Willkommen zur Vorlesung Statistik (Master)

Thema dieser Vorlesung:
Wahrscheinlichkeit und Zufallsvorgänge

Prof. Dr. Wolfgang Ludwig-Mayerhofer

Universität Siegen – Philosophische Fakultät, Seminar für Sozialwissenschaften

Über Wahrscheinlichkeitstheorie

Here I should like to quote the words of Laplace, that the theory of probability is in fact but the good common sense which is reduced to formulae. It is able to express in exact terms what the sound minds feel by a sort of instinct, sometimes without being able to give good reasons for their beliefs.

Neyman, Jerzy: On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection, in: Journal of the Royal Statistical Society, Series A 97, 1934, S. 558-625, hier S. 563.

Zufallsvorgang (Zufallsexperiment) I

Beispiele:

- Werfen einer Münze – Kopf oder Zahl
- Würfeln – eine von sechs Zahlen
- Ziehen von Lottozahlen – eine Kombination aus vielen möglichen Kombinationen
- Zeugen von Kindern (noch) – Junge oder Mädchen

oder

- Ziehen einer Zufallsstichprobe – ein Stichprobenwert von vielen möglichen Stichprobenwerten (z. B. Mittel- oder Anteilswert)

Zufallsvorgang (Zufallsexperiment) II

Als Z. bezeichnet man eine Handlung oder Situation,

- die im Prinzip beliebig oft wiederholbar,
- deren Resultat genau eines von mehreren „Ereignissen“ (eines Ereignis-, Ergebnis- oder Merkmalsraumes), und
- deren Resultat vorher unbekannt ist.

Zufallsvorgang (Zufallsexperiment) III

Das Vorbild für Stichprobenziehungen: Lotterie (Ziehung aus Urne)

- Ziehung mit Zurücklegen: Unrealistisch, aber (teilweise) einfacher nachzuvollziehen.
- Ziehung ohne Zurücklegen: Realistisch, aber manchmal komplexer.
- Bei großem Stichprobenumfang n werden Unterschiede unbedeutend.

Die Laplace-Wahrscheinlichkeit

Den Ereignissen eines Zufallsvorganges können Wahrscheinlichkeiten P(A) zugeordnet werden. Beispiele:

- Münzwurf: Wahrscheinlichkeit für Kopf bzw. Zahl jeweils $^1/_2 = 0{,}5$.
- Würfel: Wahrscheinlichkeit für jede Augenzahl $= {}^1/_6$.

Allgemein:

$$P(A) = \frac{N_A}{N}$$

(Laplace-Wahrscheinlichkeit): Die Wahrscheinlichkeit von A ist Zahl der „günstigen“ Ereignisse, geteilt durch alle Ereignisse. Diese Regel gilt für den (typischen) Fall, dass alle Einzelereignisse gleich wahrscheinlich sind.

Wahrscheinlichkeit: Theoretisch und empirisch

Die „theoretischen" Wahrscheinlichkeiten werden auch als a-priori-Wahrscheinlichkeiten bezeichnet.

„Empirisch greifbar" sind sie als Anteilswert beim häufigen Wiederholen des Experiments. Eine Wahrscheinlichkeit $P(A) = 0{,}5$ bedeutet, dass wir „auf lange Sicht" in 50 Prozent der Fälle das entsprechende Ereignis erwarten.

Die Axiome von Kolmogorow

1. $0 \leqslant P(A) \leqslant 1$
   Wahrscheinlichkeiten haben Werte im Bereich von 0 bis 1.
2. $P(\Omega) = 1$
   Die Summe der Wahrscheinlichkeiten aller Ereignisse (des Ergebnisraums) ist 1.
3. $P(A \cup B) = P(A) + P(B)$, wenn A und B disjunkt
   Die Wahrscheinlichkeit, dass A oder B auftritt, ist gleich der Summe der Einzelwahrscheinlichkeiten (wenn A und B disjunkt sind, d. h. wenn sie nicht gleichzeitig auftreten können; formal: wenn $A \cap B = \emptyset$).

Aus Axiom 2 folgt eine wichtige Regel:

$$P(\bar{A}) = 1 - P(A)$$

Die Wahrscheinlichkeit von Nicht-A ist $1 - P(A)$.

Theoreme und Regeln für komplexe Ereignisse

1. Multiplikationsregel für unabhängige Ereignisse:

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

Beispiel 1: Wahrscheinlichkeit, dass bei zweimaligem Würfeln erst eine 1 und dann eine 4 gewürfelt wird:

$$\frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

Beispiel 2: Wahrscheinlichkeit, dass erst eine 1 oder 2, dann eine 4 und dann eine 6 gewürfelt wird:

$$\frac{2}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} = \frac{2}{216} = \frac{1}{108}$$

Theoreme und Regeln für komplexe Ereignisse

2. Additionstheorem:

Für disjunkte Ereignisse: Siehe 3. Axiom von Kolmogorow.

