https://matse.paddel.xyz/spicker

Stochastik

Patrick Gustav Blaneck

Letzte Änderung: 6. Dezember 2021

Inhaltsverzeichnis

1 Wahrscheinlichkeitsrechnung

1.1 Einführung in die Kombinatorik

Bonus: Urnenmodell

Ein *Urnenmodell* ist ein Gedankenexperiment, das in der Wahrscheinlichkeitstheorie und in der Statistik verwendet wird, um verschiedene Zufallsexperimente auf einheitliche und anschauliche Weise zu modellieren.

Dazu wird ein fiktives Gefäß, Urne genannt, mit einer bestimmten Anzahl an Kugeln gefüllt, die anschließend zufällig gezogen werden. Damit ist gemeint, dass bei jedem Zug alle in der Urne befindlichen Kugeln die gleiche Wahrscheinlichkeit haben, ausgewählt zu werden. Dadurch kann die Bestimmung interessierender Wahrscheinlichkeiten auf die Lösung kombinatorischer Abzählprobleme zurückgeführt werden.

Definition: Permutation

Eine Anordnung von n verschiedenen Elementen in einer bestimmten Reihenfolge heißt eine *Permutation* der n Elemente.

Für eine n-elementige Menge gilt für die Anzahl P der Permutationen:

$$P(n) = n \cdot (n-1) \cdot (n-1) \cdot \ldots \cdot 1 = n!$$

Befinden sich unter den n Elementen jeweils $n_1, n_2, \ldots, n_k$ gleiche, dann gilt:

$$P(n; n_1; n_2; \ldots, n_k) = \frac{n!}{n_1! \cdot n_2! \cdot \ldots \cdot n_k} \quad \text{mit} \sum_{i=1}^{k} n_i = n \quad k \leq n$$

Definition: Kombination

Eine *Kombination* oder ungeordnete Stichprobe ist in der Kombinatorik eine Auswahl von Objekten aus einer gegebenen Grundmenge, die (im Gegensatz zur Permutation) nicht alle Objekte der Grundmenge enthalten muss und bei der (im Gegensatz zur Permutation und Variation) die Reihenfolge unberücksichtigt bleibt.

Darf jedes Objekt nur einmal auftreten, spricht man von einer *Kombination ohne Wiederholung*.

Um alle *Kombinationen k-ter Ordnung ohne Wiederholung* zu erhalten, müssen alle Permutationen der n Kugeln betrachtet werden, wobei ein Vertauschen der Kugeln auf Plätzen mit demselben Merkmal (wird gezogen $\rightarrow k$ bzw. wird nicht gezogen $\rightarrow n-k$) keine neue Kombination ergibt:

$$C(n;k) = P(n;k;n-k) = \frac{n!}{k! \cdot (n-k)!} = \binom{n}{k}$$

Können Objekte mehrfach ausgewählt werden, so spricht man von einer *Kombination mit Wiederholung*.

Um alle *Kombinationen k-ter Ordnung mit Wiederholung* zu erhalten, gilt: [a]

$$C_W(n;k) = P(n+k-1;k;n-1) = \frac{(n+k-1)!}{k! \cdot (n-1)!} = \binom{n+k-1}{k}$$

[a] Zum Verständnis empfehle ich dieses Video: COMBINATIONS with REPETITION - DISCRETE MATHEMATICS

Definition: Variation

Eine *Variation* oder geordnete Stichprobe ist eine Auswahl von k Objekten aus einer Menge von n Objekten, wobei die Reihenfolge der Auswahl eine Rolle spielt.

Bei einer *Variation ohne Wiederholung* sollen k Objekten (mit $k \leq n$) auf k verfügbare Plätze platziert werden, wobei jedes Objekt nur höchstens einen Platz einnehmen darf. Es gibt für den ersten Platz n mögliche Objekte, für den zweiten Platz $n-1$ Objekte usw. bis zum k-ten Platz, für den es noch $n-k+1$ mögliche Objekte gibt. Insgesamt gilt also:

$$V(n;k) = k! \cdot C(n;k) = \frac{n!}{(n-k)!}$$

Bei einer Variation mit Wiederholung werden aus n Objekten k Objekte unter Beachtung der Reihenfolge ausgewählt, wobei Objekte auch mehrfach ausgewählt werden können. Nachdem jedes der n Objekte auf jedem der k Plätze der Auswahl erscheinen kann, gilt demzufolge:

$$V_W(n;k) = n^k$$

Bonus: Kombinatoriktabelle

TO DO

1.2 Grundbegriffe

Definition: Zufallsexperiment

Damit ein Experiment ein *Zufallsexperiment* ist, muss es folgende Eigenschaften aufweisen:

- Es gibt einen genau festgelegten Plan zur Durchführung.
- Alle möglichen Ergebnisse des Experiments sind vorab bekannt.
- Das Ergebnis jedes einzelnen Experiments kann nicht vorhergesagt werden (Zufälligkeit).

