Vorlesung Statistik I

Florian Wickelmaier

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN



Wintersemester 2026/2027

Laplace-Experiment

Im Folgenden wollen wir die Ergebnisse der kombinatorischen Überlegungen im Kontext einiger beispielhafter Laplace-Experimente ausnutzen

Insbesondere betrachten wir

- Die Gewinnchancen beim Lotto 6 aus 49
- Das sogenannte Geburtstagsproblem

  Wie wahrscheinlich ist es, dass von k ausgewählten Personen mindestens zwei am gleichen Tag des Jahres Geburtstag haben?

Lotto 6 aus 49

Wir haben bereits die Anzahl der möglichen Ergebnisse einer Lottoausspielung ermittelt

$$|\Omega| = \binom{49}{6} = 13983816$$

Für das Ereignis “6 richtige” gibt es nur einen günstigen Fall, daher gilt

$$P(6 \text{ “richtige”}) = \frac{1}{13983816} = 0.715 \cdot 10^{-7}$$

Lotto 6 aus 49

Wahrscheinlichkeit des Ereignisses “5 richtige”

$$P(5 \text{ “richtige”}) = \frac{\binom{6}{5} \cdot \binom{43}{1}}{13983816} = \frac{258}{13983816} = 0.184 \cdot 10^{-4}$$

Wahrscheinlichkeit des Ereignisses “4 richtige”

$$P(4 \text{ “richtige”}) = \frac{\binom{6}{4} \cdot \binom{43}{2}}{13983816} = \frac{13545}{13983816} = 0.969 \cdot 10^{-3}$$

Wahrscheinlichkeit des Ereignisses “3 richtige”

$$P(3 \text{ “richtige”}) = \frac{\binom{6}{3} \cdot \binom{43}{3}}{13983816} = \frac{246820}{13983816} = 0.01765$$

Lotto 6 aus 49: Zusatzzahl

Früher wurde neben den "6 richtigen" auch eine Zusatzzahl gezogen, sodass die Ereignisse "5 richtige" mit bzw. ohne Zusatzzahl unterschieden wurden

Bei "5 richtigen" mit Zusatzzahl müssen die Zusatzzahl und 5 der "6 richtigen" angekreuzt sein, wofür es $\binom{6}{5}$ Möglichkeiten gibt

$$P(\text{5 "richtige" mit Zusatzzahl}) = \frac{\binom{6}{5}\binom{1}{1}\binom{42}{0}}{13983816} = \frac{6}{13983816}$$

Bei "5 richtigen" ohne Zusatzzahl muss anstatt der Zusatzzahl eine der 42 nicht gezogenen Zahlen angekreuzt sein

$$P(\text{5 "richtige" ohne Zusatzzahl}) = \frac{\binom{6}{5}\binom{1}{0}\binom{42}{1}}{13983816} = \frac{252}{13983816}$$

Urnenmodell: Ziehung ohne Zurücklegen

Eine Urne enthalte N Kugeln, von denen K rot und die restlichen $N - K$ grün sind

Es werden n Kugeln ohne Zurücklegen gezogen

Wir betrachten die Wahrscheinlichkeit für das Ereignis A_k, dass sich unter den gezogenen Kugeln k rote befinden

$$P(A_k) = \frac{\binom{K}{k} \cdot \binom{N-K}{n-k}}{\binom{N}{n}}$$

Urnenmodell: Ziehung mit Zurücklegen

Eine Urne enthalte N Kugeln, von denen K rot und die restlichen $N - K$ grün sind

Es werden n Kugeln mit Zurücklegen gezogen

Wir betrachten die Wahrscheinlichkeit für das Ereignis A_k, dass sich unter den gezogenen Kugeln k rote befinden

$$P(A_k) = \binom{n}{k} \left(\frac{K}{N}\right)^k \left(1 - \frac{K}{N}\right)^{n-k}$$

Geburtstagsproblem

Wie wahrscheinlich ist es, dass von k ausgewählten Personen mindestens zwei am gleichen Tag des Jahres Geburtstag haben (Ereignis A_k)?

