

Wahrscheinlichkeitstheorie und Statistik - Zusammenfassung

Julian Shen

26. Dezember 2021

1 Grundbegriffe

Definition: Ergebnisse und Ereignisse

- **Grundraum** ist eine nicht leere Menge $\Omega \neq \emptyset$ und enthält alle möglichen **Ergebnisse** eines Zufallsexperiments
- **Ereignisse** sind Teilmengen $A \subseteq \Omega$, denen eine Wahrscheinlichkeit zugeordnet werden kann. Falls ein ω Ergebnis ist, dann heißt $\{\omega\}$ **Elementarereignis**

Ereignisse können durch Mengenoperationen logisch verknüpft werden:

- $A \cup B$: Ereignis A oder B tritt ein („inklusives oder“)
- $A \cap B$: Ereignis A und B treffen ein
- $A \setminus B$: Ereignis A tritt ein, aber Ereignis B trifft nicht ein
- B^c : Ereignis B trifft nicht ein
- $A \subseteq B$: Wenn A eintritt, dann tritt auch B ein

Jedem Ereignis kann durch die **relative Häufigkeit** eine Wahrscheinlichkeit zugeordnet werden. Für n Wiederholungen und Ergebnisse $\omega_1, \dots, \omega_n \in \Omega$ gilt:

$$\mathbb{P}_n(A) := \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{\omega_i \in A\}}$$

Definition: Diskretes Wahrscheinlichkeitsmaß und Wahrscheinlichkeitsraum

Eine Abbildung $\mathbb{P} : \mathcal{P}(\Omega) \rightarrow [0, 1]$ heißt diskretes Wahrscheinlichkeitsmaß, falls

- $\mathbb{P}(\Omega) = 1$
- $\forall A_n \subseteq \Omega, n \in \mathbb{N}, \text{ disjunkt: } \mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_n)$ (σ -Additivität)
- es existiert eine abzählbare Menge $\Omega_0 \subseteq \Omega$ mit $\mathbb{P}(\Omega_0) = 1$

Dann heißt (Ω, \mathbb{P}) **diskreter Wahrscheinlichkeitsraum**.

Rechenregeln für diskrete Wahrscheinlichkeitsräume

- $\mathbb{P}(\emptyset) = 0$
- $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
- $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$
- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

Definition: Bernoulliverteilung

Wahrscheinlichkeitsverteilung heißt Bernoulliverteilung Ber_p mit Erfolgswahrscheinlichkeit p , wenn:

- Grundraum $\Omega = \{0, 1\}$
- $\mathbb{P}(1) = p$ für ein $p \in [0, 1]$

Es gilt $\mathbb{P}(\{0\}) = 1 - \mathbb{P}(\{1\}) = 1 - p$

Definition: Gleichverteilung

Das Wahrscheinlichkeitsmaß (Ω, \mathbb{P}) heißt Gleichverteilung oder **Laplace-Verteilung** U_A auf Ω , falls

- $\Omega \neq \emptyset$ endlich
- $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$, für $A \subseteq \Omega$

Urnenmodelle/Fächermodelle

Urnenmodell mit n Kugeln und k Ziehungen	mit Zurücklegen	ohne Zurücklegen	
mit Reihenfolge	n^k	$\frac{n!}{(n-k)!}$	unterscheidbare Murmeln
ohne Reihenfolge	$\binom{n+k-1}{k}$	$\binom{n}{k}$	ununterscheidbare Murmeln
	mit Mehrfachbelegung	ohne Mehrfachbelegung	Verteilung von k Murmeln auf n Fächer

Urnenmodelle ermöglichen es, die Wahrscheinlichkeiten zu bestimmen, falls von einer Gleichverteilung ausgegangen werden kann!

Definition: Zähldichte

Sei (Ω, \mathbb{P}) ein diskreter Wahrscheinlichkeitsraum. Dann wird die Funktion

$$f : \Omega \rightarrow [0, 1], f(\omega) = \mathbb{P}(\{\omega\})$$

Wahrscheinlichkeitsfunktion oder **Zähldichte** von \mathbb{P} genannt.

Diese besitzt folgende Eigenschaften:

- $\Omega_T := \{\omega \in \Omega \mid f(\omega) > 0\}$ ist abzählbar und heißt **Träger** von \mathbb{P} bzw. von f
- $\sum_{\omega \in \Omega} f(\omega) = 1$

Die Zähldichte ist eindeutig!