Beispiel: Wahrscheinlichkeit, dass bei einem Wurf entweder eine 1 oder eine 4 gewürfelt wird:

$$\frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

Für nicht disjunkte Ereignisse:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Beispiel: Wahrscheinlichkeit, dass bei einem Wurf eine gerade Zahl oder eine Zahl ≤ 3 gewürfelt wird:

$$\frac{1}{2} + \frac{1}{2} - \frac{1}{6} = \frac{5}{6}$$

Bedingte Wahrscheinlichkeit

Bedingte Wahrscheinlichkeiten nutzen vorhandene Informationen aus: Wie groß ist die Wahrscheinlichkeit für B, wenn A gegeben ist? A heißt das bedingende Ereignis.

Wenn die Wahrscheinlichkeit für B bei Vorliegen von A anders ist als bei Nicht-Vorliegen von A, so sind die Ereignisse A und B nicht unabhängig voneinander. Es gilt:

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \text{ (mit } P(A) > 0)$$

Daraus folgt die Multiplikationsregel für abhängige Ereignisse:

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

Stochastische Unabhängigkeit

Wir haben oben die Multiplikationsregel für unabhängige Ereignisse bereits kennengelernt. Man kann umgekehrt auch die gerade kennengelernte Multiplikationsregel für bedingte Wahrscheinlichkeiten heranziehen, um die Bedingung für (stochastische) Unabhängigkeit zu definieren:

Zwei Ereignisse sind unabhängig, wenn die Wahrscheinlichkeit für B nicht von A abhängt, wenn also gilt:

$$P(B|A) = P(B)$$

Das bedeutet aber:

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

woraus folgt:

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

Wahrscheinlichkeiten in/als Kreuztabellen

Wir beobachten über 150 Fußballspiele, ob Torwart A mitspielt oder nicht, und ob seine Mannschaft gewinnt (B) oder nicht.

	Spielt	Spielt nicht	n
Gewinnt	12	63	75
Gewinnt nicht	48	27	75
n	60	90	150

Die Anteilswerte, bezogen auf die Gesamtzahl der Fälle, entsprechen den gemeinsamen Wahrscheinlichkeiten $P(A \cap B)$ usw. bzw., bei der Randverteilung, den Wahrscheinlichkeiten $P(A)$, $P(\bar{A})$, $P(B)$ und $P(\bar{B})$.

	Spielt	Spielt nicht	n
Gewinnt	0,08	0,42	0,5
Gewinnt nicht	0,32	0,18	0,5
n	0,4	0,6	150

Bedingte Wahrscheinlichkeiten in Kreuztabellen

Die bedingten Wahrscheinlichkeiten $P(B|A)$ usw. lassen sich dann als Spalten-Anteilswerte ausdrücken:

	Spielt	Spielt nicht	n
Gewinnt	0,2	0,7	75
Gewinnt nicht	0,8	0,3	75
n	60	90	150

Die Zeilenprozente entsprechen dagegen den bedingten Wahrscheinlichkeiten $P(A|B)$ usw.

	Spielt	Spielt nicht	n
Gewinnt	0,16	0,84	75
Gewinnt nicht	0,64	0,36	75
n	60	90	150

Der Satz von Bayes: Die Fragestellung

Häufig interessieren uns „umgekehrte“ bedingte Wahrscheinlichkeiten.

Beispiel: Es sei bekannt, wie groß die Wahrscheinlichkeit ist, dass ein HIV-Test positiv ausfällt, wenn man tatsächlich HIV hat [$=P(B|A)$]. Wie ist es aber umgekehrt: Wenn jemand ein positives Ergebnis hat, wie groß ist die Wahrscheinlichkeit, dass er/sie tatsächlich infiziert ist [$=P(A|B)$]?

(Man beachte: In manchen Lehrbüchern, etwa bei Fahrmeir et al., finden sich komplexere Formeln, die mehr als zwei Zustände – also nicht nur A und nicht-A, sondern bspw. $A_1, A_2, \ldots, A_k$ – umfassen. Wir beschränken uns hier aber auf den binären Fall.)

Der Satz von Bayes: Gegebene Größen

Die Ausgangsgrößen sind:

$P(B|A) =$ Wahrscheinlichkeit, ein positives Ergebnis zu haben, falls man krank ist (Sensitivität)

$P(B|\bar{A}) =$ Wahrscheinlichkeit, ein positives Ergebnis zu haben, falls man nicht krank (=gesund) ist (Spezifität)

$P(A) =$ Wahrscheinlichkeit, krank zu sein (woraus sich die Wahrscheinlichkeit, gesund zu sein, als $1 - P(A)$ ergibt).

Gesucht ist, wie gesagt, $P(A|B)$.

