Stochastik – Wiederholung von Teil 1

Andrej Depperschmidt

Sommersemester 2016

Wahrscheinlichkeitsraum

Definition

Das Tripple $(\Omega, \mathcal{A}, P)$ heißt Wahrscheinlichkeitsraum, falls gilt:

(i) $\mathcal{A}$ ist eine σ-Algebra, d.h. $\mathcal{A} \subset \mathcal{P}(\Omega)$ ($=$ Potenzmenge von Ω) mit
  - $\Omega \in \mathcal{A}$,
  - $A \in \mathcal{A} \;\Rightarrow\; A^{c} \in \mathcal{A}$,
  - $A_1, A_2, \ldots \in \mathcal{A} \;\Rightarrow\; \cup_i A_i \in \mathcal{A}$.

(ii) P ist ein Wahrscheinlichkeitsmaß, d.h. $P : \mathcal{A} \to [0, 1]$ mit
  - $P(\Omega) = 1$
  - $A_1, A_2, \dots \in \mathcal{A}$ paarweise disjunkt, dann $P(\cup_i A_i) = \sum_i P(A_i)$

- P ist also eine Mengenfunktion, die jeder Menge $A \in \mathcal{A}$ einen Wert in $[0, 1]$ zuweist. Typischerweise reicht es P auf einer Teilmenge von $\mathcal{A}$ zu kennen um es vollständig zu beschreiben.
- Elemente von $\mathcal{A}$ heißen Ereignisse.

Laplace Modelle

- Ω endlich, $\mathcal{A} = \mathcal{P}(\Omega)$. ω's heißen Elementarereignisse (wir identifizieren ω mit Menge $\{\omega\}$).
- $P(\omega) = 1/|\Omega|$ für alle $\omega \in \Omega$ (P ist also die Gleichverteilung auf Ω). Dadurch ist P vollständig charakterisiert.
- Für alle $A \in \mathcal{A}$ ist dann ($\rightarrow$ Kombinatorik)

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\text{Anzahl günstiger Fälle}}{\text{Anzahl möglicher Fälle}}.$$

Beispiele: Münzwurf und Würfeln (fair), Lotto, viele andere Urnenmodelle (oder als solche modelierbare)

Allgemeine diskrete Wahrscheinlichkeitsräume

- Ω abzählbar (=endlich oder abzählbar unendlich), $\mathcal{A} = \mathcal{P}(\Omega)$.
- $P(\omega) := P(\{\omega\}) \in [0,1]$. Dadurch ist P vollständig beschrieben, denn es ist

$$P(A) = \sum_{\omega \in A} P(\omega), \quad A \in \mathcal{A}.$$

(Es muss natürlich $P(\Omega) = 1$ gelten!)

Beispiele:

- Bernoulli-Verteilung: $\Omega = \{0,1\}$ oder $\{\text{Miserfolg}, \text{Erfolg}\}$, $\{\text{Kopf}, \text{Zahl}\}$ etc. $P(1) = p \in [0,1]$, $P(0) = 1 - p$.
- Binomial-Verteilung: $\Omega = \{0, \ldots, n\}$,

$$P(k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

$P(k)$ ist die Wahrscheinlichkeit für genau k Erfolge bei n unabhängigen Bernoulli Experimenten.

- Poisson, Geometrische, Hypergeometrische Verteilungen ...

Bedingte Wahrscheinlichkeiten

Definition

Seien $A, B \in \mathcal{A}$ und $P(B) > 0$. Dann heißt

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

bedingte Wahrscheinlichkeit von A gegeben B.