Wir nehmen an, dass das Jahr 365 Tage hat, die als Geburtstage für jede der k Personen gleich wahrscheinlich sind

Es gilt offensichtlich

$$P(A_k) = 1 \quad \text{für} \quad k \geq 366$$

Für $k \leq 365$ betrachten wir das komplementäre Ereignis $\overline{A}_k$, das eintritt, wenn alle k Personen an unterschiedlichen Tagen Geburtstag haben

Geburtstagsproblem

Tritt das Ereignis $\overline{A}_k$ ein, dann kommen für die erste Person 365 Tage, für die zweite Person 364 Tage, ..., und für die k-te Person $365 - k + 1$ Tage in Frage

Mit $365 \cdot 364 \cdot \ldots \cdot (365 - k + 1)$ günstigen Fällen und 365^k insgesamt möglichen Fällen hat man

$$P(\overline{A}_k) = \frac{365 \cdot \ldots \cdot (365 - k + 1)}{365^k}$$

Für das Ereignis A_k gilt dann

$$P(A_k) = 1 - \frac{365 \cdot \ldots \cdot (365 - k + 1)}{365^k} \quad \text{für} \quad k \leq 365$$

Geburtstagsproblem

In Abhängigkeit von der Anzahl k der ausgewählten Personen ergeben sich folgende Wahrscheinlichkeiten

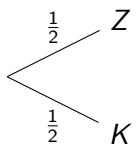
k	$P(A_k)$
5	0.027
10	0.117
20	0.411
25	0.569
30	0.706
40	0.891
60	0.994

Baumgraphen

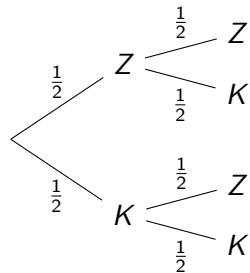
Insbesondere nacheinander unabhängig durchgeführte Zufallsexperimente lassen sich durch Baumgraphen darstellen

Wir betrachten das dreimal wiederholte Werfen einer Münze

Das erste Teilexperiment lässt sich folgendermaßen beschreiben

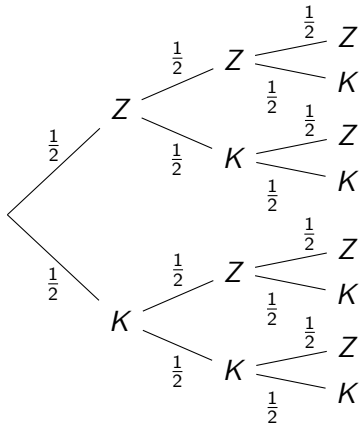


Zusätzliche Berücksichtigung des zweiten Wurfs ergibt



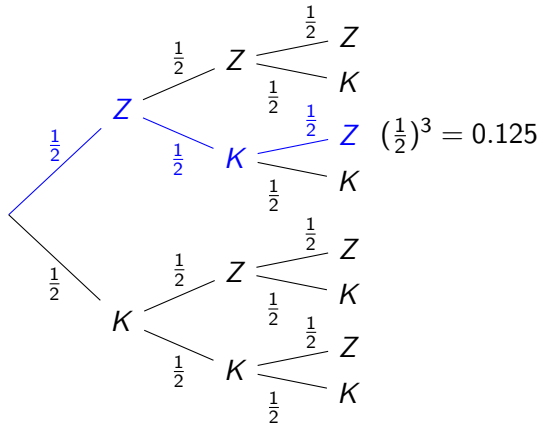
Dreimaliger Münzwurf

Berücksichtigung aller drei Teilexperimente liefert



Dreimaliger Münzwurf

Bestimmung der Wahrscheinlichkeit der Sequenz (Z, K, Z)



Dreimaliger Münzwurf

Bestimmung der Wahrscheinlichkeit der Sequenz (Z, K, Z)

- Die Wahrscheinlichkeit einer derartigen Folge von Ereignissen ergibt sich als Produkt der Einzelwahrscheinlichkeiten entlang der Pfade des Graphen, wenn die Einzelereignisse davon unabhängig sind, welche Ereignisse in den vorhergehenden Teilexperimenten eingetreten sind
- Der Begriff der (stochastischen) Unabhängigkeit von Ereignissen wird später weiter präzisiert

Bestimmung der Wahrscheinlichkeit mehrerer Sequenzen

- Die Wahrscheinlichkeit mehrerer Sequenzen ergibt sich als Summe der den einzelnen Sequenzen (d. h. den entsprechenden Pfaden) zugeordneten Wahrscheinlichkeiten (disjunkte Ereignisse)

Empirische Bestimmung von Wahrscheinlichkeiten

Nicht in jedem Fall kann die Wahrscheinlichkeit von Ereignissen theoretisch abgeleitet werden

Oft kann die Wahrscheinlichkeit $\pi(A)$ eines Ereignisses A anhand von Stichproben vom Umfang n (genaueres später) über die entsprechende relative Häufigkeit $f_n(A)$ bestimmt werden

- Voraussetzung dafür ist, dass die relative Häufigkeit aus unabhängigen Zufallsexperimenten bestimmt wird

Generell wird die relative Häufigkeit mit wachsendem Stichprobenumfang die Wahrscheinlichkeit zunehmend genauer beschreiben