Auf dem Weg zum Satz von Bayes I

Nach der Formel für bedingte Wahrscheinlichkeiten gilt:

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

Die Größe im Zähler kann man nach der Multiplikationsregel bestimmen:

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

Im Beispiel: Die Wahrscheinlichkeit, krank zu sein und ein positives Ergebnis zu haben, ist gleich der Wahrscheinlichkeit, krank zu sein, mal der Sensitivität des Tests.

Auf dem Weg zum Satz von Bayes II

Die Größe $P(B)$ – die Gesamtwahrscheinlichkeit, positive Ergebnisse zu haben – setzt sich zusammen aus der eben berechneten Wahrscheinlichkeit eines positiven Ergebnisses bei Krankheit und der Wahrscheinlichkeit eines positiven Ergebnisses bei Gesundheit, also der Wahrscheinlichkeit, gesund zu sein, mal der Spezifität des Tests; formal:

$$P(B) = P(A \cap B) + P(\bar{A} \cap B)$$

Für den rechten Term gilt dann nach der Multiplikationsregel:

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

Der Satz von Bayes (binärer Fall)

Damit ergibt sich:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|\bar{A}) \cdot P(\bar{A})}$$

Sieht kompliziert aus, lässt sich aber leicht anhand eines Wahrscheinlichkeitsbaums oder einer Kreuztabelle nachvollziehen.

Permutationen

Wie viele Möglichkeiten gibt es, n Elemente anzuordnen? Beispiel: Die Zahlen 1 bis 4

Zug 1:		1			2			3			4	
Zug 2:	2	3	4	1	3	4	1	2	4	1	2	3
Zug 3:	3 4	2 4	2 3		usw.			usw.			usw.	
Zug 4:	4 3	4 2	3 2									

Also: Im ersten Schritt 4 Möglichkeiten, im zweiten 3, im dritten 2 und zuletzt nur noch 1:

$$n \cdot (n-1) \cdot \ldots \cdot (n-[n-1]) = n!$$

n! wird gesprochen: n Fakultät.

Für später: Es gilt per definitionem: $0! = 1$

Variationen I

Angenommen, wir ziehen eine Stichprobe von k Elementen aus einer Gesamtheit von n Elementen? Wie viele Möglichkeiten gibt es? Beispiel: Ziehen von 2 aus 4 Elementen.

Fall 1: Mit Zurücklegen

Zug 1:		1				2				3				4		
Zug 2:	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4

Zahl der Variationen mit Zurücklegen: n^k

Variationen II

Angenommen, wir ziehen eine Stichprobe von k Elementen aus einer Gesamtheit von n Elementen? Wie viele Möglichkeiten gibt es? Beispiel: Ziehen von 2 aus 4 Elementen.

Fall 2: Ohne Zurücklegen

Zug 1:		1			2			3			4	
Zug 2:	2	3	4	1	3	4	1	2	4	1	2	3

Zahl der Variationen ohne Zurücklegen:

$$n \cdot (n-1) \cdot \ldots \cdot (n-[k-1]) = \frac{n!}{(n-k)!}$$

Kombinationen

Die Reihenfolge der Ziehung der Elemente ist (für unsere Zwecke) irrelevant. Bspw. sind die Stichproben (1, 2) und (2, 1) als gleichwertig zu betrachten, ebenso (1, 3) und (3, 1) usw.

Das gleiche gilt für Stichproben mit Zurücklegen, wo auch $(1_1, 1_2)$ und $(1_2, 1_1)$ gleichwertig sind; da der Fall mit Zurücklegen für die Sozialwissenschaften irrelevant ist, übergehen wir ihn im Folgenden.

Beispiel: 3 Elemente aus 4

$$\begin{array}{ccccccl}
1,2,3 & 1,3,2 & 2,1,3 & 2,3,1 & 3,1,2 & 3,2,1 & \rightarrow (1\ \ 2\ \ 3) \\
1,2,4 & 1,4,2 & 2,1,4 & 2,4,1 & 4,1,2 & 4,2,1 & \rightarrow (1\ \ 2\ \ 4) \\
1,3,4 & 1,4,3 & 3,1,4 & 3,4,1 & 4,1,3 & 4,3,1 & \rightarrow (1\ \ 3\ \ 4) \\
2,3,4 & 2,4,3 & 3,2,4 & 3,4,2 & 4,2,3 & 4,3,2 & \rightarrow (2\ \ 3\ \ 4)
\end{array}$$

Kombinationen: Der Binomialkoeffizient

Es gibt also $\frac{n!}{(n-k)!}$ Variationen, aber jeweils k! Permutationen werden als identisch aufgefasst. Die Zahl der Kombinationen ist also:

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

Dies ist der sog. Binomialkoeffizient, gesprochen „n über k". Das bekannteste Beispiel ist die Lottoziehung (z. B. 6 Richtige im Lotto).