- $P(\,\cdot \mid B)$ ist ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{A})$ mit $P(A \mid B) = 0$ für alle $A \in \mathcal{A}$ mit $A \cap B = \emptyset$.
- Gilt $P(A), P(B) > 0$ dann folgt („einfache Bayes-Formel“)

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)} = \frac{P(A \cap B)}{P(B)} \frac{P(B)}{P(A)} = \frac{P(A \mid B)P(B)}{P(A)}.$$

Sätze von der totalen Wahrscheinlichkeit und Bayes

Satz (Satz von der totalen Wahrscheinlichkeit)

Sei $(B_i)_{i \in I}$ *eine disjunkte Zerlegung von* Ω. *Dann gilt für alle* $A \in \mathcal{A}$

$$P(A) = \sum_{i \in I : P(B_i) > 0} P(A \mid B_i) P(B_i).$$

Satz (Satz von Bayes)

Sei $(B_i)_{i \in I}$ *eine disjunkte Zerlegung von* Ω *und* $A \in \mathcal{A}$ *mit* $P(A) > 0$*, dann gilt*

$$P(B_i \mid A) = \frac{P(B_i) P(A \mid B_i)}{\sum_{i \in I} P(A \mid B_i) P(B_i)}.$$

Unabhängigkeit

Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum.

Definition

Eine Familie von Ereignissen $(A_i)_{i \in I}$ heißt unabhängig wenn für *alle* endlichen $J \subset I$ gilt

$$P(\cap_{i \in J} A_i) = \prod_{i \in J} P(A_i).$$

Bemerkungen:

- „alle endlichen" ist in der Definition oben wichtig! Man kann Beispiele mit $\{A_1, A_2, A_3\}$ konstruieren mit

$$\textit{paarweise Unabhängigkeit} \quad \begin{array}{c} \not\Leftarrow \\ \not\Rightarrow \end{array} \quad P(\cap_{i=1}^{3} A_i) = \prod_{i=1}^{3} P(A_i).$$

- Sind A und B unabhängig und gilt $P(A), P(B) > 0$, dann folgt

$$P(A \mid B) = P(A) \quad \text{und} \quad P(B \mid A) = P(B).$$

Zufallsvariablen

Definition

Eine Abbildung $X : \Omega \to \mathbb{R}$ mit

$$X^{-1}([a,b]) = \{\omega \in \Omega \colon X(\omega) \in [a,b]\} \in \mathcal{A}, \quad \text{für alle } a,b \in \mathbb{R}$$

heißt Zufallsvariable.

Bemerkungen:

- Statt $[a,b]$ hätte man hier auch $(a,b]$, $(-\infty,a]$, alle offenen Mengen, alle abgeschlossenen Mengen und viel mehr fordern können.
- Zufallsvariablen sind messbare Abbildungen, hier messbar bezüglich der Borel-σ-Algebra auf $\mathbb{R}$, Bezeichnung $\mathcal{B}(\mathbb{R})$. Jedes Urbild einer Borel-Menge ist Element von $\mathcal{A}$.
- Mengen der Form $X^{-1}(B)$, $B \in \mathcal{B}$ kann auf $(\Omega, \mathcal{A}, P)$ die Wahrscheinlichkeit $P(X^{-1}(B)) = P(X \in B)$ zugeordnet werden.

Verteilung und Verteilungsfunktion

Definition

Ist X eine Zufallsvariable, so heißt $F : \mathbb{R} \to [0,1]$ definiert durch

$$F(x) := P(X \leq x), \quad x \in \mathbb{R}$$

die Verteilungsfunktion von X. Das auf $\mathcal{B}(\mathbb{R})$ definierte W-Maß (=Mengenfunktion)

$$P_X(B) := P(X \in B), \quad B \in \mathcal{B}(\mathbb{R})$$

heißt Verteilung von X.

Bemerkungen:

- Verteilungsfunktionen sind monoton, rechststetig, $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to \infty} F(x) = 1$.
- Verteilungsfunktion von X bestimmt die Verteilung von X eindeutig! ($\to$ Maßtheorie).

