Wahrscheinlichkeitsrechnung für Sozialwissenschaftler ?

1. Probabilistisches Denken (d.h. das Denken in Wahrscheinlichkeiten) unerlässlich! Strenge Kausalitäten (wenn A dann folgt immer B) findet man bestenfalls vereinzelt in Naturwissenschaften, in den Sozialwissenschaften gilt typischerweise nur: wenn A dann folgt eher B als C.

2. Wahrscheinlichkeiten und Umgang mit Unsicherheit spielen in der Gesellschaft eine wichtige Rolle. Bei naiver Herangehensweise (ohne Wahrscheinlichkeitsrechnung) kann man sich leicht täuschen. Z.B. bewerten sowohl medizinische Experten als auch Laien Risiken oft falsch.

3. Stichprobenverfahren und statistische Modelle spielen in den (empirisch orientierten) Sozialwissenschaften eine zentrale Rolle. Für das Verständnis sind Grundlagenkenntnisse in Wahrscheinlichkeitsrechnung zentral

Literatur

- Götz Rohwer, Ulrich Pötter: Wahrscheinlichkeit: Begriff und Rhetorik in der Sozialforschung, Weinheim (u.a.): Juventa-Verlag 2002
- Gigerenzer, G.: Das Einmaleins der Skepsis. BTB, 2.Aufl 2005.
- Nate Silver: The Signal and the Noise: Why So Many Predictions Fail but Some Don‘t. Penguin Press 2012.

Zufallsvorgänge

Ein Zufallsvorgang (Zufallsexperiment) führt zu einem von mehreren, sich gegenseitig ausschließenden Ergebnissen. Es ist vor der Durchführung ungewiss, welches Ergebnis eintreten wird.
Was benötigen wir zur Beschreibung eines Zufallsvorganges?

Zwei wesentliche Aspekte:

a) Welche Ergebnisse eines Zufallsvorgangs sind möglich? (Was kann alles passieren?)

b) Mit welcher Wahrscheinlichkeit treten die einzelnen Ergebnisse ein?

Formale Beschreibung

Festlegen eines *Ergebnisraums* (Grundraum, Stichprobenraum) Ω, der alle möglichen *Ergebnisse* ω enthält.

Beispiele:

- $\Omega = \{1, \ldots, 6\}$ beschreibt die möglichen Ergebnisse eines Würfelexperiments

  Ein mögliches Ergebnis: $\omega = 4$; $\omega = 17$ ist kein mögliches Ergebnis.

- $\Omega = \mathbb{R}_0^+$ beschreibt die möglichen Erwerbseinkommen

  Ein mögliches Ergebnis: $\omega = 17513$ €

- Ziehung einer Person: $\Omega = \{1, \ldots, N\}$

  Ein mögliches Ergebnis: $\omega = 17$

Ereignisse

Ereignisse sind **Teilmengen** von Ω

Beispiele:

1. „gerade Zahl“ $= \{2, 4, 6\}$
2. „1 oder 2“ $= \{1, 2\}$
3. „Einkommen zwischen 1000 und 2000 h €“ $= \{\omega | 1000 \leq \omega \leq 2000\}$
4. „Person ist weiblich“ $=$ {alle Nummern, die zu Frauen gehören}

Ereignissen sollen Wahrscheinlichkeiten zugeordnet werden.
Wir bezeichnen Ereignisse mit A,B,C,...

Grundlagen: Mengen und Mengenoperationen

Definition

Eine Menge ist eine Zusammenfassung verschiedener Objekte zu einem Ganzen. Die einzelnen Objekte einer Menge werden Elemente genannt.

Mengen werden üblicherweise mit Großbuchstaben bezeichnet, z.B. A, B, C, Ω, ...

Mengen werden benutzt, um den Ausgang von Zufallsexperimenten zu beschreiben.

Beispiele für Mengen

- $\Omega = \{\text{Hörer einer Vorlesung}\}$.
- $\Omega = \{1, 2, 3, 4, 5, 6\}$ (Menge der Ergebnisse eines Würfelwurfs).

  Die Reihenfolge der Aufzählung spielt (im Gegensatz zu Tupeln) keine Rolle:

  $$\{1, 2, 3, 4, 5, 6\} \;=\; \{1, 3, 5, 2, 4, 6\}$$

- $B = \{K, Z\}$ (Menge der Ergebnisse eines Münzwurfs, K=Kopf, Z=Zahl).
- Charakterisierung von Mengen mit einer gewissen Eigenschaft:

  $$\{1, 2, 3, \ldots, 10\} = \{x \,|\, x \text{ ist eine natürliche Zahl } \leq 10\}$$

  Die Menge aller x mit der Eigenschaft „x ist eine natürliche Zahl ≤ 10“.

Beispiele: Standardmengen

$\mathbb{N} = \{1, 2, 3, \ldots\}$: Menge der natürlichen Zahlen,

$\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\}$: Menge der natürlichen Zahlen inklusive 0,

$\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$: Menge der ganzen Zahlen,

$\mathbb{R} = (-\infty, \infty)$: Menge der reellen Zahlen,

$\emptyset$: leere Menge.

Grundlegende Begriffe der Mengenlehre

Illustration anhand folgender Mengen

$$
\begin{aligned}
\Omega &= \{\text{CDU/CSU, SPD, FDP, Grüne, Linke, Sonstige}\} \\
A &= \{\text{CDU/CSU, SPD, FDP, Grüne}\} \\
B &= \{\text{CDU/CSU, SPD, FDP}\} \\
C &= \{\text{SPD, FDP, Grüne}\}
\end{aligned}
$$

- **Elementeigenschaft:**

  x ist Element der Menge A: $x \in A$

  x ist nicht Element der Menge A: $x \notin A$

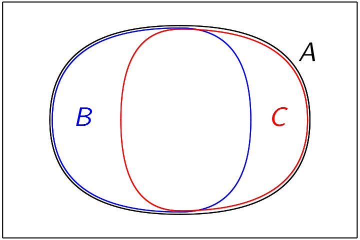
  Es gilt SPD $\in C$, aber CDU/CSU $\notin C$.

