zufallssituationen.
  - *Statistik*: Zieht Schlussfolgerungen aus Beobachtungen.

  Statistik benötigt Modelle der Wahrscheinlichkeitstheorie. Wahrscheinlichkeitstheorie 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

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

σ-Additivität für paarweise disjunkte Ereignisse$A_1, A_2, \dots \in \mathscr{F} \Rightarrow \mathbb{P}(\bigcup_{i\geq 1} A_i) = \sum_{i\geq 1} \mathbb{P}(A_i)$ (A)

(N), (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 (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 $(\Omega, \mathscr{F}), A, B, A_1, A_2, \dots \in \mathscr{F}$. 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:
$$\mathbb{P}\left(\bigcup_{i\geq 1} A_i\right) \leq \sum_{i\geq 1} \mathbb{P}(A_i)$$
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 (*)
$$\begin{aligned}\mathbb{P}(A \cup B) + \mathbb{P}(A \cap B) &= \mathbb{P}(A \uplus (B \setminus (A \cap B)) + \mathbb{P}(A \cap B) \\ &= \mathbb{P}(A) + \mathbb{P}(B \setminus (A \cap B)) + \mathbb{P}(A \cap B) \\ &= \mathbb{P}(A) + \mathbb{P}(B).\end{aligned}$$
Im letzten Schritt wurde (*) verwendet.
4. Folgt aus endlicher 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
$$\mathbb{P}(A) = \mathbb{P}\left(\biguplus_{i=1}^{\infty} F_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(F_i) = \lim \lim_{m\to\infty} \mathbb{P}\left(\biguplus_{i=1}^{m} F_i\right) = \lim \lim_{m\to\infty} \mathbb{P}(A_m).$$ □

■ 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 ↗ später)

Satz 1.8 (Konstruktion von Wahrscheinlichkeitsmaßen 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) Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $(\Omega, \mathscr{F})$. ρ wird als Zähldichte bezeichnet.

- Umgekehrt definiert jedes Wahrscheinlichkeitsmaß $\mathbb{P}$ 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 Wahrscheinlichkeitsmaß $\mathbb{P}$ 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 (↗ Schilling MINT, Lemma 8.9) □

▶ **Bemerkung**

- Die Eineindeutige Beziehung zwischen Dichte und Wahrscheinlichkeitsmaß überträgt sich nicht auf den stetigen Fall.
  - Nicht jedes Wahrscheinlichkeitsmaß auf $(\Omega, \mathscr{B}(\Omega)), \Omega \subset \mathbb{R}^n$ besitzt eine Dichte.
  - Zwei Dichtefunktionen definieren dasselbe Wahrscheinlichkeitsmaß, 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 Wahrscheinlichkeitsmaß auf $(\mathbb{R}^n, \mathscr{B}(\Omega))$ lässt mit den Wahrscheinlichkeitsmaß auf $(\Omega,)$ identifizieren.
- Mittels Dirac-Maß δ_x können auch jedes diskrete Wahrscheinlichkeitsmaß auf $\Omega \subset \mathbb{R}^n$ als

Wahrscheinlichkeitsmaß 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 Wahrscheinlichkeitsmaße 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 Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$.

Abschließend erinnern wir uns an:

Satz 1.9 (Eindeutigkeitssatz für Wahrscheinlichkeitsmaße)
Sei $(\Omega, \mathscr{F})$ Ereignisraum und $\mathbb{P}$ ein Wahrscheinlichkeitsmaß 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 Wahrscheinlichkeitsmaß mit konstanter Zähldichte $\rho(\omega) = {}^1/_{|\Omega|}$ die (diskrete) Gleichverteilung auf Ω und wird mit $\mathrm{U}(\Omega)$ notiert (U = Uniform). Ist $\Omega \subset \mathbb{R}^n$ eine Borelmenge mit Lebesgue-Maß $0 < \lambda^n(\Omega) < \infty$ so heißt das Wahrscheinlichkeitsmaß 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 $\mathbb{P}$ Wahrscheinlichkeitsmaß 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 Wahrscheinlichkeitsmaß auf $(\Omega', \mathscr{F}')$ auf $(\Omega', \mathscr{F}')$, welches wir als Wahrscheinlichkeitsverteilung 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.