(Absolut) Stetige Zufallsvariablen

Definition

Eine Zufallsvariable X heißt absolut stetig, falls es eine Funktion $f : \mathbb{R} \to [0, \infty)$, genannt Dichte von X, gibt mit

$$\int_{\mathbb{R}} f(x)\, dx = 1 \quad \text{und} \quad F(b) = P(X \leq b) = \int_{-\infty}^{b} f(x)\, dx, \text{ für alle } b \in \mathbb{R}.$$

Beispiele:

- *Gleicheverteilung auf* $[a, b]$: $f(x) = \frac{1}{b-a} \mathbb{1}_{[a,b]}(x)$
- *Exponentialverteilung mit Parameter* $\lambda > 0$: $f(x) = \mathbb{1}_{[0,\infty)}(x) \lambda e^{-\lambda x}$
- *Normalverteilung mit Parametern* $\mu \in \mathbb{R}$ *und* $\sigma > 0$:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\Big\{-\frac{(x-\mu)^2}{2\sigma^2}\Big\}.$$

Erwartungswert diskreter Zufallsvariablen

Definition (Erwartungswert diskreter Zufallsvariablen)

Der Erwartungswert einer diskreten Zufallsvariablen X mit Werten in S ist definiert durch

$$E[X] = \sum_{a \in S} aP(X = a),$$

sofern die Summe wohldefiniert ist ($\pm\infty$ sind als Werte zugelassen).

Lemma (Transformation von Erwartungswerten)

Sei X eine diskrete Zufallsvariable mit $P(X \in S) = 1$ und $h : S \to \mathbb{R}$. Dann ist $h(X)$ eine diskrete Zufalls variable und für den Erwartungswert gilt (sofern wohldefiniert)

$$E[h(X)] = \sum_{a \in S} h(a)P(X = a).$$

Erwartungswert (absolut) stetiger Zufallsvariablen

Definition (Erwartungswert stetiger Zufallsvariablen)

Der Erwartungswert einer stetiger Zufallsvariablen X mit Dichte f ist gegeben durch

$$E[X] = \int_{\mathbb{R}} x f(x)\, dx,$$

sofern das Integral wohldefiniert ist ($\pm\infty$ sind als Werte zugelassen).

Lemma (Transformation von Erwartungswerten)

Sei X eine stetige Zufallsvariable mit $P(X \in S) = 1$ und $h : S \to \mathbb{R}$. Dann gilt für den Erwartungswert von $h(X)$ (sofern wohldefiniert)

$$E[h(X)] = \int_{\mathbb{R}} h(x) f(x)\, dx.$$

Rechnen mit Erwartungswerten

- Linearität: $E[aX + bY] = aE[X] + bE[Y]$
- Positivität: Sei $X \geq 0$ dann gilt
  (i) $E[X] \geq 0$,
  (ii) $E[X] = 0$ genau dann, wenn $P(X = 0) = 1$.
- Monotonie: Ist $X \leq Y$, dann $E[X] \leq E[Y]$.
- Jensen-Ungleichung: $h : \mathbb{R} \to \mathbb{R}$ konvex und $E[X]$ endlich, dann gilt

$$h(E[X]) \leq E[h(X)].$$

  Insbesondere gilt

$$|E[X]| \leq E[|X|] \quad \text{und} \quad (E[X])^2 \leq E[X^2].$$

Varianz und Rechnen mit Varianzen

Definition (Varianz)

Die Varianz einer Zufallsvariablen X ist definiert durch (Wohldefiniertheit vorausgesetzt)

$$\mathrm{Var}[X] = E[(X - E(X))^2]$$

Rechenregeln:

- $\mathrm{Var}[aX + b] = a^2 \mathrm{Var}[X]$
- $\mathrm{Var}[X] = 0$ genau dann, wenn $P(X = E[X]) = 1$
- $\mathrm{Var}[X] = E[X^2] - (E[X])^2$

Unabhängige Zufallsvariablen

Definition

Zufallsvariablen $X_1, \dots, X_n$ heißen (stochastisch) unabhängig falls alle Ereignisse $\{X_i \in A_i\}$, $A_i \in \mathcal{B}(\mathbb{R})$ die Produktform

$$P(X_1 \in A_1, \dots, X_n \in A_n) = P(X_1 \in A_1) \cdots P(X_n \in A_n)$$

haben.

Bemerkung:

1. Ist die Familie $X_1, \dots, X_n$ unabhängig, dann ist jede Teilfamilie unabhängig.
2. Ist die Familie $X_1, \dots, X_n$ unabhängig, dann auch die Familie $h_1(X_1), \dots, h_n(X_n)$ (für sinnvolle Funktionen $h_1, \dots, h_n$).