

1.2.4 Die „induktive Brücke“ I

„Trivial oder genial, jedenfalls fundamental:“

Wir werden, mit wachsender Komplexität und Praxisbezogenheit immer wieder innehalten und uns mit der „Grundfrage der induktiven Statistik“ auseinandersetzen:

Warum und wie können wir aus Stichproben etwas über die dahinterstehende Grundgesamtheit lernen?

Man betrachte (und illustriere) dazu folgendes Beispiel:

Bsp. 1.15. [„Wahlbeispiel“]

Betrachtet werde ein Land, in dem die Wahlberechtigten die Wahl zwischen den Parteien Nr 1, Nr 2, ..., Nr 5 und Nr 6 (Nichtwähler) haben. Dabei entfallen auf die Parteien 2, 4, 6 jeweils 25% der Stimmen; die restlichen Stimmen verteilen sich gleichmäßig auf die Parteien 1, 3, 5. Seien f_1, \dots, f_6 die entsprechenden relativen Häufigkeiten

$$f_2 = f_4 = f_6 = \frac{1}{4} \quad , \quad f_1 = f_3 = f_5 = \frac{1 - 3 \cdot \frac{1}{4}}{3} = \frac{1}{12}$$

Es wird zufällig (im Sinne einer reinen Zufallsauswahl) eine Person ausgewählt und ihre Parteipräferenz ermittelt.

Geben Sie die sich ergebende Wahrscheinlichkeitsbewertung an!

Bem. 1.16. [Fortpflanzung relativer Häufigkeiten als Wahrscheinlichkeiten in der Stichprobe]

Gegeben sei eine Gesamtheit \mathcal{G} und ein Merkmal X mit den Ausprägungen a_1, \dots, a_k und der relativen Häufigkeitsverteilung f_1, \dots, f_k .

Zieht man eine „reine Zufallsauswahl“ vom Umfang 1 aus \mathcal{G} , hat also jedes Element aus \mathcal{G} die gleiche Wahrscheinlichkeit, gezogen zu werden, so gilt für die Ereignisse A_j : “die gezogene Person hat Merkmalsausprägung a_j “, $j = 1, \dots, k$:

$$P(A_j) = f_j$$

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 Beobachten der Stichprobe etwas über die Häufigkeitsverhältnisse in der Grundgesamtheit lernen.

1.2.5 Das Axiomensystem von Kolmogoroff und wichtige Rechenregeln

Warum reichen Laplace-Wahrscheinlichkeiten nicht?

Essentielle Voraussetzung bei der Festlegung von Wahrscheinlichkeiten nach Laplace:

Definition 1.17. [Axiome von Kolmogorov (1933)]

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

$$(K1) \quad P(A) \geq 0 \quad \text{für alle Ereignisse } A \subset \Omega.$$

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

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

Bem. 1.18.

- Die Axiome von Kolmogorov stellen zunächst eine reine Definition dar, die festlegt, was eine Wahrscheinlichkeit sein soll.
- Es gibt verschiedene Interpretationen, die die Axiomatik mit Leben füllen sollen und verschiedene Versuche, Wahrscheinlichkeiten operational zu definieren (also durch eine Messvorschrift).
- 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.
 - * Wir werden uns darum aber nicht dezidiert kümmern. (Nur daran denken, wenn man in etwas formale Bücher schaut.)

Rechenregeln für Wahrscheinlichkeiten

- $$P(\bar{A}) = 1 - P(A) \quad (1.2)$$

Beweis:

- Für nicht notwendigerweise disjunkte Mengen A, B gilt

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (1.3)$$

Veranschaulichung im Venn-Diagramm:

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

$$P\left(\bigcup_{i=1}^n A_i\right) = P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n) \quad (1.4)$$

Beweisidee:

- Als Spezialfall folgt, dass, sofern Ω endlich ist, die Wahrscheinlichkeitsbewertung durch die Bewertung auf den Elementarereignissen vollständig bestimmt ist:

$$P(A) = \sum_{\omega \in A} P(\{\omega\}) \quad (1.5)$$

Sind alle Elementarereignisse gleichwahrscheinlich ($P(\{\omega\}) \equiv \frac{1}{|\Omega|}$), so gilt $P(A) = \sum_{\omega \in A} P(\{\omega\}) = \frac{|A|}{|\Omega|}$; der Fall der Laplace-Wahrscheinlichkeit ist also als Spezialfall enthalten.