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

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 Wahrscheinlichkeitsmaß. □

▶ **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 Wahrscheinlichkeitsmaß und Wahrscheinlichkeitsverteilung zusammen.
- In der (weiterführenden) Literatur zu Wahrscheinlichkeitstheorie 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

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

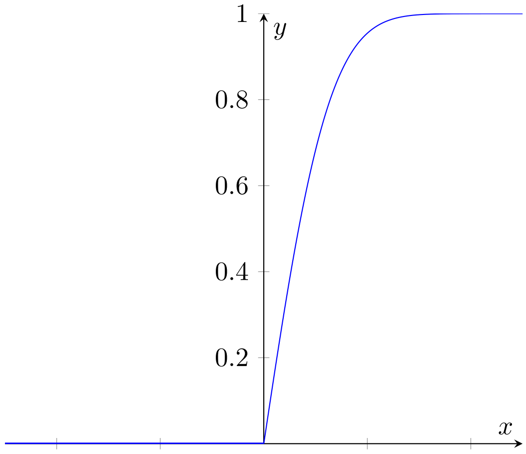
die (kommulative) Verteilungsfunktion von X.

■ **Beispiel 2.6**
Sei $(\mathbb{R}, \mathscr{B}(\mathbb{R}), \mathbb{P})$ mit $\mathbb{P}$ Exponentialverteilung mit Parameter $\lambda > 0$

$$\mathbb{P}(A) = \int_{A \cap [0,\infty)} \lambda e^{-\lambda x} \, \mathrm{d}x \quad A \in \mathscr{B}(\mathbb{R}).$$

Dann ist

$$F(x) = \mathbb{P}((-\infty, x)) = \begin{cases} 0 & x \leq 0 \\ \int_0^x \lambda e^{-\lambda y} \, \mathrm{d}y = 1 - e^{-\lambda x} & x > 0 \end{cases}.$$

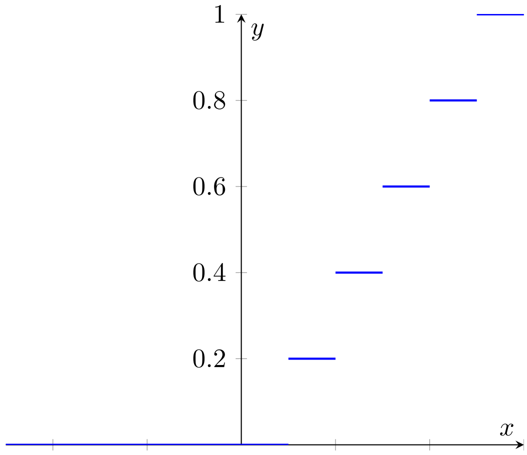


■ **Beispiel 2.7**
Das Würfeln mit einem fairen, sechseitigen Würfel kann mittels einer reellen Zufallsvariablen

$$X : \{1, 2, \ldots, 6\} \to \mathbb{R} \text{ mit } x \mapsto x$$

modelliert werden. Es folgt als Verteilungsfunktion

$$\begin{aligned} F(x) &= \mathbb{P}'(X \leq x) = \mathbb{P}(X^{-1}(-\infty, x]) = \mathbb{P}((-\infty, x]) \\ &= \frac{1}{6} \sum_{i=1}^{6} \mathbb{1}_{i \leq x}. \end{aligned}$$



Allgemein:

Satz 2.8
Ist $\mathbb{P}$ ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ und F die zugehörige Verteilungsfunktion, so gelten

1. F ist monoton wachsend
2. F ist rechtsseitig stetig
3. $\lim\limits_{x\to-\infty} F(x) = 0,\ \lim\limits_{x\to\infty} F(x) = 1$

Umgekehrt existiert zu jeder Funktion $F : \mathbb{R} \to [0,1]$ mit Eigenschaften 1-3 eine reelle Zufallsvariable auf $((0,1), \mathscr{B}((0,1)), \mathrm{U}((0,1))$ mit Verteilungsfunktion F.

Beweis. Ist F Verteilungsfunktion, so folgt mit Satz 1.4

$$x \leq y \Rightarrow F(x) = \mathbb{P}((-\infty, x]) \overset{1.4.3}{\leq} \mathbb{P}((-\infty, y]) = F(y)$$

und

$$\lim_{x \searrow c} F(x) = \lim_{x \searrow c} \mathbb{P}((-\infty, x]) \overset{\sigma\text{-Stetigkeit}}{=} \mathbb{P}((-\infty, c]) = F(c)$$

sowie

$$\lim_{x\to-\infty} F(x) \overset{1.4.5}{=} \mathbb{P}(\varnothing) \overset{1.4.1}{=} 0$$
$$\lim_{x\to\infty} F(x) \overset{1.4.5}{=} \mathbb{P}(\mathbb{R}) = 1.$$