Ein Zufallsexperiment kann einmalig und unwiederholbar sein oder auch Serien von Durchführungen mit gleichwertigen und von Durchführung zu Durchführung voneinander unabhängigen Versuchen ermöglichen.

Weiter kann ein Zufallsexperiment *einstufig* oder *mehrstufig* sein.

Definition: Ergebnismenge

Die Menge aller möglichen Ergebnisse eines Zufallsexperiments bezeichnen wir als *Ergebnismenge* Ω.

Definition: Ereignis

Interessieren wir uns nicht für alle möglichen Ergebnisse des Zufallsexperiments, sondern nur für bestimmte, z.B. „Würfeln einer geraden Zahl", so sprechen wir von einem *Ereignis* ω.

Spezialfall: Einelementige Teilmengen von Ω sind *Elementarereignisse*

$$\omega \subseteq \Omega \text{ mit } |\omega| = 1$$

Ein Ereignis ist A ist entweder:

- Das *unmögliche* Ereignis (A enthält kein Element von Ω)
- Ein *Elementarereignis* (A enthält genau ein Element von Ω)
- Eine *Zusammenfassung* mehrerer Elementarereignisse (A enthält mehrere Elemente von Ω)
- Das *sichere* Ereignis (A enthält alle Elemente von Ω bzw. $A = \Omega$)

Definition: Ereignisraum

Die Menge aller Ereignisse, die sich aus der Ergebnismenge bilden lässt, heißt *Ereignisraum*.

Bonus: Zusammengesetzte Ereignisse

Da Ereignisse Teilmengen der Ergebnismenge sind, lassen sie sich auch wie Mengen verknüpfen. Wir erhalten dadurch zusammengesetzte Ereignisse:

1. *Vereinigung*:
$$A \cup B = \{\omega \mid \omega \in A \vee \omega \in B\}$$
2. *Durchschnitt* bzw. *Schnittmenge*:
$$A \cap B = \{\omega \mid \omega \in A \wedge \omega \in B\}$$
3. *Gegenereignis* bzw. *Komplement*:
$$\overline{A} = \{\omega \mid \omega \notin A\}$$
4. *Differenz*:
$$A \setminus B = \{\omega \mid \omega \in A \wedge \omega \notin B\}$$
5. *Disjunkte Ereignisse*:
$$A \cap B = \emptyset \quad \Longleftrightarrow \quad A \text{ und } B \text{ sind disjunkt}$$

Analog zur Mengenalgebra gelten hier natürlich auch die Regeln von de Morgan[a].

[a] Siehe: De-morgansche Gesetze

1.3 Wahrscheinlichkeit

Definition: Laplace-Experiment

Es gibt eine Reihe von Zufallsexperimenten, bei denen keines der Elementarereignisse gegenüber einem anderen bevorzugt ist, d.h. bei ausreichend häufiger Wiederholung des Experimentes tritt jedes Elementarereignis mit nahezu gleicher Häufigkeit auf. Ein derartiges Experiment bezeichnen wir als *Laplace-Experiment*.

Einem Elementarereignis ω_i aus einer Ergebnismenge Ω mit m möglichen Elementarereignissen wird definitionsgemäß die positive Zahl

$$P(\{\omega_i\}) = p(\omega_i) = \frac{1}{m} = \frac{1}{|\Omega|}$$

als *Wahrscheinlichkeit* zugeordnet. Damit gilt für ein Ereignis A direkt:

$$P(A) = \sum_{\omega_i \in A} P(\{\omega_i\}) = p(\omega_i) = \frac{|A|}{m} = \frac{|A|}{|\Omega|}$$

Das gilt allerdings nur, wenn:

- Die Ergebnismenge Ω endlich ist.
- Alle Elementarereignisse gleichwahrscheinlich sind.

Definition: Absolute Häufigkeit

Die *absolute Häufigkeit* ist das Ergebnis einer einfachen Zählung von Objekten oder Ereignissen (besser Elementarereignissen). Sie gibt an, wie viele Elemente mit dem gleichen interessierenden Merkmal gezählt wurden.