- **Teilmengen:** A ist Teilmenge von B, in Zeichen $A \subset B$, wenn jedes Element von A auch in B ist.

$B \subset A$

C ist keine Teilmenge von B

Veranschaulichung von Mengen im Venn-Diagramm



$\Omega \leftarrow$ größte Menge als Kasten

Im Beispiel:
$A \subset \Omega$, $B \subset \Omega$, $C \subset \Omega$, $B \subset A$, $C \subset A$,
aber beispielsweise: weder $B \subset C$ noch $C \subset B$ noch $C = B$

Einfache Eigenschaften von Teilmengen

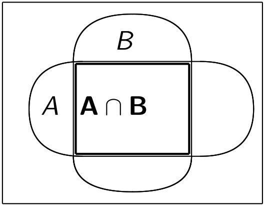
Für jede Menge A gilt:

$\emptyset \subset A$	Die leere Menge ist Teilmenge jeder Menge, denn jedes Element in $\emptyset$ ist in A
$A \subset A$	d.h. „ $\subset$ “ enthält implizit „ $=$ “, deshalb in Literatur manchmal auch $\subseteq$ statt $\subset$

Schnittmenge

Die Schnittmenge $A \cap B$ ist die Menge aller Elemente, die sowohl in A als auch in B enthalten sind:

$$A \cap B = \{x | x \in A \text{ und } x \in B\}$$



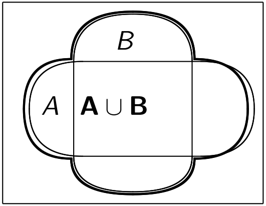
Eigenschaften

* Gilt $A \subset B$, so ist $A \cap B = A$.
* Für jede Menge A gilt: $A \cap A = A$ und $A \cap \emptyset = \emptyset$.
* Zwei Mengen A und B mit $A \cap B = \emptyset$, d.h. zwei Mengen, die kein gemeinsames Element haben, heißen *disjunkt*.
* Die Schnittmenge aus n Mengen $A_1, \ldots, A_n$ enthält alle Elemente, die in jeder der Mengen $A_1, \ldots, A_n$ enthalten sind und wird bezeichnet mit
$$\bigcap_{i=1}^{n} A_i := A_1 \cap A_2 \cap \ldots \cap A_n.$$

Vereinigungsmenge: Die Vereinigungsmenge $A \cup B$ ist die Menge aller Elemente, die in A oder B enthalten sind:

$$A \cup B = \{x | x \in A \text{ oder } x \in B\}$$

Weitere Eigenschaften



* Vorsicht: Das „oder“ ist *nicht* exklusiv gemeint, also nicht „entweder oder“, sondern als „in A oder in B oder in beiden“.
* Die Vereinigungsmenge aus n Mengen $A_1, \ldots, A_n$ enthält alle Elemente, die in mindestens einer der Mengen $A_1, \ldots, A_n$ enthalten sind und wird bezeichnet mit

$$\bigcup_{i=1}^{n} A_i := A_1 \cup A_2 \cup \ldots \cup A_n$$

Ereignisoperationen

$A \cup B$:	Vereinigung	= „A oder B“
$A \cap B$:	Durchschnitt	= „A und B“
A^C:	Komplement	= „Nicht A“

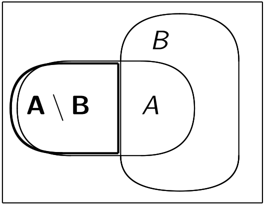
Beispiele:

Ω	$= \{1,2,3,4,5,6\}$	
A	$= \{2,4,6\}$	„gerade“
B	$= \{4,5,6\}$	„groß“
$A \cup B$	$= \{2,4,5,6\}$	„gerade oder groß“
$A \cap B$	$= \{4,6\}$	„gerade und groß“
A^C	$= \{1,3,5\}$	„ungerade“
B^C	$= \{1,2,3\}$	„klein“

Differenzmenge

Die Differenzmenge $A \setminus B$ ist die Menge aller Elemente, die in A, aber nicht in B enthalten sind:

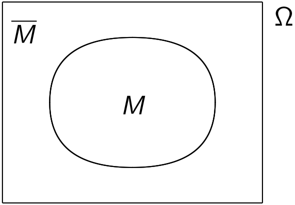
$$A \setminus B = \{x | x \in A \text{ aber } x \notin B\}$$



Komplementärmenge

Die Komplementärmenge $\overline{A}$ bezüglich einer Grundmenge Ω ist die Menge aller Elemente von Ω, die nicht in A sind:

$$\overline{A} = \{x \in \Omega | x \notin A\} = \{x : x \notin A\}$$



Bemerkung

* Die Komplementärmenge ist nur unter Bezugnahme auf eine Grundmenge Ω definierbar.
* Es gilt $\overline{A} = \Omega \setminus A$.
* Es existieren noch weitere Schreibweisen für die Komplementärmenge,
  z.B. A^C, $\mathcal{C}A$.
* „Tertium non datur“ (Grundlegendes Prinzip der Mengenlehre (und der Logik): Für jedes Element $x \in \Omega$ gilt entweder $x \in A$ oder $x \in \overline{A}$

Potenzmenge: Die Potenzmenge $\mathcal{P}(A)$ ist die Menge aller Teilmengen von A:

$$\mathcal{P}(A) = \{M | M \subset A\}.$$

Beispiel

Im Beispiel:

$$\begin{aligned} \mathcal{P}(B) &= \{\emptyset, \{\text{CDU/CSU}\}, \{\text{SPD}\}, \{\text{FDP}\}, \\ &\quad \{\text{CDU/CSU, SPD}\}, \{\text{CDU/CSU,FDP}\}, \{\text{SPD, FDP}\}, B\} \end{aligned}$$