Umgekehrt wähle

$$X(u) := \inf\{x \in \mathbb{R} \colon F(x) \geq u\}, \quad u \in (0,1)$$

Dann ist X eine "linkseitige Inverse" von F (auch Quantilfunktion / verallgemeinerte Inverse). Wegen 3 gilt:

$$-\infty < X(u) < \infty$$

und zudem

$$\{X \leq x\} = (0, F(x)) \cap (0,1) \in \mathscr{B}((0,1)).$$

Da diese halboffene Mengen ein Erzeugendensystem von $\mathscr{B}(\mathbb{R})$ bilden, folgt bereits die Messbarkeit von X, also ist X eine ZV. Insbesondere hat die Menge $\{X \leq x\}$ gerade LEBESGUE-Maß $F(x)$ und damit hat X die Verteilungsfunktion F. □

Folgerung 2.9
Ist $\mathbb{P}$ Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ und F die zugehörige Verteilungsfunktion. Dann besitzt $\mathbb{P}$ genau eine Dichtefunktion ρ, wenn F stetig differenzierbar ist, denn dann gelten

$$F(x) = \int_{-\infty}^{x} \rho(x)\,\mathrm{d}x,\ \text{ bzw. } \rho(x) = F'(x)$$

Beweis. Folgt aus Satz 1.8, der Definition 2.5 der Verteilungsfunktion und dem Eindeutigkeitssatz Satz 1.9. □

Kapitel II

Erste Standardmodelle der Wahrscheinlichkeitstheorie

Diskrete Verteilungen

1. Diskrete Gleichverteilungen

Erinnerung:

▶ **Erinnerung (Definition I.1.10)**
Ist Ω endlich, so heißt Wahrscheinlichkeitsmaß mit Zähldichte

$$\rho(\omega) = \frac{1}{\omega} \quad , \omega \in \Omega$$

(diskrete) Gleichverteilung auf $\Omega \to U(\Omega)$

Es gilt das für jedes $A \in \mathcal{P}(\Omega)$

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|}$$

Anwendungsbeispiele sind faires Würfeln, fairer Münzwurf, Zahlenlotto, ...

2. Urnenmodelle

Ein "Urnenmodell" ist eine abstrakte Darstellung von Zufallsexperimenten, bei denen zufällig Stichproben aus einer gegebenen Menge "gezogen" werden.

Definition (Urne)
Eine Urne ist ein Behältnis in welchem sich farbige/nummerierte Kugeln befinden, die ansonsten ununterscheidbar sind.

Aus der Urne ziehe man blind/zufällig eine oder mehrere Kugeln und notiere Farbe/Zahl.

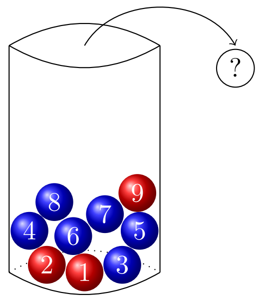


Abbildung II.1: Urnenmodell

2.1. Urnenmodell mit Zurücklegen: Multinomial-Verteilung

Gegeben: Urne mit N Kugeln, verschiedenfarbig mit Farben aus E, $|E| \geq 2$

Ziehe: n Stichproben/Kugeln, wobei nach jedem Zug die Kugel wieder zurückgelegt wird. Uns interessiert die Farbe in jedem Zug, setze also

$$\Omega = E^n \text{ und } \mathscr{F} = \mathcal{P}(\Omega)$$

Zur Bestimmung einer geeigneten Wahrscheinlichkeitsmaßes, nummerieren wir die Kugeln mit $1, \dots, N$, so dass alle Kugeln der Farbe $a \in E$ eine Nummer aus $F_a \subset \{1, \dots, N\}$ tragen. Würden wir die Nummern notieren, so wäre

$$\bar{\Omega} = \{1, \dots, N\}^n \text{ und } \overline{\mathscr{F}} = \mathcal{P}(\overline{\Omega})$$

und wir könnten die Gleichverteilung $\overline{\mathbb{P}} = \mathrm{U}(\overline{\Omega})$ als Wahrscheinlichkeitsmaß für einem einzelnen Zug verwenden. Für den Übergang zu Ω konstruieren wir Zufallsvariablen. Die Farbe im i-ten Zug wird beschrieben durch

$$X_i : \overline{\Omega} \to E \text{ mit } \overline{\omega} = (\overline{\omega}_1, \dots, \overline{\omega}_n) \mapsto a \text{ falls } \overline{\omega}_i \in F_a$$