Bsp. 1.19. [Würfelwurf mit fairem Würfel]

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

Alle Elementarereignisse sind gleich wahrscheinlich, d.h.

$$\begin{aligned}
 P(\{1\}) = P(\{2\}) = \dots P(\{6\}) & \quad := \quad a; \\
 & \quad \text{wegen } P(\Omega) \stackrel{(K2)}{=} 1 \\
 \implies P(\{1\}) + P(\{2\}) + \dots + P(\{6\}) & \quad = \quad 1 \\
 \implies 6 \cdot a & \quad = \quad 1 \\
 \implies a & \quad = \quad \frac{1}{6}
 \end{aligned}$$

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

$$\begin{aligned}
 P(A) & \quad = \quad P(\{2, 4, 6\}) = \\
 & \quad =
 \end{aligned}$$

Alternativ hier möglich: Berechnung über Laplace-Wahrscheinlichkeiten

$$P(A) =$$

Vergleiche oben:

Die Axiomatik ist insbesondere nötig, um mit Situationen mit nicht gleichwahrscheinlichen Elementarereignissen rechnen zu können.

Bsp. 1.20.

Betrachtet werde der Wurf eines verfälschten Würfels mit

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

Gegeben seien ferner die Ereignisse

A der Würfel zeigt 1 oder 3

B der Würfel zeigt 4 oder 6

C der Würfel zeigt 3 oder 4

Berechnen Sie $P(A)$, $P(B)$, $P(B \cap C)$ und $P(A \cup C)$.

Prinzipiell kann man in diesem Beispiel zwar immer noch mit Laplace-Wahrscheinlichkeiten rechnen, indem man die Betrachtung auf gleich wahrscheinliche Fälle (hier Wähler) zurückführt. Dies erweist sich aber später als immer umständlicher beim Rechnen, und auch -bei möglichen Ziehungswahrscheinlichkeiten- im Detail verwirrend. Für das Verständnis ist diese Vorstellungsweise aber oft hilfreich, zumal man oft die Stichprobe so zieht, dass alle Einheiten gleich wahrscheinlich sind.

Bem. 1.21. [Vollständige Zerlegung:]

Wichtig für später ist auch die Möglichkeit, Wahrscheinlichkeiten zu bestimmen, indem man Ereignisse geeignet zerlegt und dann die Additivität der Wahrscheinlichkeiten disjunkter Ereignisse ausnützt.

Ist A_1, A_2, \dots, A_n eine *vollständige Zerlegung* von Ω , d.h. gilt

$$\bigcup_{i=1}^n A_i = \Omega \text{ und } A_i \cap A_j = \emptyset \text{ für } i \neq j,$$

so gilt für jedes Ereignis B :

$$P(B) = \sum_{i=1}^n P(B \cap A_i). \quad (1.6)$$

Diese Regel, die ein Spezialfall der allgemeinen Additionsformel für paarweise disjunkte Ereignisse ist (vgl. 1.4), ist für später sehr wichtig.² Als veranschaulichendes Beispiel denke man etwa an die zufällige Auswahl einer Person ω aus der Menge Ω aller wahlberechtigten Bundesbürger.

Sei	B	das Ereignis	ω ist für die Einführung der „Autobahn-Maut“
	A_1	”	” ω ist SPD-Wähler
	A_2	”	” ω ist CDU-Wähler
	\vdots		
	A_n	”	” ω ist Nichtwähler

A_1, \dots, A_n zerlegt Ω vollständig (paarweise disjunkt; $\bigcup_{i=1}^n A_i = \Omega$).

$B \cap A_i$ ist dann das Ereignis „ ω ist für die Einführung der „Autobahn-Maut“ und wählt Partei i “.

²Man vergleiche dies auch mit 1.5

1.2.6 Grundlegendes zum Begriff „Wahrscheinlichkeit“

- Was ist eigentlich „Wahrscheinlichkeit“?
- Was bedeutet: „Mit Wahrscheinlichkeit $\frac{2}{3}$ wird es morgen regnen?“

Ausführlichere und weiterführende Literatur:

- Rohwer, G., Pötter, U. (2002): *Wahrscheinlichkeit, Begriff und Rhetorik in der Sozialforschung*. Juventa, Weinheim und München.
- Schneider, I. (Hg.) (1988): *Die Entwicklung der Wahrscheinlichkeit von den Anfängen bis 1933. Einführungen und Texte*. Wissenschaftliche Buchgesellschaft, Darmstadt.
- Weichselberger, K. (2001): *Elementare Grundbegriffe einer allgemeineren Wahrscheinlichkeitsrechnung I. Intervallwahrscheinlichkeit als umfassendes Konzept*. Physika, Heidelberg S. 30-63.

a) **Wahrscheinlichkeit in der Alltagssprache**

- 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“

b) **Klassische Aspekte und Meilensteine**

- Wahrscheinlichkeit im Glücksspiel
- Wahr-schein-lichkeit (Prove-ability → probability)

Zwei historische Wurzeln

- Mathematisierung von Glücksspielen, v.a. Würfelspielen: Profanisierung erst im Mittelalter, dort erst als Zufall gedeutet, vorher oft als Gottesurteil etc.
- als philosophischer/theologischer Begriff: „Wahr-schein-lichkeit (Prove-ability → probability)“

c) Laplace'scher Wahrscheinlichkeitsbegriff (kombinatorischer Wahrscheinlichkeitsbegriff): siehe Kapitel 1.2.2