Mächtigkeit: Die Mächtigkeit $|A|$ einer Menge A ist die Anzahl der Elemente von A
Im Beispiel:

$$|B| = 3$$

Für jede Menge B gilt $|P(B)| = 2^{|B|}$; im Beispiel 2^3

Rechenregeln für Mengen

1. Kommutativgesetze (Vertauschung):

$$A \cap B = B \cap A, \quad A \cup B = B \cup A.$$

2. Assoziativgesetze (Zusammenfassen):

$$(A \cap B) \cap C = A \cap (B \cap C).$$
$$(A \cup B) \cup C = A \cup (B \cup C).$$

Rechenregeln für Mengen

1. Distributivgesetze (Ausklammern/Ausmultiplizieren):

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C).$$
$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C).$$

2. De Morgansche Regeln:

$$(\overline{A \cup B}) = \overline{A} \cap \overline{B}$$
$$(\overline{A \cap B}) = \overline{A} \cup \overline{B}$$

3. Aus $A \subset B$ folgt $\overline{B} \subset \overline{A}$.
4. Für die Differenzmenge gilt $A \setminus B = A \cap \overline{B}$.
5. Für die Potenzmenge gilt $|\mathcal{P}(A)| = 2^{|A|}$.

   vgl. Bsp: $|B| = 3 \quad |\mathcal{P}(B)| = 2^3 = 8$

Das kartesische Produkt

Das kartesische Produkt zweier Mengen

$$A = \{a_1, a_2, a_3, \ldots, a_k\}$$
$$B = \{b_1, b_2, b_3, \ldots, b_m\}$$

ist die Menge

$$A \times B := \Big\{(a_i, b_j) \;\Big|\; i = 1, \ldots, k, j = 1, \ldots, m\Big\}$$

Sie besteht also aus allen möglichen Kombinationen, so dass

$$\begin{aligned} A \times B \quad = \quad & \{(a_1, b_1), (a_1, b_2), (a_1, b_3), \ldots, (a_1, b_m), \\ & (a_2, b_1), (a_2, b_2), (a_2, b_3), \ldots, (a_2, b_m), \\ & \quad\quad \vdots \\ & (a_k, b_1), (a_k, b_2), (a_k, b_3), \ldots, (a_k, b_m)\} \end{aligned}$$

Kartesisches Produkt: Beispiel

$$
\begin{aligned}
A &= \{1, 2, 3\} \\
B &= \{1, 2, 3, 4\} \\
A \times B &= \big\{(1, 1), (1, 2), (1, 3), (2, 1), \ldots\big\}
\end{aligned}
$$

Achtung: Bei den Elementen von $A \times B$ handelt es sich um Tupel, das heißt die Reihenfolge ist wichtig! (z.B. $(1, 2)$ ist etwas anderes als $(2, 1)$, vgl. Statistik I Kapitel 5)

Verallgemeinerungen:

- Das kartesische Produkt der Mengen $\Omega_1, \Omega_2, \ldots, \Omega_n$ wird mit

$$\times_{i=1}^{n} \Omega_i = \Omega_1 \times \Omega_2 \times \ldots \times \Omega_n$$

  bezeichnet und besteht aus allen möglichen n-Tupeln, die sich (unter Beachtung der Reihenfolge) aus Elementen aus $\Omega_1, \Omega_2, \ldots, \Omega_n$ bilden lassen.
- Die Mengen $\Omega_1, \Omega_2, \ldots, \Omega_n$ müssen nicht endlich sein; für endliche Mengen gilt

$$|\times_{i=1}^{n} \Omega| = |\Omega_1| \cdot |\Omega_2| \cdot \ldots \cdot |\Omega_n|$$

- Kartesische Produkte werden verwendet, um Ergebnisse komplexer Experimente aus Einzelexperimenten zusammenzusetzen.

Wahrscheinlichkeit (formale Definition)

Definition

Eine Wahrscheinlichkeitsfunktion ordnet jedem Ereignis seine Wahrscheinlichkeit zu. Eine Wahrscheinlichkeit ist also eine Abbildung von Ereignissen (Elementen der Potenzmenge von Ω) auf reelle Zahlen:

$$\begin{aligned} P : \mathcal{P}(\Omega) &\rightarrow \mathbb{R} \\ A &\mapsto P(A) \end{aligned}$$

Dabei sollen gewisse fundamentale Rechenregeln gelten, z.B.

- 108 kann keine Wahrscheinlichkeit sein, nur Zahlen zwischen 0 und 1.
- $P(\{2,3\})$ muss mindestens so groß sein wie $P(\{3\})$.

Elementarereignisse

Die einelementigen Teilmengen (also die Ereignisse, die ein Ergebnis in ω enthalten) werden als *Elementarereignisse* bezeichnet.
Beispiel: Bei einem fairen Würfel gilt für die Elementarereignisse

$$\begin{aligned}
P(\text{„Augenzahl 1“}) &= P(\{1\}) = \tfrac{1}{6} \\
P(\text{„Augenzahl 2“}) &= P(\{2\}) = \tfrac{1}{6} \\
&\vdots \\
P(\text{„Augenzahl 6“}) &= P(\{6\}) = \tfrac{1}{6}
\end{aligned}$$

Laplace-Wahrscheinlichkeiten

Laplace-Wahrscheinlichkeiten oder kombinatorische Wahrscheinlichkeiten

Häufig: Alle möglichen Elementarereignisse sind *gleich* wahrscheinlich, d.h. $P(\{w_j\}) = \frac{1}{|\Omega|}$. In diesem Fall sprechen wir von einem *Laplace-Experiment*.

Abzählregel:

$$P(A) = \frac{\text{Anzahl der für } A \text{ günstigen Ergebnisse}}{\text{Anzahl aller möglichen Ergebnisse}}$$

Laplace-Wahrscheinlichkeit: In einem Laplace-Experiment gilt für $P(A)$ mit $|A| = M$ und $|\Omega| = N$:

$$P(A) = \frac{|A|}{|\Omega|} = \frac{M}{N}.$$

Beispiel: Dreimaliger Münzwurf

Wir werfen dreimal unabhängig voneinander eine faire Münze und notieren jeweils, ob die Münze Wappen oder Zahl anzeigt. Man beschreibt den Ergebnisraum und berechne die Wahrscheinlichkeit, mindestens einmal Wappen zu erhalten und genau zweimal Wappen zu erhalten.