Definition: Binomialverteilung

Das Wahrscheinlichkeitsmaß $\mathbb{P} = \text{Bin}_{(n,p)}$ auf $\{0, \dots, n\}$ mit der Zähldichte

$$f(k) = \binom{n}{k} p^k \cdot (1-p)^{n-k} \quad \forall k \in \{0, \dots, n\}$$

heißt **Binomialverteilung** mit Parametern $n \in \mathbb{N}$ und $p \in [0, 1]$.

Definition: Geometrische Verteilung

Das Wahrscheinlichkeitsmaß $\mathbb{P} = \text{Geo}_p$ auf \mathbb{N}_0 mit der Zähldichte

$$f(k) = (1-p)^k \cdot p \quad \forall k \in \mathbb{N}_0$$

heißt **geometrische Verteilung** mit Parameter $p \in (0, 1]$.

2 Bedingte Wahrscheinlichkeit und Unabhängigkeit

Definition: Bedingte Wahrscheinlichkeit

Für (Ω, \mathbb{P}) diskreter Wahrscheinlichkeitsraum und $A, B \in \Omega$ mit $\mathbb{P}(B) > 0$ heißt

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

die **bedingte Wahrscheinlichkeit** von A gegeben B.

Multiplikationsformel

Seien $A_1, \dots, A_n \subseteq \Omega$ Ereignisse mit $\mathbb{P}(A_1 \cap \dots \cap A_{n-1}) > 0$, dann gilt

$$\mathbb{P}(A_1 \cap \dots \cap A_{n-1}) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2 \mid A_1) \cdot \mathbb{P}(A_3 \mid A_1 \cap A_2) \cdots \mathbb{P}(A_n \mid A_1 \cap \dots \cap A_{n-1})$$

Im Fall von $n=2$ gilt: $\mathbb{P}(A \cap B) = \mathbb{P}(B) \cdot \mathbb{P}(A \mid B)$

Satz von der totalen Wahrscheinlichkeit und Satz von Bayes

Sei (Ω, \mathbb{P}) ein diskreter Wahrscheinlichkeitsraum, I eine abzählbare Indexmenge, $B_i \subseteq \Omega, i \in I$, disjunkt mit $\mathbb{P}(B_i) > 0$ und $\bigcup_{i \in I} B_i = \Omega$ und $A \subseteq \Omega$ beliebig.

- Es gilt der **Satz von der totalen Wahrscheinlichkeit**:

$$\mathbb{P}(A) = \sum_{i \in I} \mathbb{P}(A \mid B_i) \cdot \mathbb{P}(B_i)$$

- Falls $\mathbb{P}(A) > 0$ und $k \in I$, dann gilt der **Satz von Bayes**:

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

Definition: Stochastische Unabhängigkeit

Sei (Ω, \mathbb{P}) ein diskreter Wahrscheinlichkeitsraum. Zwei Ereignisse $A, B \subseteq \Omega$ heißen **stochastisch unabhängig**, falls

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$$

Ereignisse $A_1, \dots, A_n \subseteq \Omega$ heißen **stochastisch unabhängig**, wenn für jede Indexmenge $I \subseteq \{1, \dots, n\}, I \neq \emptyset$, gilt

$$\mathbb{P}\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} \mathbb{P}(A_i)$$

Achtung: Mehr als zwei Ereignisse A_1, \dots, A_n sind im Allgemeinen nicht stochastisch unabhängig, wenn nur $\mathbb{P}\left(\bigcap_{i=1}^n A_i\right) = \prod_{i=1}^n \mathbb{P}(A_i)$ gilt! Gleiches gilt, wenn jeweils nur zwei der Ereignisse stochastisch unabhängig sind.

3 Zufallsvariablen und ihre Verteilungen

Definition: S -wertige Zufallsvariable

Ist (Ω, \mathbb{P}) ein diskreter Wahrscheinlichkeitsraum und $S \neq \emptyset$ eine beliebige Menge, so wird die Abbildung $X : \Omega \rightarrow S$ **S -wertige Zufallsvariable** genannt.