- Laplace (1749 - 1827)
- Aufbauend auf Symmetrieüberlegungen:
- Intuitiv einleuchtend, aber beschränkte Anwendbarkeit

- Von Laplace generell angewendet auf beliebige Unsicherheitssituationen (Sonnenaufgang, Jurisprudenz)
- Laplace strenger Determinist (Laplace'scher Dämon!): Unsicherheit entsteht allein durch unvollständiges und ungenaues Wissen über die Anfangsbedingungen!! → Subjektivismus

d) Aktuelle interpretatorische Hauptrichtungen / typische Wahrscheinlichkeitsbegriffe

1. 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“ als objektbezogen zu verstehen (wie z.B. spezifisches Gewicht, Länge)
- Häufigkeitsinterpretation bzw. sogar -Definition (von Mises, 1883 - 1953): Wahrscheinlichkeit als relative Häufigkeiten in „unendlich langen“ reproduzierbaren Experimenten

- Exkurs: natürliche Häufigkeiten (nach Gigerenzer, MPI für Bildungsforschung)
 - * Anmerkung: „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 . (Aber man weiß natürlich nicht, welche 1756 Personen.)
- + einfachere Kommunikation von Wahrscheinlichkeiten und Risiken, reduziert Fehler beim Rechnen mit Wahrscheinlichkeiten
„mathematisch triviale Umformung“, aber erfolgreich: Experimente mit Ärzten zeigen, dass die Darstellungsform Wahrscheinlichkeiten vs. natürliche Häufigkeiten einen starken Einfluss auf die Korrektheit von Berechnungen und Abschätzungen hat.

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

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).

– Anmerkung: Gefahr der Verschleierung von Unsicherheit: die „natürlichen Häufig-

keiten“ sind zu *erwartende Durchschnittswerte*, wenn man sehr viele Stichproben hätte, superrepräsentative Stichproben gibt es praktisch nicht. Individuelle Aussagen sind natürlich zufällig. Einer von zehn leiden an Krankheit A bedeutet natürlich nicht, dass nach 9 nicht an A leidenden Patienten zwingend der nächste krank ist.

Auf Beipackzettel wird typischerweise das Risiko von Nebenwirkungen so kommuniziert:

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

2. subjektivistische Richtungen

- Wahrscheinlichkeit hat ausschließlich mit Unsicherheit, nicht mit Zufälligkeit zu tun (vgl. Laplace) (Man kann auch über völlig deterministische Aspekte unsicher sein!)

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.

- „subjektiv“ als „subjektivbezogen“ zu verstehen. Wahrscheinlichkeitsbewertung ist Eigenschaft des untersuchenden Subjekts
⇒ 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 ⇒ Wettdefinition / -Interpretation
- Wahrscheinlichkeit von A gleich $P(A)$ bedeutet gedanklich:
- Wichtig: subjektiv sind die Bewertungen, nicht die Rechenregeln.
Kohärenzkriterien!

3. logischer Wahrscheinlichkeitsbegriff

Wahrscheinlichkeit kommt Schlüssen zu: Wahrscheinlichkeit als logischer Grad mit dem aus einer Prämisse auf die Konklusion geschlossen werden darf

e) („kalter“) mathematisch-formaler Wahrscheinlichkeitsbegriff: Axiomatik von Kolmogorov (siehe auch Kapitel 1.2.5)

- Die Kolmogorovsche Axiomatik ist eine reine Definition, die sich zunächst im „luftleeren“ Raum bewegt. Es wird rein formal festgelegt, was eine Wahrscheinlichkeit sein soll.

Es gab/gibt (wie eben ausgeführt), verschiedene Versuche, Wahrscheinlichkeiten operational zu definieren (also durch eine Messvorschrift) und verschiedene *Interpretationen*, die die Axiomatik mit Leben füllen (sollen).

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 eine Frage des Wahrscheinlichkeitsbegriffs.

- In der Tat gibt es auch Kritik an dieser Axiomatik: „zu streng und überpräzise“ → aktueller Forschungsgegenstand (*Imprecise Probabilities, Intervallwahrscheinlichkeit*); hier nicht näher thematisiert: für Statistik II Kolmogoroff als „absolute Wahrheit“.
Kritik:
 - * Modellierung unsicheren (partiell widersprüchlichen, unvollständigen) Expertenwissens
 - * Ökonomie: Entscheidungen unter komplexer Unsicherheit widersprechen Prognosen aus der üblichen Wahrscheinlichkeitsrechnung („Bounded rationality“)
 - * komplexere Erhebung in schwierigem Feld: wirklich jede Unsicherheit präzise mit Wahrscheinlichkeiten beschreibbar?