Beispiel: Dreimaliger Münzwurf

Der zugehörige Ergebnisraum lautet:

$$\Omega = \{(W,W,W),(W,W,Z),(W,Z,W),(Z,W,W),\\ (W,Z,Z),(Z,W,Z),(Z,Z,W),(Z,Z,Z)\}$$

Ω enthält die acht Ergebnisse dieses Zufallsexperiments, d.h. $|\Omega| = 8$. Da wir vorausgesetzt haben, dass die Münze fair ist, kann dieser Zufallsvorgang als Laplace-Experiment angesehen werden, und es gilt:

$$p(\{\omega\}) = \frac{1}{|\Omega|} = \frac{1}{8}, \quad \omega \in \Omega.$$

Beispiel: Dreimaliger Münzwurf

Berechnung der Wahrscheinlichkeit

Damit besitzt jedes mögliche Ergebnis die gleiche Wahrscheinlichkeit $1/8$, d.h. dreimal hintereinander Wappen zu erhalten, ist genauso wahrscheinlich wie etwa die Kombination (W, Z, W).

Die Wahrscheinlichkeit für das Ereignis $A =$ „mindestens einmal Wappen“ berechnet sich nun direkt als Quotient der Anzahl der für A günstigen Ergebnisse und der Anzahl aller möglichen Ergebnisse. Da bei sieben Elementarereignissen von Ω mindestens einmal Wappen auftritt, ist $|A| = 7$ und somit

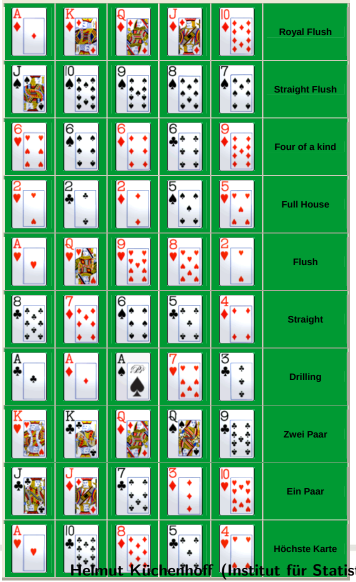
$$P(A) = \frac{7}{8}.$$

Für das Ereignis $B = \{\text{genau zweimal Wappen}\}$ gilt:

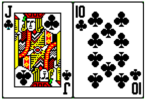
$$B = \{(W, W, Z), (W, Z, W), (Z, W, W)\}$$

also $|B| = 3$ und damit $P(B) = \frac{3}{8}$.

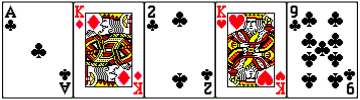
Beispiel: Poker



Poker: Texas Holdem



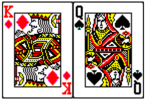
Deine Hand



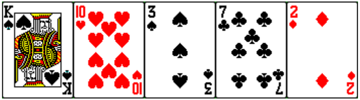
Board

In diesem Beispiel hat man ein Flush, denn man kann die drei Kreuzkarten in der Mitte zusammen mit den zwei Kreuzkarten auf der Hand verwenden.

Beispiel 2



Deine Hand



Board

In diesem Beispiel hat man ein Paar Könige mit einer Dame als Kicker.

Beispiel: Poker

Berechnung von Wahrscheinlichkeiten

Meine Hand: 2 Asse
Ein Gegner: Alle Möglichkeiten für 2 Karten des Gegners:
Alle Möglichkeiten für die 5 gemeinsmen Karten.
Berechnung mit PokerStove (frei im Internet)
2,097,572,400 Möglichkeiten
Günstige Fälle 1781508418
$P(Gewinn) = \frac{1781508418}{2097572400} = 85.204\%$

Ziehen aus einer Grundgesamtheit

Beispiel: Es wird ein Studierender der Vorlesung gezogen und nach seiner Wahlabsicht gefragt.

Dazu nehmen wir an, dass es N Studierende in der Vorlesung gibt und dass sie durchnummeriert sind $n = 1, ..., N$

$$P(\text{Student Nr n wird gezogen}) = 1/N$$

Alle haben die gleiche Ziehungswahrscheinlichkeit.

- Wie groß ist die Wahrscheinlichkeit, dass er/sie ein SPD Wähler ist?
- Wie groß ist die Wahrscheinlichkeit, dass eine Frau gezogen wird?

Lösung nach Laplace

- Wahrscheinlichkeit für „SPD-Wähler"

$$
\begin{aligned}
P(SPD) &= \frac{\text{Anzahl der für } SPD \text{ günstigen Ergebnisse}}{\text{Anzahl aller möglichen Ergebnisse}} \\
&= \frac{\text{Anzahl der SPD Wähler}}{\text{Anzahl aller Studierenden der Vorlesung}}
\end{aligned}
$$

  Die Wahrscheinlichkeit ist also die relative Häufigkeit f_{SPD} der SPD Wähler in der Grundgesamtheit.

- Wahrscheinlichkeit für „Frau"?

Relative Häufigkeiten und Wahrscheinlichkeiten

Die Argumentation des Beispiels gilt ganz allgemein.

$$P(\text{Eine Person mit der Eigenschaft E wird gezogen }) = f_E$$

- Die relativen Häufigkeiten/Anteile aus der Grundgesamtheit pflanzen sich also in der entsprechenden Wahrscheinlichkeitsverteilung in der Stichprobe fort.
- Dies ist ganz entscheidend, denn dadurch kann man also durch eine Stichprobe etwas über die Häufigkeitsverhältnisse in der Grundgesamtheit lernen.
- Wichtig ist dabei, dass die Stichprobe zufällig gezogen wird!