Definition: Verteilung

Ist $X : \Omega \rightarrow S$ eine Zufallsvariable auf einem diskreten Wahrscheinlichkeitsraum (Ω, \mathbb{P}) , dann wird durch

$$\mathbb{P}^X(B) := \mathbb{P}(X^{-1}(B)) \quad \forall B \subseteq S$$

ein Wahrscheinlichkeitsmaß \mathbb{P}^X auf S definiert, welches Verteilung von X genannt wird. (S, \mathbb{P}^X) ist ein diskreter Wahrscheinlichkeitsraum. Notation für Urbilder:

- $\{X \in B\} := \{\omega \in \Omega \mid X(\omega) \in B\} = X^{-1}(B)$
- $\{X = x\} := \{\omega \in \Omega \mid X(\omega) = x\} = X^{-1}(\{x\})$
- $\{X > x\} := \{\omega \in \Omega \mid X(\omega) > x\} = X^{-1}((x, \infty))$

Zudem schreibt man $\mathbb{P}(X \in B) := \mathbb{P}(\{X \in B\})$

Definition: Stochastische Unabhängigkeit von Zufallsvariablen

Sei (Ω, \mathbb{P}) ein diskreter Wahrscheinlichkeitsraum und $S_i, i \in \{1, \dots, n\}$ nichtleere Zufallsvariablen. Zufallsvariablen $X_i : \Omega \rightarrow S_i, i \in \{1, \dots, n\}$ heißen **stochastisch unabhängig**, wenn für beliebige $B_i \subseteq S_i$ die Ereignisse $\{X_1 \in B_1\}, \dots, \{X_n \in B_n\}$ stochastisch unabhängig sind.

Auch der Vektor $(X_1, \dots, X_n) : \Omega \rightarrow S_1 \times \dots \times S_n$ ist eine Zufallsvariable mit Verteilung $\mathbb{P}^{(X_1, \dots, X_n)}$ auf $S_1 \times \dots \times S_n$.

Bemerkung zur Schreibweise: $(X_1 \in B_1, X_2 \in B_2) = (X_1 \in B_1) \cap (X_2 \in B_2)$

Satz für stochastisch unabhängige Zufallsvariablen

Es sind äquivalent:

- X_1, \dots, X_n sind stochastisch unabhängig
- $\forall B_i \subseteq S_i : \mathbb{P}(X_i \in B_i \mid \forall 1 \leq i \leq n) = \prod_{i=1}^n \mathbb{P}(X_i \in B_i)$
- Bezeichne mit f_{X_i} die Zähldichte von \mathbb{P}^{X_i} auf S_i . Dann hat die Zähldichte $f_{(X_1, \dots, X_n)}$ von $\mathbb{P}^{(X_1, \dots, X_n)}$ die Form: $f_{(X_1, \dots, X_n)}(t_1, \dots, t_n) = \prod_{i=1}^n f_{X_i}(t_i) \quad \forall t_i \in S_i$

Definition: Hypergeometrischen Verteilung

Das Wahrscheinlichkeitsmaß $\mathbb{P} = \text{Hyp}_{(N,M,n)}$ auf \mathbb{N}_0 gegeben durch die Zähldichte

$$\mathbb{P}^X(\{m\}) = \mathbb{P}(X = m) = \frac{\binom{M}{m} \cdot \binom{N-M}{n-m}}{\binom{N}{n}} \mathbb{1}_S(m) \quad \forall m \in \mathbb{N}_0$$

heißt **hypergeometrische Verteilung**.

Zusammenhang hypergeometrische Verteilung und Binomialverteilung

- Die **hypergeometrische Verteilung** $\text{Hyp}_{(N,M,n)}$ beschreibt die Anzahl der markierten Gegenstände bei n -maligem **Ziehen ohne Zurücklegen** aus N Gegenständen, von denen M markiert sind
- Die **Binomialverteilung** $\text{Bin}_{(n,M/N)}$ beschreibt die Anzahl der markierten Gegenstände bei n -maligem **Ziehen mit Zurücklegen** aus N Gegenständen, von denen M markiert sind

Falls $n \ll N$, dann ist Ziehen mit oder ohne Zurücklegen fast identisch und daher $\text{Hyp}_{(N,M,n)}(\{m\}) \approx \text{Bin}_{(n, \frac{M}{N})}(\{m\}) \quad \forall 0 \leq m \leq n$

Poisson'scher Grenzwertsatz

Für eine große Anzahl an Experimenten n und eine kleine Erfolgswahrscheinlichkeit p kann $\text{Bin}_{(n,p)}$ durch eine strukturell einfachere Verteilung approximiert werden:

$$\lim_{n \rightarrow \infty} \text{Bin}_{(n,p)}(\{k\}) = \lim_{n \rightarrow \infty} \binom{n}{k} p^k (1-p)^{n-k} = e^{-\lambda} \cdot \frac{\lambda^k}{k!} =: \text{Poiss}_\lambda(\{k\}) \quad k \in \mathbb{N}_0$$

Das Wahrscheinlichkeitsmaß Poiss_λ heißt Poissonverteilung mit Parameter λ .

