Stochastik

Barth, Friedrich

München, [20]03

17. 2. Stichproben

es handelt sich um das Ziehen einer Stichprobe – und *schließt* auf Grund des eingetretenen Ereignisses zurück auf den Anteil p der schwarzen Kugeln. Dabei unterscheidet man zwei Situationen.

1. Das Schätzproblem. Man hat keinerlei Vermutung über den Anteil p der schwarzen Kugeln in der Urne. In diesem Fall *schätzt* man den Anteil p auf Grund des eingetretenen Ereignisses (**Hochrechnung**). Man gibt als Schätzergebnis entweder einen einzigen Wert für p an (**Punktschätzung**) oder ein ganzes Intervall, in dem p liegen soll (**Intervallschätzung**). Das so abgegebene Urteil über den Anteil p ist mit einer gewissen Unsicherheit behaftet. Die Berechnung dieses Unsicherheitsgrades ist eine der wesentlichen Aufgaben der *Beurteilenden Statistik*.

2. Das Testproblem. Man hat von vornherein gewisse Vermutungen, Hypothesen genannt, über den Anteil p der schwarzen Kugeln in der Urne. Auf Grund des eingetretenen Ereignisses wird nun *entschieden*, welche dieser Hypothesen man beibehält oder verwirft. Auch hier ist es wesentlich, sich darüber klarzuwerden, mit welcher Sicherheit ein solches Urteil ausgesprochen werden kann.

Bevor wir uns diesen beiden Problemen zuwenden, wollen wir erst den Begriff der Stichprobe klären.

17.2. Stichproben

Die Grundlage aller Anwendungen der Stochastik ist die Möglichkeit, einen Versuch unter gleichen Bedingungen mehrmals zu wiederholen. Wollen wir z. B. über die Einkommensverteilung in einer Bevölkerung Ω etwas erfahren, so nützt es so gut wie nichts, wenn wir nur von einem zufällig ausgewählten Bürger ω das Einkommen wissen. Wir müssen eine Stichprobe von mehreren Personen ziehen. Dabei muß jede Person die gleiche Chance haben, in die Stichprobe aufgenommen zu werden. Man spricht dann von einer **Zufallsstichprobe***.
Nun sei X die Zufallsgröße »Einkommen der ausgewählten Person in DM«. Sie habe die Wertemenge $\mathfrak{S} := \{x_1, \ldots, x_s\}$ und die Wahrscheinlichkeitsverteilung W mit

$$W(x_j) = \frac{\text{Zahl der Personen mit } x_j \text{ DM Einkommen}}{\text{Zahl aller Personen}} \qquad \text{für } j = 1, \ldots, s\,.$$

Um über W etwas zu erfahren, wählen wir n-mal eine Person aus der Gesamtbevölkerung aus. Wir erhalten als Ergebnis ein n-Tupel von Zahlen, die sämtlich der Wertemenge $\mathfrak{S}$ von X angehören. Diese n Zahlen hängen vom Zufall ab. Wir haben es also mit n verschiedenen Zufallsgrößen $X_1, \ldots, X_n$ zu tun:

$X_i :=$ »Einkommen der i-ten ausgewählten Person in DM«

mit $i = 1, \ldots, n$. Die Wertemengen aller X_i stimmen mit der von X überein; die

* Das Wort *Stichprobe* entstammt der Bergmannssprache. Die alten Schmelzöfen wurden angestochen, um die Schmelze auf ihren Zustand zu prüfen.

X_i haben sogar die *gleiche* Verteilung W wie X, da ihre Werte – vom Experiment her gesehen – unter den gleichen Bedingungen angenommen werden. Außerdem sind die X_i insgesamt *unabhängig*. In die mathematische Theorie des Stichprobenziehens gehen allein diese Eigenschaften der X_i ein. Wir halten daher fest:

> **Definition 335.1:** Das n-Tupel $(X_1 | X_2 | \ldots | X_n)$ der Zufallsgrößen X_i heißt **Stichprobe der Länge n aus der Zufallsgröße** X, wenn gilt:
> **1.** Die X_i sind stochastisch unabhängig.
> **2.** Jedes X_i hat dieselbe Wahrscheinlichkeitsverteilung wie X.

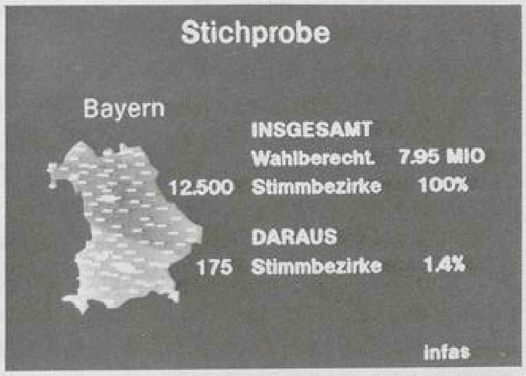

Bild 335.1 Landtagswahl in Bayern 1982 – ARD

Die Verwendung des Wortes »Stichprobe« ist in der Literatur nicht einheitlich. Oft nennt man auch das einzelne Werte-n-Tupel, das sich beim Stichprobenziehen ergibt, eine Stichprobe. Wir nennen gemäß Definition 335.1 das ganze Verfahren »Stichprobe«; das einzelne n-Tupel von Werten ist demgemäß **Stichprobenergebnis** zu nennen. Die Menge aller Stichprobenergebnisse kann als neuer Ergebnisraum genommen werden, der in diesem Zusammenhang auch manchmal **Stichprobenraum** (englisch: *sample space*) heißt. Ist $\mathfrak{S}$ die Wertemenge der Zufallsgröße X, so ist der Stichprobenraum das n-fache kartesische Produkt der Faktoren $\mathfrak{S}$, d. h.