Urnenmodelle

Es hat sich eingebürgert, gewisse Grundsituationen, die in der praktischen Stichprobenziehung immer wieder vorkommen, als „Urnenmodelle“ zu beschreiben. Man stellt sich eine Urne mit Kugeln vor und zieht daraus dann in einer bestimmten Art und Weise eine bestimmte Anzahl von Kugeln (*Stichprobe*). In der Sozialwissenschaft entspricht jede „Kugel“ einer interessierenden Einheit (Person, Haushalt) und die „Urne“ einer gedachten Gesamtliste dieser Einheiten. Eine typische Unterscheidung der verschiedenen Ziehungsvorgänge besteht darin, ob eine Einheit mehrfach in eine Stichprobe gelangen kann (*„Ziehen mit Zurücklegen“*, die gezogene Kugel wird also wieder in die Urne zurückgelegt) oder *„Ziehen ohne Zurücklegen“* (die gezogenen Kugeln bleiben also außerhalb der Urne).

Praktische Aspekte

1. das Ziehen von n Kugeln ohne Zurücklegen entspricht einem n-fachen einmaligen Zug aus der Urne.
2. Typischerweise sind Stichproben ohne Zurücklegen praktisch einfacher zu realisieren und zu rechtfertigen.
3. Für sehr große Grundgesamtheiten sind die Unterschiede zwischen mit und ohne Zurücklegen verschwindend gering.

Die praktische Umsetzung erfolgt mit Hilfe von Computerprogrammen (Pseudozufallszahlen)

Ziehen mit Zurücklegen

- Grundgesamtheit mit N Zahlen $G = \{1, \ldots, N\}$.
- Ziehe Stichprobe vom Umfang n **mit** Zurücklegen.
- Zur Beschreibung des Zufallsvorgangs müssen wir die Anzahl der potentiell möglichen Stichprobenergebnisse bestimmen (jede Stichprobe ist gleichwahrscheinlich).
- $\Omega = \{(\omega_1, \ldots, \omega_n) | \omega_j \in \{1, \ldots, N\}\}$, das selbe Element kann mehrfach vorkommen.
- $|\Omega| = \underbrace{N \cdot N \cdot \ldots \cdot N}_{n\text{-mal}} = N^n$, d.h. N^n potentiell mögliche Stichproben vom Umfang n.

Ziehen ohne Zurücklegen

- Grundgesamtheit mit N Einheiten(mit Nummern identifiziert) $G = \{1, \ldots, N\}$.
- Ziehe Stichprobe vom Umfang n **ohne** Zurücklegen.
- $\Omega = \{(\omega_1, \ldots, \omega_n) : \omega_j \in \{1, \ldots, N\}, \omega_j \neq \omega_i \text{ für } i \neq j\}$, jedes Element kann nur einmal vorkommen.
- Anzahl möglicher Stichproben:

$$|\Omega| = \underset{\substack{\uparrow \\ \text{1. Ziehung}}}{N} \cdot \underset{\substack{\uparrow \\ \text{2. Ziehung}}}{(N-1)} \cdot \ldots \cdot \underset{\substack{\uparrow \\ n\text{-te Ziehung}}}{N-n+1} =$$

$$= \frac{N \cdot (N-1) \cdot \ldots \cdot (N-n+1)(N-n) \cdot \ldots \cdot 2 \cdot 1}{(N-n)(N-n-1) \cdot \ldots \cdot 2 \cdot 1} =$$

$$\frac{N!}{(N-n)!}$$

Einfache Zufallsstichprobe

Ziehen **ohne** Zurücklegen **ohne** Berücksichtigung der Reihenfolge

- Ziehe n Kugeln aus einer Urne mit N nummerierten Kugeln. Die Reihenfolge der Ziehungen spielt keine Rolle, d.h. die Stichprobe „4,1,7“ wird nicht unterschieden von „7,1,4“.
- $\Omega = \{\{\omega_1, \ldots, \omega_n\} : \omega_j \in \{1, \ldots, N\}, \omega_j \neq \omega_i$ für $j \neq i\}$
- Anzahl der Stichproben:

$$|\Omega| = \frac{N!}{(N-n)!n!} = \binom{N}{n}$$

Herleitung

Man berücksichtigt zunächst die Reihenfolge. Dann gibt es $N!$ Möglichkeiten, die Kugeln anzuordnen. Da nur die ersten n Kugeln zählen, sind jeweils $(N-n)!$ Anordnungen, die sich nur ab der $(n+1)$-ten Kugel unterscheiden, äquivalent, also hat man

$$\frac{N!}{(N-n)!}$$

Möglichkeiten, geordnete Stichproben zu erzeugen.
Da die Reihenfolge unter den ersten n Kugeln aber keine Rolle spielen soll, sind jeweils $n!$ geordnete Stichproben wiederum äquivalent, es gibt also insgesamt

$$\frac{N!}{(N-n)!n!}$$

verschiedene Stichproben.

Binomialkoeffizient

Der *Binomialkoeffizient* $\binom{N}{n}$ ist definiert als

$$\binom{N}{n} = \frac{N!}{(N-n)! \cdot n!}\,.$$

Es gilt:

$$\binom{N}{0} = 1, \binom{N}{1} = N, \binom{N}{N} = 1\,, \quad \binom{N}{n} = 0, \text{ falls } N < n\,.$$

Beispiel: Lottozahlen

- Frage: Wie viele verschiedene mögliche Ziehungen gibt es?
- Antwort: Entspricht der Ziehung ohne Zurücklegen von 6 Zahlen aus 49, d.h.

$$|\Omega| = \binom{49}{6} = \frac{49!}{43! \cdot 6!} = 13983816.$$

Dann gilt

$$P(\text{„6 Richtige“}) = \frac{1}{13983816} = 0.000000072$$

Beispiel: Lottozahlen, 5 richtige

- Frage: Wie groß ist die Wahrscheinlichkeit, 5 Richtige zu bekommen?
- Anzahl der günstigen Fälle:

$$\binom{6}{5} \cdot \binom{43}{1}$$

$$\uparrow \qquad\qquad \uparrow$$

Anzahl der Möglichkeiten: 5 Richtige aus 6 Zahlen

Anzahl der Möglichkeiten: 1 Falsche aus 43 Zahlen

$$P(\text{„5 Richtige“}) = \frac{\binom{6}{5}\binom{43}{1}}{\binom{49}{6}} = \frac{6 \cdot 43}{13983816} = 0.0000184.$$

Axiomensystem von Kolmogoroff (1933)

Warum reichen Laplace-Wahrscheinlichkeiten nicht?