Definition: Relative Häufigkeit

Die *relative Häufigkeit* gibt den Anteil der Elemente einer Menge wieder, bei denen eine bestimmte Merkmalsausprägung vorliegt. Sie wird berechnet, indem die absolute Häufigkeit eines Merkmals in einer zugrundeliegenden Menge durch die Anzahl der Objekte in dieser Menge geteilt wird. Die relative Häufigkeit ist also eine Bruchzahl und hat einen Wert zwischen 0 und 1.

Definition: Wahrscheinlichkeitsaxiome

Seien A, B, E Ereignisse. Es gilt allgemein:

- $P(\emptyset) = 0$
- $P(\bar{E}) = 1 - P(E)$
- $0 \leq P(E) \leq 1$
- $A \subseteq B \implies P(A) \leq P(B)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B) \leq P(A) + P(B)$

Definition: Ereignisalgebra

Es sei Ω eine nichtleere Menge. Eine Ereignisalgebra über Ω ist eine nichtleere Menge S von Teilmengen von Ω, für die gilt:

- Für jedes $E \in S$ ist $\bar{E} \in S$.
- Für jede Folge $E_1, E_2, \ldots \in S$ ist $\bigcup_{E_i \in \Omega} E_i \in S$.

Folgerungen:

- Für jede Folge $E_1, E_2, \ldots \in S$ ist $\bigcap_{E_i \in \Omega} E_i \in S$.
- $\emptyset \in S$ und $\Omega \in S$.
- Für jede nichtleere Menge Ω ist die Potenzmenge $\mathcal{P}(\Omega)$ eine Ereignisalgebra über Ω.

Ist Ω endlich oder abzählbar unendlich, wählt man als Ereignisalgebra stets die Potenzmenge.

Bonus: Kolmogoroff-Axiome

Es sei Ω eine Ergebnismenge und S eine Ereignisalgebra über Ω. Eine Zuordnungsvorschrift $P : S \to \mathbb{R}$ heißt *Wahrscheinlichkeitsmaß*, wenn gilt:

1. $\forall E \in S : 0 \leq P(E) \leq 1$.
2. $P(\Omega) = 1$
3. Falls $E_1, E_2, \ldots$ disjunkte Ereignisse sind, gilt

$$P\left(\bigcup_i E_i\right) = \sum_i P(E_i) \quad \text{„}\sigma\text{-Additivität“}$$

Das Tripel (Ω, S, P) mit Ω als Ergebnismenge, S als Ereignisalgebra und P als Wahrscheinlichkeitsmaß bezeichnet man als *Wahrscheinlichkeitsraum*.

Definition: Bedingte Wahrscheinlichkeit

In vielen Anwendungen ist das Eintreten eines Ereignisses A nicht unabhängig davon, ob vorher ein anderes Ereignis B eingetreten ist oder nicht. Man spricht dann von einer *bedingten Wahrscheinlichkeit* und schreibt $P(A \mid B)$. Es gilt:

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

Generell gilt $P(A \mid B) \neq P(B \mid A)$, aber:

$$\begin{aligned} & P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B \cap A)}{P(B)} = \frac{P(B \cap A)}{P(A)} \cdot \frac{P(A)}{P(B)} \\ \Longrightarrow \quad & P(A \mid B) = P(B \mid A) \cdot \frac{P(A)}{P(B)} \end{aligned}$$

Alle Axiome für eine Wahrscheinlichkeit werden von der bedingten Wahrscheinlichkeit erfüllt. Damit gilt auch:

- $E_1 \subseteq E_2 \implies P(E_1 \mid B) \leq P(E_2 \mid B)$
- $P(\overline{A} \mid B) = 1 - P(A \mid B)$

Definition: Multiplikationssatz

Löst man die Definition der bedingten Wahrscheinlichkeit auf nach der Wahrscheinlichkeit für das gleichzeitige Eintreten zweier Ereignisse A und B, so ergibt sich der *Multiplikationssatz* der Wahrscheinlichkeitsrechnung:

$$P(A \cap B) = P(A \mid B) \cdot P(B) = P(B \mid A) \cdot P(A)$$

bzw.

$$P(A \cap B \cap C) = P(A) \cdot P(B \mid A) \cdot P(C \mid A \cap B)$$

Definition: Stochastische Unabhängigkeit

In einigen Anwendungen ist die Wahrscheinlichkeit für das Eintreten des Ereignisses A unabhängig davon, ob B eingetreten ist oder nicht:

$$P(A \mid B) = P(A \mid \overline{B}) = P(A)$$

Aus dem Multiplikationssatz ergibt sich damit als Definition für die *stochastische Unabhängigkeit* zweier Ereignisse A und B:

$$P(A \cap B) = P(A) \cdot P(B)$$

Bonus: Vierfeldertafel

	B	$\overline{B}$	Zeilensumme
A	$P(A \cap B)$	$P(A \cap \overline{B})$	$P(A)$
$\overline{A}$	$P(\overline{A} \cap B)$	$P(\overline{A} \cap \overline{B})$	$P(\overline{A})$
Spaltensumme	$P(B)$	$P(\overline{B})$	1