Diese Beziehung wird durch das Bernoulli-Theorem beschrieben

Empirische Bestimmung von Wahrscheinlichkeiten

Bernoulli-Theorem (Gesetz der großen Zahlen)

$$P\left(|f_n(A) - \pi(A)| \geq \epsilon\right) \to 0 \quad \text{für} \quad n \to \infty$$

Diese Beziehung zeigt, dass die relative Häufigkeit $f_n(A)$ nicht zwangsläufig für wachsenden Stichprobenumfang n gegen die Wahrscheinlichkeit $\pi(A)$ konvergiert

Größere Abweichungen sind auch bei großem n stets möglich

Derartige Abweichungen werden bei wachsendem n aber zunehmend unwahrscheinlicher

Vollständige Ereignisdisjunktion

Partition bzw. Zerlegung von Ω

Gegeben sei eine Folge

$$B_1, B_2, \ldots, B_n$$

von Ereignissen mit folgenden Eigenschaften

- $B_i \cap B_j = \emptyset$ für alle $i \neq j$
- $\bigcup_{i=1}^{n} B_i = \Omega$

Jedes Ergebnis aus Ω liegt in genau einem der Ereignisse $B_1, \ldots, B_n$

Vollständige Ereignisdisjunktion

Partition bzw. Zerlegung von Ω

Im Allgemeinen kann die Ergebnismenge Ω hinsichtlich der Ausprägungen eines oder mehrerer Merkmale zerlegt werden

Seien die Folgen

$$A_1, \ldots, A_n \quad \text{und} \quad B_1, \ldots, B_m$$

Ereignisdisjunktionen von Ω

Beispiel: Körpergröße und Geschlecht

	B_1	B_2	
A_1			< 1.65
A_2			$1.65 - 1.80$
A_3			> 1.80
	m	w	

Verbundene Ereignisse und deren Wahrscheinlichkeit

Die Ereignisse

$$A_i \cap B_j$$

für alle $i = 1, \ldots, n$ und $j = 1, \ldots, m$ heißen *verbundene Ereignisse*

Die entsprechenden Wahrscheinlichkeiten

$$P(A_i \cap B_j)$$

für alle $i = 1, \ldots, n$ und $j = 1, \ldots, m$ werden als *verbundene Wahrscheinlichkeiten* bezeichnet

Satz der totalen Wahrscheinlichkeit

Wegen

$$A_i = (A_i \cap B_1) \cup \ldots \cup (A_i \cap B_m) \quad \text{und} \quad B_j = (B_j \cap A_1) \cup \ldots \cup (B_j \cap A_n)$$

folgt aus der Additivität der Wahrscheinlichkeit

$$P(A_i) = \sum_{j=1}^{m} P(A_i \cap B_j) \quad \text{und} \quad P(B_j) = \sum_{i=1}^{n} P(B_j \cap A_i)$$

Diese Formeln werden als *Satz der totalen Wahrscheinlichkeit* bezeichnet

- Die unbedingten Wahrscheinlichkeiten werden dabei in Beziehung gesetzt zu den aus einer Ereignisdisjunktion resultierenden verbundenen Wahrscheinlichkeiten

Verbundene Ereignisse und deren Wahrscheinlichkeit

Tabellendarstellung

	B_1	$\dots$	B_j	$\dots$	B_m	
A_1	$P(A_1 \cap B_1)$					$P(A_1)$
$\vdots$						$\vdots$
A_i			$P(A_i \cap B_j)$			$P(A_i)$
$\vdots$						$\vdots$
A_n					$P(A_n \cap B_m)$	$P(A_n)$
	$P(B_1)$	$\dots$	$P(B_j)$	$\dots$	$P(B_m)$	

Die Wahrscheinlichkeiten $P(A_1), \dots, P(A_n)$ und $P(B_1), \dots, P(B_m)$ werden als *Randwahrscheinlichkeiten* bezeichnet

Beispiel

Für das einmalige Werfen eines Würfels betrachten wir die Ereignisdisjunktionen

$$A_1 = \{1, 3, 5\} \quad \text{und} \quad A_2 = \{2, 4, 6\}$$

bzw.

$$B_1 = \{1, 2, 3\} \quad \text{und} \quad B_2 = \{4, 5, 6\}$$

Damit erhält man folgende Tabelle der verbundenen Ereignisse

	B_1	B_2
A_1	$\{1, 3\}$	$\{5\}$
A_2	$\{2\}$	$\{4, 6\}$

Als Tabelle der verbundenen Wahrscheinlichkeiten ergibt sich

	B_1	B_2	
A_1	$\frac{2}{6}$	$\frac{1}{6}$	$\frac{1}{2}$
A_2	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{1}{2}$
	$\frac{1}{2}$	$\frac{1}{2}$	