Essentielle Voraussetzung: alle Fälle müssen „gleich möglich“(also gleich wahrscheinlich) sein

Beispiel: Wie wird das Wetter morgen? 3 Möglichkeiten:

$$\{\text{Sonne,Regen,Gemischt}\} \Longrightarrow P(\text{„Sonne“}) = \frac{1}{3}$$

Axiome von Kolmogorov

Eine Funktion P (P steht für Probability), die Ereignissen aus Ω reelle Zahlen zuordnet, heißt *Wahrscheinlichkeit*, wenn gilt

(K1) $P(A) \geq 0$ für alle Ereignisse $A \subset \Omega$.

(K2) $P(\Omega) = 1$.

(K3) Falls $A \cap B = \emptyset$, dann gilt $P(A \cup B) = P(A) + P(B)$

Axiome von Kolmogorov (1933)

- Die Axiome von Kolmogorov stellen zunächst eine reine Definition dar, die festlegt, was eine Wahrscheinlichkeit sein soll.
- Es gibt verschiedene Versuche Wahrscheinlichkeiten operational zu definieren (also durch eine Messvorschrift) und verschiedene Interpretationen, die die Axiomatik mit Leben füllen sollen.
- Aus hier nicht zu erörternden mathematischen Gründen
  - darf man bei überabzählbar unendlichen Ergebnisräumen, z.B. also im Fall $\Omega = \mathbb{R}$, nicht alle Teilmengen von Ω als Ereignisse zulassen. Alle Mengen, „an die man normalerweise denkt“ sind aber zugelassen.
  - muss man bei unendlichen Ergebnisräumen in $(K3)$ eigentlich unendliche Summen zulassen.

Rechenregeln für Wahrscheinlichkeiten

- $P(\overline{A}) = 1 - P(A)$
- Für nicht notwendigerweise disjunkte Mengen A, B gilt

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

- Falls $A_1, A_2, \ldots, A_n$ paarweise disjunkt sind, also $A_i \cap A_j = \emptyset$ für $i \neq j$, dann gilt:

$$P\left(\bigcup_{i=1}^{n} A_i\right) = P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n)$$

Es folgt, dass, sofern Ω endlich ist, die Wahrscheinlichkeit durch die Wahrscheinlichkeit der Elementarereignisse vollständig bestimmt ist:

$$P(A) = \sum_{\omega \in A} P(\{\omega\})$$

Würfelwurf mit fairem Würfel

Ergebnisraum: $\Omega = \{1, 2, 3, 4, 5, 6\}$

Alle Elementarereignisse sind gleich wahrscheinlich, d.h.

$$\begin{aligned} p(\{1\}) = p(\{2\}) = \ldots p(\{6\}) &:= a; \\ \text{wegen } p(\{1\}) + p(\{2\}) + \ldots + p(\{6\}) &= 1 \\ \Leftrightarrow 6 \cdot a &= 1 \\ \Leftrightarrow a &= \frac{1}{6} \end{aligned}$$

Sei $A = \{\text{gerade Augenzahl}\} = \{2, 4, 6\}$:

$$\begin{aligned} P(A) &= P(\{2, 4, 6\}) = P(\{2\} \cup \{4\} \cup \{6\}) \\ &= P(\{2\}) + P(\{4\}) + P(\{6\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2} \end{aligned}$$

Gefälschter Würfel

Die Axiomatik ist insbesondere nötig, um mit Situationen mit nicht gleich wahrscheinlichen Elementarereignissen rechnen zu können.
Betrachtet werde ein verfälschter Würfel mit

$$\begin{aligned} P(\{1\}) = P(\{3\}) = P(\{5\}) &= \frac{1}{12} \\ P(\{2\}) = P(\{4\}) = P(\{6\}) &= \frac{1}{4} \end{aligned}$$

Übung:
Gegeben seien ferner die Ereignisse

A	die Person wählt Partei	1 oder 3
B	"	4 oder 6
C	"	3 oder 4

und berechnen Sie $P(A)$, $P(B)$, $P(B \cap C)$ und $P(A \cup C)$ mit Hilfe der Axiome.

Grundlegendes zum Begriff „Wahrscheinlichkeit“

- Was ist eigentlich Wahrscheinlichkeit?
- Was bedeutet: „Mit Wahrscheinlichkeit $\frac{2}{3}$ wird es morgen regnen?“
- Stark gebräuchlich in der Umgangssprache als graduelle Abschwächung von Sicherheit („wahrscheinlich kommt Max“). Weg vom simplen Ja/Nein.
- Teilweise sogar quantifiziert:
  „Die Niederschlagswahrscheinlichkeit für morgen beträgt 30%“
- Medizinische Beipackzettel: „seltene Nebenwirkungen“

Literatur zum Begriff „Wahrscheinlichkeit“

Ausführlichere und weiterführende Literatur:

- Rohwer, G., Pötter, U. (2002): *Wahrscheinlichkeit, Begriff und Rhetorik in der Sozialforschung.* Juventa, Weinhein und München.
- Schneider, I. (Hg.) (1988): *Die Entwicklung der Wahrscheinlichkeit von den Anfängen bis 1933. Einführungen und Texte.* Wissenschaftliche Buchgesellschaft, Darmstadt.