Bei vollständiger Unabhängigkeit der Ereignisse A und B voneinander gilt:

	B	$\overline{B}$	Zeilensumme
A	$P(A) \cdot P(B)$	$P(A) \cdot P(\overline{B})$	$P(A)$
$\overline{A}$	$P(\overline{A}) \cdot P(B)$	$P(\overline{A}) \cdot P(\overline{B})$	$P(\overline{A})$
Spaltensumme	$P(B)$	$P(\overline{B})$	1

Bonus: Zusammenhangskoeffizient

Der *Zusammenhangskoeffizient* bzw. das *Assoziationsmaß* $Q \in [-1;1]$ sei definiert als:

$$Q = \frac{P(A \cap B) \cdot P(\overline{A} \cap \overline{B}) - P(A \cap \overline{B}) \cdot P(\overline{A} \cap B)}{P(A \cap B) \cdot P(\overline{A} \cap \overline{B}) + P(A \cap \overline{B}) \cdot P(\overline{A} \cap B)}$$

Definition: Mehrstufiges Zufallsexperiment

Kompliziertere Zufallsprozesse bestehen häufig aus mehreren nacheinander ablaufenden Zufallsexperimenten bzw. einem *mehrstufigen Zufallsexperiment*.

Ein wichtiges Hilfsmittel dabei sind *Ereignisbäume*.

Beispiel: Ereignisbaum

TO DO

Definition: Totale Wahrscheinlichkeit

Sind nur bedingte Wahrscheinlichkeiten und die Wahrscheinlichkeiten des bedingenden Ereignisses bekannt, ergibt sich die totale Wahrscheinlichkeit von B aus:

$$P(B) = \sum_{i=1}^{n} P(A_i) \cdot P(B \mid A_i)$$

TO DO (Graph, siehe Folie 67)

Definition: Bayes'sche Formel

Für den Zusammenhang zwischen $P(A \mid B)$ und $P(B \mid A)$ ergibt sihc direkt aus der Definition und dem Multiplikationssatz die *Bayes'sche Formel* bzw. der *Satz von Bayes*:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B \cap A)}{P(B)} = \frac{P(B \mid A) \cdot P(A)}{P(B)} = \frac{P(B \mid A) \cdot P(A)}{\sum_{i=1}^{n} P(A_i) \cdot P(B \mid A_i)}$$

Die Wahrscheinlichkeit, dass B über einen bestimmten Pfad eintritt, ergibt sich als Verhältnis der Wahrscheinlichkeit für diesen Pfad zur totalen Wahrscheinlichkeit von B.

1.4 Wahrscheinlichkeitsverteilung einer Zufallsvariablen

Definition: Zufallsvariable

Definition: Verteilungsfunktion

Definition: Wahrscheinlichkeitsfunktion

Definition: Dichtefunktion

Definition: Transformierte Zufallsvariable

Algorithmus: Transformation der Verteilungsfunktion

Algorithmus: Transformation der Dichtefunktion

Definition: Erwartungswert

Definition: Quantile

Definition: Varianz

Definition: Tschebyscheff-Ungleichung

1.5 Spezielle Wahrscheinlichkeitsverteilungen

Definition: Gleichverteilung

Definition: Bernoulli-Experiment

Definition: Binomialverteilung

Definition: Hypergeometrische Verteilung

Definition: Poisson-Verteilung

Definition: Gaußsche Normalverteilung

Definition: Standardnormalverteilung

Bonus: Geometrische Verteilung

Bonus: Negative Binomialverteilung

Bonus: Pascal-Verteilung

Definition: Approximation von Verteilungen

1.6 Mehrdimensionale Zufallsvariablen

1.7 Kovarianz und Korrelation

1.8 Gesetze der großen Zahlen und Grenzwertsätze

2 Beschreibende Statistik

2.1 Merkmale und weitere wichtige Begriffe

2.2 Darstellung der Beobachtungsergebnisse

2.3 Statistische Maßzahlen

3 Schließende Statistik

3.1 Grundbegriffe

3.2 Punktschätzungen

3.3 Intervallschätzungen

3.4 Statistische Testverfahren

Index

Beispiele