Definition: Faltung

Sind X, Y \mathbb{R} -wertige Zufallsvariablen auf einem diskreten Wahrscheinlichkeitsraum mit Zähldichten f_X von \mathbb{P}^X und f_Y von \mathbb{P}^Y , dann heißt

$$(f_X * f_Y)(z) = \sum_{x \in \mathbb{R}: f_X(x) > 0} f_X(x) \cdot f_Y(z - x) \quad \forall z \in \mathbb{R}$$

die **Faltung** von f_X und f_Y . Hierbei ist $f_X * f_Y$ wieder eine Zähldichte mit Träger $\Omega_T := \{z \in \mathbb{R} \mid \exists x, y \in \mathbb{R} : z = x + y, f_X(x) > 0, f_Y(y) > 0\}$ und die zugehörige diskrete Verteilung $\mathbb{P}^X * \mathbb{P}^Y$ heißt **Faltung** von \mathbb{P}^X und \mathbb{P}^Y .

Satz für die Faltung

Sind X, Y unabhängige \mathbb{R} -wertige Zufallsvariablen auf einem diskreten Wahrscheinlichkeitsraum, so gilt

$$\mathbb{P}^X * \mathbb{P}^Y = \mathbb{P}^{X+Y}$$

4 Wahrscheinlichkeitsmaße auf \mathbb{R}

Kontinuierliche Ergebnisse lassen sich nicht mehr durch eine abzählbare Anzahl an Versuchsausgängen beschreiben. Wahrscheinlichkeiten kann man dann nur noch „gutartigen Mengen“ zuordnen, u.a.:

- Intervalle sind gutartig
- Komplemente gutartiger Mengen sind gutartig
- Abzählbare Vereinigungen gutartiger Mengen sind gutartig

Bezeichne nun mit $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ das System aller „gutartigen“ Mengen.

Definition: σ -Algebra

Sei $\Omega \neq \emptyset$ ein beliebiger Grundraum. Eine Menge $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ heißt **σ -Algebra** auf Ω , falls:

- $\emptyset, \Omega \in \mathcal{A}$
- $A \in \mathcal{A} \implies A^c \in \mathcal{A}$
- $A_n \in \mathcal{A} \forall n \in \mathbb{N} \implies \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$

(Ω, \mathcal{A}) heißt dann **messbarer Raum**. Die Mengen $A \in \mathcal{A}$ heißen **Ereignisse**.

Definition: Borel- σ -Algebra

Die **Borel- σ -Algebra** \mathcal{B} auf \mathbb{R} beschreibt die kleinste σ -Algebra, welche alle Intervalle $(a, b]$ für beliebige $a, b \in \mathbb{R}$ enthält.

Definition: Wahrscheinlichkeitsmaß und Wahrscheinlichkeitsraum

Sei (Ω, \mathcal{A}) ein Messraum mit Grundraum $\Omega \neq \emptyset$ und σ -Algebra \mathcal{A} . Eine Abbildung $\mathbb{P} : \mathcal{A} \rightarrow [0, 1]$ heißt **Wahrscheinlichkeitsmaß** auf (Ω, \mathcal{A}) , falls

- $\mathbb{P}(\Omega) = 1$
- $A_n \in \mathcal{A}, n \in \mathbb{N}, \text{ disjunkt} \implies \mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} \mathbb{P}(A_n)$ (σ -Additivität)

$(\Omega, \mathcal{A}, \mathbb{P})$ heißt dann **Wahrscheinlichkeitsraum**.

Sätze und Definitionen für allgemeine Wahrscheinlichkeitsräume

Folgende Sätze und Definitionen übertragen sich sinngemäß, wobei als Ereignisse jeweils nur Mengen aus \mathcal{A} betrachtet werden:

- Rechenregeln für diskrete Wahrscheinlichkeitsmaße
- Bedingte Wahrscheinlichkeiten
- Satz von der Totalen Wahrscheinlichkeit und von Bayes
- Stochastische Unabhängigkeit von Ereignissen

Unterschied: Während diskrete Wahrscheinlichkeitsmaße \mathbb{P} vollständig durch die Zähldichte $f(\omega) := \mathbb{P}(\{\omega\}), \omega \in \Omega$ bestimmt sind, ist dies für allgemeine Wahrscheinlichkeitsmaße falsch!