Klassische Aspekte und Meilensteine

- Wahrscheinlichkeit im Glücksspiel, v.a. Würfelspiel: Profanisierung erst im Mittelalter, dort erst als Zufall gedeutet, vorher oft als Gottesurteil etc.
  - Cardano (1501-1576)
  - Gallilei (1546-1642)
  - Briefwechsel zwischen Pascal (1623-1662) und Fermat (1601-1665), erste systematische Wahrscheinlichkeitsrechnung: Lösung für Frage, wie Einsätze gerecht aufzuteilen sind, wenn Spiel unterbrochen wurde
  - Huygens (1629-1695)
- Wahr-schein-lichkeit (Prove-ability $\rightarrow$ probability)

Historische Wurzeln

- Mathematisierung von Glücksspiel
- als philosophischer/theologischer Begriff

- der Philosophie des Unsicheren und
- der Mathematik der Glücksspiele
- Jacob Bernoulli (1654 - 1705)

  Binomialverteilung.

  Theorem von Bernoulli: durch genügend große Versuchsreihen kann der Unterschied zwischen der relativen Häufigkeit eines Ereignisses und seiner Wahrscheinlichkeit beliebig gering gemacht werden.

Wahrscheinlichkeitsbegriffe

Objektivistisch // frequentistische Richtungen // aleatorische Wahrscheinlichkeiten

- Anschluss an die göttliche Ordnung
- Wahrscheinlichkeiten beschreiben tatsächlich vorhandene, zufällige Gesetzmäßigkeiten
- Objektbezogen: Wahrscheinlichkeit ist eine Eigenschaft des untersuchten Objekts (z.B. Würfel), objektiv $\longleftrightarrow$ objektbezogen (wie z.B. spezifisches Gewicht, Länge)
- Häufigkeitsinterpretation bzw. sogar -Definition Wahrscheinlichkeit als relative Häufigkeiten in „unendlich langen" reproduzierbaren Experimenten

Frequentistischer Wahrscheinlichkeitsbegriff I

R. von Mises (1883 - 1953):

„Die Wahrscheinlichkeit eines Ereignisses ist die langfristige relative Häufigkeit seines Auftretens“

Für ein Ereignis A:

$$P(A) = \lim_{n \to \infty} \frac{n_A}{n}$$

n_A : Anzahl der Erfolge
n : Anzahl der Versuche

Frequentistischer Wahrscheinlichkeitsbegriff II

Probleme bei der Definition:

- Einmalige Ereignisse
- Grenzwertdefinition
- Experimentdurchführung

Subjektivistische Richtungen

- Wahrscheinlichkeit hat ausschließlich mit Unsicherheit, nicht mit Zufälligkeit zu tun (Man kann auch über völlig deterministische Aspekte unsicher sein!)
- Wahrscheinlichkeit ist Eigenschaft des untersuchenden Subjekts $\Longrightarrow$ verschiedene Subjekte können durchaus zu unterschiedlichen Bewertungen kommen.
- Anwendung auch auf Aussagen. Bsp: Die Wahrscheinlichkeit, dass die Regierungskoalition die gesamte Legislaturperiode hält, ist...
- behaviouristischer Standpunkt: Wahrscheinlichkeiten äußern sich im Verhalten und können so gemessen werden
  z.B. bei Wetten
- Wichtig: subjektiv sind die Wahrscheinlichkeiten aber nicht die Rechenregeln.

Subjektiver Wahrscheinlichkeitsbegriff I

Laplace, Ramsey, de Finetti:

„Die Wahrscheinlichkeit eines Ereignisses ist der Grad der Überzeugung, mit der ein Beobachter aufgrund eines bestimmten Informationsstandes an das Eintreten eines Ereignisses glaubt“

$P(A)$ ist der Wetteinsatz in Euro, den eine Person höchstens einzugehen bereit ist, falls diese bei Eintreten von A einen Euro gewinnt.

Beispiele:

Münzwurf:	Einsatz auf „Zahl“ bis zu 0.5 € sinnvoll
Würfel:	Einsatz auf „5 oder 6“ bis zu 1/3 € sinnvoll

Subjektiver Wahrscheinlichkeitsbegriff II

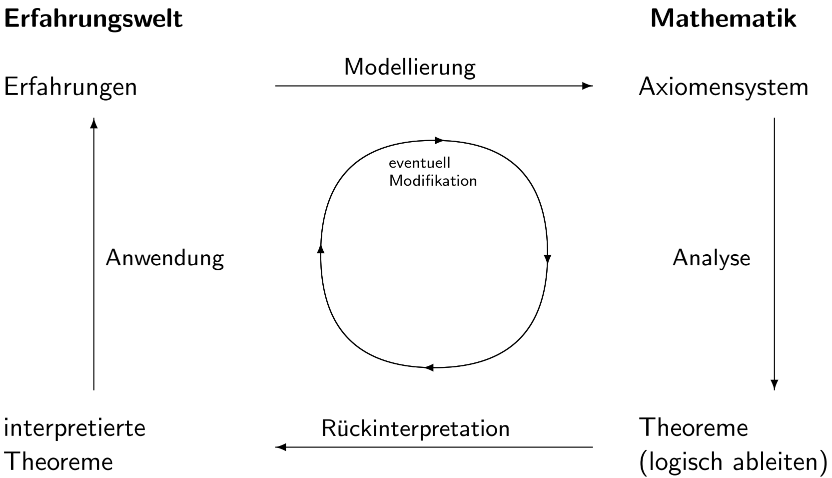
Probleme:

- subjektiv $=$ unwissenschaftlich ?
- Wettdefinition
- Informationsstand

Mathematisch-formaler Wahrscheinlichkeitsbegriff

- Axiomatik nach Kolmogoroff
- typische Anwendung der axiomatischen Methode:
  Axiom: Nicht bezweifelte Grundannahme für Kalkül
- Die Kolmogorffsche Axiomatik ist eine reine Definition, die sich zunächst im „luftleeren" Raum bewegt. Es wird rein formal festgelegt, was eine Wahrscheinlichkeit sein soll.
- Die Axiomatik ist *verträglich* sowohl mit der *Häufigkeits*- als auch mit der *Wettinterpretation*.
- Die Axiome von Kolmogoroff geben an, wie man mit Wahrscheinlichkeiten rechnet.
- Welche Phänomene man durch Wahrscheinlichkeiten beschreiben darf und wie die Ergebnisse zu interpretieren sind, ist aber damit nicht geklärt.