Der Zufallsvektor

$$X = (X_1, \dots, X_n) : \overline{\Omega} \to \Omega$$

beschreibt dann die Abfolge der Farben. Für jedes $\omega \in \Omega$ gilt dann

$$\{X = \omega\} = F_{\omega_1} \times \cdots \times F_{\omega_n} = \bigtimes_{i=1}^{n} F_{\omega_i}$$

und damit

$$\begin{aligned}\mathbb{P}(\{\omega\}) &= \overline{\mathbb{P}}(X^{-1}(\{\omega\})) = \mathbb{P}(X = \omega) \\ &= \frac{|F_{\omega_1}| \cdots |F_{\omega_n}|}{\left|\overline{\Omega}\right|} \\ &= \prod_{i=1}^{n} \frac{|F_{\omega_i}|}{N} =: \prod_{i=1}^{n} \rho(\omega_i)\end{aligned}$$

Zähldichten, die sich als Produkt von Zähldichten schreiben lassen, werden auch als Produktdichten bezeichnet ($\nearrow$ §3 Unabhängigkeit).

Sehr oft interessiert bei einem Urnenexperiment nicht die Reihenfolge der gezogenen Farben, sondern nur die Anzahl der Kugeln in Farbe $a \in E$ nach n Zügen. Dies enspricht

$$\hat{\Omega} = \left\{ k = (k_a)_{a \in E} \in \mathbb{N}_0^{|E|} \colon \sum_{a \in E} k_a = n \right\} \text{ und } \hat{\mathscr{F}} = \mathcal{P}\left(\hat{\Omega}\right)$$

Den Übergang $\Omega \to \hat{\Omega}$ beschreiben wir durch die Zufallsvariablen

$$Y_a(\omega) : \Omega \to \mathbb{N}_0 \text{ mit } \omega = (\omega_1, \ldots, \omega_n) \mapsto \sum_{a \in E} \mathbb{1}_{\{a\}}(\omega_i)$$

und

$$Y = (Y_a)_{a \in E} : \Omega \to \hat{\Omega} = \left\{ k = (k_a)_{a \in E} \colon \sum_{a \in E} k_a = n \right\}$$

Wir erhalten

$$\begin{aligned}\mathbb{P}(Y = k) &= \mathbb{P}(Y_a = k_a, \;\; a \in E) && (1) \\ &= \sum_{\omega \in \Omega : Y(\omega) = k} \prod_{i=1}^{n} \rho(\omega_i) && (2) \\ &= \sum_{\omega \in \Omega : Y(\omega) = k} \prod_{a \in E} \rho(a) = \left(\begin{matrix} n \\ (k)_{a \in E} \end{matrix} \right) \prod_{a \in E} \rho(a)^{k_a}\end{aligned}$$

wobei

$$\left(\begin{matrix} n \\ (k_1, \ldots, k_l) \end{matrix} \right) = \begin{cases} \frac{n!}{k_1!\, k_2! \cdots k_l!} & \sum_{i=1}^{l} k_i = n \\ 0 & \text{sonst} \end{cases}$$

der Multinomialkoeffizient ist, welcher die Anzahl der Möglichkeiten beschreibt, n Objekte in l Gruppen aufzuteilen, so dass Gruppe i gerade k_i Objekte beinhaltet.

Definition 2.1
Sei $l > 2, p = (p_1, \dots, p_l)$ eine Zähldichte und $n \in \mathbb{N}$, dann heißt die Verteilung auf $\left\{k = (k_i)_{i=1,\dots,l} \in \mathbb{N}_0^l : \sum_{i=1}^l k_i = n\right\}$ mit Zähldichte

$$m((k_1, \dots, k_l)) = \binom{n}{k_1, \dots, k_l} \prod_{i=1}^{l} p_i^{k_i}$$

Multinomialverteilung mit Parametern n und p. Wir schreiben auch Multi(n, p).