$$\mathfrak{S}^n = \underbrace{\mathfrak{S} \times \mathfrak{S} \times \ldots \times \mathfrak{S}}_{n \text{ Faktoren}}$$

Beispiel: Die Zufallsgröße $X :=$ »Anzahl der Adler beim Wurf einer Münze« hat die Wertemenge $\mathfrak{S} = \{0; 1\}$. Eine Stichprobe der Länge n aus X ist das n-Tupel $(X_1 | X_2 | \ldots | X_n)$ mit $X_i :=$ »Anzahl der Adler beim i-ten Wurf«. Die X_i sind stochastisch unabhängig und besitzen dieselbe Wahrscheinlichkeitsverteilung wie X. Der Stichprobenraum $\mathfrak{S}^n$ ist die Menge aller n-Tupel, die aus 0 und 1 gebildet werden können.

Für verschiedene Fragestellungen der Praxis erweist es sich als zweckmäßig, auch dann noch von Stichproben zu sprechen, wenn die Zufallsgrößen $X_1, X_2, \ldots, X_n$ nicht mehr gleichverteilt oder nicht mehr stochastisch unabhängig sind. Das ist z. B. der Fall beim Ziehen *ohne Zurücklegen*. Man würde in unserem Beispiel alle n Personen auf einmal auswählen, so daß prinzipiell niemand die Chance hätte, zweimal gewählt zu werden. Die Zufallsgrößen $X_1, \ldots, X_n$ sind dann zwar noch gleichverteilt, aber nicht mehr unabhängig (siehe Aufgabe 223/**2**). Die für die Statistik wichtigen Formeln werden dadurch im allgemeinen komplizierter als beim Ziehen *mit Zurücklegen*. Glücklicherweise ist der Unterschied zwischen beiden Arten von Stichproben bei großen Grundgesamtheiten Ω (Bevölkerun-

gen, Urnen usw.) verschwindend gering, so daß es meist genügt, die einfachere Variante zu untersuchen.
Für den Statistiker ist die Frage wichtig, ob eine Stichprobe **repräsentativ** ist, d. h., ob sie eine genügend genaue Auskunft über die »Urne« geben kann, aus der sie stammt. Die Gewinnung einer repräsentativen Stichprobe gehört mit zu den schwierigsten Aufgaben der Beschreibenden Statistik. In der Markt- und Meinungsforschung nimmt man meist Stichproben der Größenordnung 2000.
Nach der Interpretationsregel für Wahrscheinlichkeiten und dem Gesetz der großen Zahlen kann man vermuten, daß genügend lange Stichproben im Sinne der Definition 335.1 auch repräsentativ sind. Eine genauere Auskunft hierüber gibt der 1933 von *Waleri Iwanowitsch Gliwenko* (1897–1940)* bewiesene

Hauptsatz der Mathematischen Statistik: Die mit Hilfe von Stichproben der Länge n gewonnenen empirischen Verteilungsfunktionen einer Zufallsgröße konvergieren mit Wahrscheinlichkeit 1 gleichmäßig gegen die wahre Verteilungsfunktion dieser Zufallsgröße, falls der Stichprobenumfang n gegen Unendlich strebt.

Dieser interessante Satz besagt also, daß die Aussagekraft einer Zufallsstichprobe von ihrer absoluten Länge abhängt und daß die Mächtigkeit der Grundgesamtheit, aus der sie gezogen wird, erstaunlicherweise keine Rolle spielt.

17.3. Test bei zwei einfachen Hypothesen

Das älteste Entscheidungsproblem der Wahrscheinlichkeitsrechnung ist *Blaise Pascal*s (1623–1662) *Infini-rien* (siehe Seite 343). Als ersten Test kann man *John Arbuthnot*s (1667–1735)** mathematischen Gottesbeweis *An Argument for Divine Providence, taken from the constant Regularity observ'd in the Births of both Sexes* aus dem Jahre 1710 auffassen. (Siehe Aufgabe 369/**32.**) Auch die Versuche von *Daniel Bernoulli* (1700–1782) aus dem Jahre 1770, das wahre Geburtsverhältnis der Geschlechter zu finden, kann man als Test im heutigen Sinne deuten. Die eigentliche Testtheorie entwickelten jedoch erst in der ersten Hälfte des 20. Jahrhunderts *Jerzy Neyman* (1894–1981) und *Egon Sharpe Pearson* (1895–1980). Beginnend mit einem einfachen Beispiel wollen wir ihre Gedanken nachvollziehen.

Beispiel 1: An eine Werkstatt werden Schachteln mit Schrauben geliefert. Ein Teil davon enthält Erste Qualität, das sind Schrauben, von denen nur 15% die vorgeschriebenen Maßtoleranzen nicht einhalten. Die restlichen Schachteln enthalten Zweite Qualität, mit einem Ausschußanteil von 40%. Die Lieferfirma hat vergessen, die Schachteln nach ihrem Inhalt zu kennzeichnen. Für die Verarbeitung ist es aber wichtig, die Qualität der Schrauben zu kennen. Man braucht also ein **Entscheidungsverfahren**, mit dessen Hilfe man die Schachteln der jeweiligen Qualität zuordnet. Über die Qualität der Schrauben gibt es nur 2 Vermutungen, **Hypothesen*** ** genannt, nämlich

* Гливенко
** *John Arbuthnot* war Leibarzt der Königin *Anna*. Als politischer Satiriker schuf er die Figur des John Bull (1712).
*** *ὑπόθεσις* = das Untergelegte; die Annahme.