Die axiomatische Methode



Ausblick

- In der Tat gibt es auch Kritik an dieser Axiomatik: „zu streng und überpräzise“ $\longrightarrow$ aktueller Forschungsgegenstand (*Imprecise Probabilities, Intervallwahrscheinlichkeit*); hier nicht näher thematisiert: Kolmogoroff als absolute „Wahrheit“. Kritik:
  * Modellierung unsicheren (partiell widersprüchlichen, unvollständigen) Expertenwissens
  * Ökonomie: Entscheidungen unter komplexer Unsicherheit widersprechen Prognosen aus der üblichen Wahrscheinlichkeitsrechnung
- <u>logischer Wahrscheinlichkeitsbegriff</u>
  Wahrscheinlichkeit kommt Schlüssen zu: Wahrscheinlichkeit als logischer Grad mit dem aus einer Prämisse auf die Konklusion geschlossen werden darf (frühere Formalisierungsversuche gelten heute als gescheitert; aber Renaissance der Thematik)

Zur Kommunikation von Wahrscheinlichkeiten

Darstellung durch natürliche Häufigkeiten (nach Gigerenzer)

* „Superrepräsentative Stichprobe vorstellen", in der sich genau die Häufigkeitsverhältnisse in der Grundgesamtheit wiederfinden, z.B. 10 000 Personen
* Dann $P(A) = 0.1756$ vorstellen als: 1756 Personen haben die Eigenschaft A.

\+ einfachere Kommunikation von Wahrscheinlichkeiten und Risiken, reduziert Fehler beim Rechnen mit Wahrscheinlichkeiten Experimente mit Ärzten zeigen, dass die Darstellungsform Wahrscheinlichkeiten vs. natürliche Häufigkeiten) einen starken Einfluss auf die Korrektheit von Berechnungen hat.

– Gefahr der Verschleierung von Unsicherheit: die „natürlichen Häufigkeiten" sind zu *erwartende Durchschnittswerte*, wenn man sehr viele Stichproben hätte.

Beispiel: Beipackzettel

Angabe des Risikos von Nebenwirkungen auf Beipackzetteln:

sehr häufig:	mehr als 1 von 10 Behandelten
häufig:	weniger als 1 von 10, aber mehr als 1 von 100 Behandelten
gelegentlich:	weniger als 1 von 100, aber mehr als 1 von 1000 Behandelten
selten	weniger als 1 von 1000, aber mehr als 1 von 10000 Behandelten
sehr selten:	1 Fall oder weniger von 10000 Behandelten, einschließlich Einzelfälle

Welche Nebenwirkungen können bei der Anwendung von * Auftreten?**

Gelegentlich wurde über das Auftreten von Mundschleimhautentzündungen, Kopfschmerzen, Ohrengeräuschen berichtet.

Selten können auftreten: Beschwerden im Magen-Darm-Bereich (z.B. Sodbrennen, Übelkeit, Erbrechen oder Durchfall).

Beispiel: Lotto

- Beim Lotto ist die Wahrscheinlichkeit bei einem Spiel einen 6er zu bekommen:

$$\frac{1}{\binom{49}{6}} = \frac{1}{13983816} = 0.000000072$$

- „Einmal in 14 Millionen Spielen“
- „Einmal in 20.000 Jahren bei wöchentlichem Spielen“
- „Es ist wahrscheinlicher, den Tag der Ziehung nicht mehr zu erleben, als zu gewinnen“
- Simulationsexperiment

Risikodarstellung

Es gibt drei Arten der Beschreibung von Risiken für die menschliche Gesundheit:

- **Absolutes Risiko:**
  Angabe von Krankheitswahrscheinlichkeiten, jeweils getrennt für die Gruppe mit und ohne Risikofaktor
- **Relatives Risiko:**
  Verhältnis der Krankheitswahrscheinlichkeiten mit und ohne Risikofaktor
- **Anzahl der zusätzlich geschädigten Personen**
  (erwarteter Effekt)

Beispiel: WM Studie

- **Absolutes Risiko:**
  Ohne WM-Spiel: 14.6/1.34 Millionen = 1.09:100.000
  Mit WM-Spiel: 2.9:100000
  (Zum Vergleich: Tägliches Sterberisiko laut Sterbetafel für M, 50: 0.00460442/365 = 1.26/100.000)
- **Relatives Risiko:**
  Bei Spielen der deutschen Mannschaft um den Faktor 2.9/1.09 = 2.66 erhöht
  Konfidenzintervall (2.33 - 3.04)
- **Anzahl der zusätzlichen Fälle:**
  Gesamtfälle im Studiengebiet bei D-Spielen: 302
  Zu erwartende Fälle: 302/2.66 = 114
  Zusätzliche Fälle: 302 - 114 = 188
  Konfidenzintervall (172 - 203)

Hochrechnung auf Deutschland: 8000 - 11000

Beispiel: Wirkung von Pravastatin

„Menschen mit hohem Cholesterinspiegel können das Risiko eines erstmaligen Herzinfarkts sehr schnell um 22 Prozent vermindern, wenn sie einen häufig angewandten Wirkstoff namens Pravastatin einnehmen“

- Reduktion der Todesfälle von 41 auf 32 pro 1000 Patienten mit hohem Chorlesterin ($32 = 41 \cdot (1\text{-}0.22) = 41 \cdot 0.78$)
  Wahrscheinlichkeit für Todesfall: Reduktion von 4.1% auf 3.2%
  Absolute Risikodifferenz: 0.9%
- Reduktion um 22% (relatives Risiko 0.78) „22% werden gerettet“
- Es müssen 111 Patienten behandelt werden, um ein Menschenleben zu retten.
  Number needed to treat $=$ 1/Absolute Risikodifferenz $= 1/0.009 = 111.11$