■ **Beispiel 2.2**
Eine Urne enthalte nur schwarze "1" und weiße "0" Kugeln, d.h. $E = \{0, 1\}$, und es sei $\rho(1) = p$ gerade die Proportion der schwarzen Kugeln (= Wahrscheinlichkeit bei einem Zug schwarz zu ziehen), dann ist Wahrscheinlichkeit in n Zügen k-mal schwarz zu ziehen:

$$\binom{n}{k} \prod_{i=0,1} \rho(i)^{k_i} = \binom{n}{k} p^k (1-p)^{n-k}.$$

Ein solches (wiederholtes) Experiment mit nur zwei möglichen Ereignissen und fester Wahrscheinlichkeit $p \in [0, 1]$ für eines der Ergebnisse nennen wir auch (wiederholtes) Bernoulliexperiment.

Definition 2.3
Sei $p \in [0, 1]$ un $n \in \mathbb{N}$, dann heißt die Verteilung mit Zähldichte

$$\rho(k) = \binom{n}{k} p^k (1-p)^{n-k} \text{ mit } k \in \{0, 1, \dots, n\}.$$

Binomialverteilung auf $\{0, \dots, n\}$ mit Parameter p (auch Erfolgswahrscheinlichkeit). Wir schreiben auch Bin(n, p). Im Fall $n = 1$ nennen wir die Verteilung mit Zähldichte

$$\rho(0) = 1 - p \text{ und } \rho(1) = p$$

auch Bernoulliverteilung mit Parameter p und schreiben Bernoulli(p).

Urnenmodell ohne Zurücklegen: Hypergeometrische Verteilung
Gegeben: Urne mit N Kugeln verschiedener Farben aus E,

$$|E| \geq 2.$$

Es werden $n \leq N$ Stichproben entnommen, wobei die gezogenen Kugeln werde nicht in die Urne zurückgelegt.

2.2. Urnenmodell ohne Zurücklegen: Hypergeometrische Verteilung

Gegeben: Urne mit N Kugeln verschiedener Farben aus E, $|E| \geq 2$. Es werden $n \leq N$ Stichproben entnommen, wobei die gezogenen Kugeln werde nicht in die Urne zurückgelegt.

■ **Beispiel 2.4**
Eine Urne enthalte S schwarze "1" und W weiße Kugeln "0" Kugeln, $(E = \{0, 1\}, S + W = N)$.

Dann ist die Wahrscheinlichkeit in n Zügen ohne Zurücklegen gerade s schwarze und w weiße Kugeln zu ziehen

$$\rho(w) = \frac{\binom{W}{w}\binom{S}{s}}{\binom{N}{n}}, \quad 0 \leq s \leq S, 0 \leq w \leq W, s + w = n, S + W = N.$$

Beweis. Hausaufgabe! □

Definition 2.5
Seinen $N \in \mathbb{N}, W \leq N, n \leq N$, dann heißt die Verteilung auf $\{0, \dots, n\}$ mit Zähldichte

$$\rho(w) = \frac{\binom{wW}{w}\binom{N-W}{n-w}}{\binom{N}{n}}, \quad w = \max\{0, n = N + W\}, \dots, \min\{W, n\},$$

die Hypergeometrische Verteilung mit Parametern N, W, n. Wir schreiben $\mathrm{Hyper}(N, W, n)$.

3. Poisson-Approximation und Poisson-Verteilung

$\mathrm{Bin}(n, p)$ ist zwar explizit und elementar definiert, jedoch für große n mühsam auszuwerten. Für seltene Ereignisse (n groß, p klein) verwende daher:

Satz 3.1 (Poisson-Approximation)
Sei $\lambda > 0$ und $(p_n)_{n \in \mathbb{N}}$ eine Folge in $[0, 1]$ mit

$$np_n \to \lambda, \quad n \to \infty.$$

Dann gilt $\forall k \in \mathbb{N}_0$ für die Zähldichte der $\mathrm{Bin}(n, p_n)$-Verteilung

$$\lim_{n \to \infty} \binom{n}{k} p_n^k (1 - p)^{n-k} = e^{-\lambda} \frac{\lambda^k}{k!}.$$

Beweis. Sei $k \in \mathbb{N}_0$ fix, dann

$$\begin{aligned}
\binom{n}{k} = \frac{n!}{k!(n-k)!} &= \frac{n^k}{k!} \frac{n(n-1)\cdots(n-k+1)}{n^k} \\
&= \frac{n^k}{k!} \cdot 1 \cdot (1 - \frac{1}{n} \cdots \frac{k-1}{n}) \\
&\overset{n \to \infty}{\sim} \frac{n^k}{k!},
\end{aligned}$$

wobei $a(l) \overset{n \to \infty}{\sim} b(l) \Leftrightarrow \frac{a(l)}{b(l)} \xrightarrow{n \to \infty} 1$. Damit

$$\begin{aligned}
\binom{n}{k} p^k (1-p)^{n-k} &\overset{n \to \infty}{\sim} \frac{n^k}{k!} p_n^k (1 - p_n)^{n-k} \\
&\overset{n \to \infty}{\sim} \frac{\lambda^k}{k!} (1 - p_n)^n \\
&= \frac{\lambda^n}{k!} \left(1 - \frac{np_n}{n}\right)^n \\
&\xrightarrow{n \to \infty} \frac{\lambda^n}{k!} e^{-\lambda}.
\end{aligned}$$

□

Der erhaltene Grenzwert liefert die Zähldichte auf $\mathbb{N}_0$, denn

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

Definition 3.2

Sei $\lambda > 0$. Dann heißt das auf $(\mathbb{N}_0, \mathcal{P}(\mathbb{N}_0))$ definiert Wahrscheinlichkeitsmaß mit

$$\mathbb{P}(\{k\}) = \frac{\lambda^k}{k!} e^{-\lambda} \quad k \in \mathbb{N}_0,$$

<u>Poissonverteilung mit Parameter λ</u>. Schreibe Poisson(λ).

Die Poissonverteilung ist ein natürliches Modell für die Anzahl von zufälligen, seltenen Ereignissen (z.B. Tore im Fußballspiel, Schadensfälle einer Versicherung, ...).

Kapitel III

Bedingte Wahrscheinlichkeiten und (Un)-abbhängigk

1. Bedingte Wahrscheinlichkeiten

■ **Beispiel 1.1**

Das Würfeln mit zwei fairen, sechseitigen Würfeln können wir mit

$$\Omega = \{(i, j,), i, j \in \{1, \ldots, 6\}\}$$

und $\mathbb{P} = \mathrm{U}(\Omega)$. Da $|\Omega| = 36$ gilt also

$$\mathbb{P}(\{\omega\}) = \frac{1}{36} \quad \forall \omega \in \Omega.$$

Betrachte das Ereignis

$$A = \{(i, j) \in \Omega : i + j = 8\},$$

dann folgt

$$\mathbb{P}(A) = \frac{5}{36}.$$

Werden die beiden Würfel nach einander ausgeführt, so kann nach dem ersten Wurf eine Neubewertung der Wahrscheinlichkeit von A erfolgen.
Ist z.B.:

$$B = \{(i, j) \in \Omega, i = 4\}$$

eingetreten, so kann die Summe 8 nur durch eine weitere 4 realisiert werden, also mit Wahrscheinlichkeit

$$\frac{1}{6} = \frac{|A \cap B|}{|B|}.$$

Das Eintreten von B führt also dazu, dass das Wahrscheinlichkeitsmaß $\mathbb{P}$ durch ein neues Wahrscheinlichkeitsmaß $\mathbb{P}_B$ ersetzt werden muss. Hierbei sollte gelten:

Renormierung: $\mathbb{P}_B = 1$ (R)

Proportionalität: Für alle$A \subset \mathscr{F}$ mit $A \subseteq B$ gilt $\mathbb{P}_B(A) = c_B \mathbb{P}(A)$ mit einer Konstante c_B. (P)

Lemma 1.2
Sei $(\Omega, \mathscr{F}, \mathbb{P})$ Wahrscheinlichkeitsraum und $B \in \mathscr{F}$ mit $\mathbb{P}(B) > 0$. Dann gibt es genau ein Wahrscheinlichkeitsmaß $\mathbb{P}_B$ auf $(\Omega, \mathscr{F})$ mit den Eigenschaften (R) und (P). Dieses ist gegeben durch

$$\mathbb{P}_B(A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \quad \forall A \in \mathscr{F}.$$

Beweis. Offenbar erfüllt $\mathbb{P}_B$ wie definiert (R) und (P). Umgekehrt erfüllt $\mathbb{P}_B$ (R) und (P). Dann folgt für $A \in \mathscr{F}$:

$$\mathbb{P}_B(A) = \mathbb{P}_B(A \cap B) + \underbrace{\mathbb{P}_B(A \setminus B)}_{=0,\text{ wegen (R)}} \overset{(P)}{=} c_B \mathbb{P}(A \cap B).$$

Für $A = B$ folgt zudem aus (R)

$$1 = \mathbb{P}_B(B) = c_B \mathbb{P}(B)$$

also $c_B = \mathbb{P}(B)^{-1}$. $\square$

Kapitel IV

Test

Anhang